Compact modeling and hybrid circuit design for spintronic devices based on current-induced switching

Yue Zhang

UNIVERSITÉ PARIS-SUD

ÉCOLE DOCTORALE
Sciences et Technologies de l'Information des Télécommunications et des Systèmes (STITS)

THESE EFFECTUEE AU SEIN DE
L'Institut d'Électronique Fondamentale (IEF), Univ. Paris-Sud/CNRS UMR 8622

DISCIPLINE
PHYSIQUE

THÈSE DE DOCTORAT

Soutenue le 11 juillet 2014
par

YUE ZHANG

MODÉLISATION COMPACTE ET CONCEPTION DE CIRCUIT HYBRIDE POUR LES DISPOSITIFS SPINTRONIQUES BASÉS SUR LA COMMUTATION INDUITE PAR LE COURANT

Directeur de thèse	Dafiné RAVELOSONA	Directeur de recherche, CNRS
Co-directeur de thèse	Weisheng ZHAO	Chargé de recherche, CNRS

Composition du jury :

Rapporteur	Ian O'CONNOR	Professeur, École Centrale de Lyon
Rapporteur	Julie GROLLIER	Chargée de recherche, HDR, Unité Mixte de Physique CNRS/Thales
Examinateur	Russell COWBURN	Professeur, University of Cambridge
Examinateur	Lionel TORRES	Professeur, Université Montpellier 2
Invité	Jacques-Olivier KLEIN	Professeur, Université Paris-Sud

Remerciements

Bien que cette thèse soit écrite en anglais, il me tient à cœur de rédiger cette partie en français.

D'abord, je souhaite adresser mes remerciements les plus sincères aux personnes qui m'ont apporté leur aide et qui ont contribué à l'élaboration de ce manuscrit. Les travaux présentés dans ce mémoire de thèse ont été réalisés à l'Institut Fondamentale Electronique (IEF), au sein de l'équipe NANOARCHI dans le département NANOELECTRONIQUE.

J'exprime toute ma reconnaissance à Monsieur Dafiné Ravelosona, directeur de recherche au Centre national de la recherche scientifique (CNRS), pour son accueil dans son département et pour avoir accepté de diriger cette thèse. Je tiens à le remercier pour son enthousiasme et son soutien sans faille.

Mes remerciements vont également à Monsieur Weisheng Zhao, chargé de recherche au CNRS, pour avoir effectué l'encadrement de ma thèse et pour avoir participé à mon jury. Nos échanges réguliers m'ont permis de progresser efficacement dans la conduite de ma thèse et j'ai beaucoup appris de ses conseils avisés.

Je souhaite remercier les membres du jury pour avoir consacré une partie de leur temps à ma thèse. Je remercie particulièrement Madame Julie Grollier, chargée de recherche (HDR) à l'Unité Mixte de Physique CNRS/Thales, et Monsieur Ian O'Connor, professeur à l'Ecole Centrale de Lyon, qui m'ont fait l'honneur d'accepter d'être rapporteurs de ma thèse. Je remercie également Monsieur Russell Cowburn, professeur à l'Université de Cambridge, et Monsieur Lionel Torres, professeur à l'Université Montpellier 2, pour avoir accepté d'examiner ma thèse.

Je tiens à adresser toute ma gratitude à Monsieur Jacques-Olivier Klein, professeur à l'Université Paris Sud et responsable de l'équipe NANOARCHI, pour ses conseils scientifiques et son aide au quotidien. Il est toujours gentil et a m'a donné de son temps pour m'aider à corriger mon manuscrit français.

Je remercie mes collèges de l'IEF, Claude Chappert, Sylvain Eimer (demi-chinois), Damien Querlioz, Djaafar Chabi, Nicolas Vernier, Thibaut Devolder, Joo-Von Kim, Jean-Paul Adam, Liza Herrera Diez, Karin Garcia et Felipe Garcia-Sanchez. Ce fût un réel plaisir de travailler, jouer au basket et sortir avec eux.

Je tiens à remercier chaleureusement toutes les personnes à l'Ambassade de Chine en France, qui m'ont soutenu pour travailler dans l'association UCECF en tant que président. Et tout particulièrement le ministre conseilleur Yansheng Ma et l'ancien ministre conseilleur Xiaoyu Zhu, avec qui j'ai eu une excellente expérience associative.

Naturellement, je remercie aussi tous mes amis français et chinois : Christelle Camus, qui m'a montré un autre Paris, Chunshi Guo, Ailun Wei, Ke Wang, Zhaohao Wang, Erya Deng, Na Lei, Weiwei Lin, Chenxing Deng et You Wang … Je suis très chanceux de vous avoir rencontrés.

Un grand merci à l'ensemble de ma famille, et surtout à mes parents Liguo Zhang et Huizhi Su (je suis si fier d'être votre fils), pour leur soutien et encouragements inconditionnels. Sans eux, cette thèse n'aurait pu toucher à sa fin.

Pour finir, une pensée pour mes grands-parents qui nous ont quittés au cours de ces trois ans. Cette thèse est pour eux.

Table of Contents

Résumé

La miniaturisation du nœud technologique de CMOS en dessous de 90 nm conduit à une forte consommation statique pour les mémoires et les circuits logiques, due aux courants de fuite de plus en plus importants. La spintronique, une technologie émergente, est d'un grand intérêt pour remédier à ce problème grâce à sa non-volatilité, sa grande vitesse d'accès et son intégration facile avec les procédés CMOS. Comparé à la commutation induite par le champ magnétique, le transfert de spin (STT), une approche de commutation induite par le courant, non seulement simplifie le processus de commutation mais aussi permet un fonctionnement sans précédent en termes de consommation et de vitesse. Cette thèse est consacrée à la modélisation compacte et la conception de circuit hybride pour les dispositifs spintroniques basés sur la commutation induite par le courant. La jonction tunnel magnétique (JTM), élément fondamental de la mémoire magnétique (MRAM), et la mémoire racetrack, nouveau concept fondé sur la propagation des parois de domaine induites par le courant, sont particulièrement étudiés. Ces dispositifs et circuits spintroniques sont basés sur les matériaux à anisotropie magnétique perpendiculaire (AMP) qui ouvrent la perspective d'une miniaturisation submicronique tout en conservant une grande stabilité thermique. De nombreux modèles physiques et paramètres réalistes sont intégrés dans la modélisation compacte pour obtenir une bonne cohérence avec les mesures expérimentales. En utilisant ces modèles compacts précis, certaines applications pour la logique et les mémoires magnétiques, tels que l'additionneur complet magnétique (ACM) et la mémoire adressable par contenu (CAM), sont conçues et simulées. Nous analysons et évaluons leur potentiel de performance en termes de surface, vitesse et consommation d'énergie par rapport aux circuits classiques. Enfin, afin de lutter contre la limitation de capacité entravant la large application, nous proposons deux optimisations de conception : la mémoire multivaluée (MLC) pour la STT-MRAM et l'assistance par champ magnétique pour la mémoire racetrack. Ce concept de MLC utilise le comportement stochastique des STT pour atteindre une haute vitesse tout en augmentant la densité de STT-MRAM. La mémoire racetrack assistée par champ magnétique est fondée sur l'observation d'une propagation des parois de domaine en dessous du courant critique, propagation est attribué à l'effet « Walker breakdown ». Ceci ouvre une nouvelle voie pour

réduire le courant de propagation et augmenter la capacité des mémoires racetrack au-delà des améliorations des circuits périphériques et des matériaux.

Mots clés : spintronique, jonction tunnel magnétique, mémoire racetrack, transfert de spin, anisotropie magnétique perpendiculaire, conception de circuit hybride spintronique/CMOS

Abstract

Title: Compact modeling and hybrid circuit design for spintronic devices based on current-induced switching

The shrinking of complementary metal oxide semiconductor (CMOS) fabrication node below 90 nm leads to high static power in memories and logic circuits due to the increasing leakage currents. Emerging spintronic technology is of great interest to overcome this issue thanks to its non-volatility, high access speed and easy integration with CMOS process. Spin transfer torque (STT), a current-induced switching approach, not only simplifies the switching process but also provides an unprecedented speed and power performances, compared with the field-induced switching. This thesis is dedicated to the compact modelling and hybrid circuit design for current-induced switching spintronic devices. Magnetic tunnel junction (MTJ), the basic element of magnetic random access memory (MRAM), and racetrack memory, a novel concept based on current-induced domain wall (CIDW) motion, are particularly investigated. These spintronic devices and circuits are based on the materials with perpendicular-magnetic-anisotropy (PMA) that promises the deep submicron miniaturization while keeping a high thermal stability. Numbers of physical models and realistic parameters are integrated in the compact modeling to achieve a good agreement with experimental measurements. By using these accurate compact models of PMA STT MTJ and PMA racetrack memory, some magnetic logic and memory applications, such as magnetic full adder (MFA) and content addressable memory (CAM), are designed and simulated. We analyze and assess their performance potential in terms of speed, area and power consumption compared with the conventional circuits. Finally, in order to tackle the capacity bottleneck hindering the wide application, we propose two design optimizations: MLC for MRAM and magnetic field assistance for racetrack memory. This MLC design benefits from the STT stochastic behavior to achieve an ultra-high speed while increasing the density. The racetrack memory with magnetic field assistance is based on the observation that CIDW motion can be triggered below the critical current due to “Walker breakdown” effect. This opens a new route to reduce the propagation current and increase the capacity of racetrack memory beyond the improvements of peripheral circuits or materials.

Keywords: spintronics, magnetic tunnel junction, racetrack memory, spin transfer toque, perpendicular magnetic anisotropy, hybrid spintronics/CMOS circuit design

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Motivation

The manipulation of the charge of electron has dominated the electronic world for over six decades. Thanks to its solid physical foundation, incalculable charge-based electronic devices have been designed or truly applied for our life. One of the most well-known stories is about complementary metal-oxide-semiconductor (CMOS) technology, which plays a predominant role for integrated circuits nowadays. This relatively mature technology has been widely used in not only digital but also analog integrated circuits, for example, microprocessors, static random access memory (SRAM), image sensors and data converters. The evolution of CMOS technology is commonly described by the famous Moore's law that was observed by Gordon Moore in 1965. It predicts that the number of transistors in integrated circuits doubles approximately every two years, which is, however, often quoted as 18 months afterwards. This prediction based on observation has continued for a long time, in turn, it has become a motive force to drive the researchers to innovate and develop the technologies.

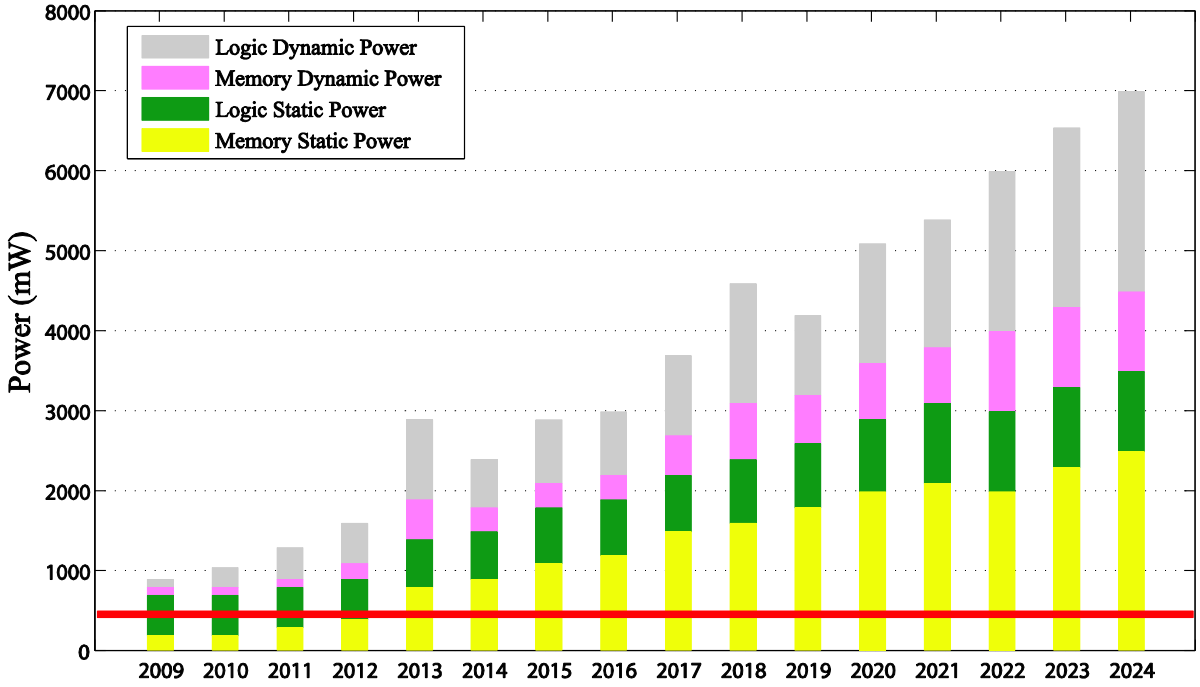


Figure 1.1 ITRS power consumption roadmap. (Red line: Requirement of dynamic plus static power)

However, with the shrinking of CMOS technology node below 90 nm, the growth of transistor number began to slow down. This is mainly due to the high static power dissipation in memory

and logic chip caused by the increasing leakage currents (see Figure 1.1). This power issue limits greatly the shrinking and improvement of electronic devices. For example, the design of multicore microprocessors for CPU in computer by many large semiconductor chips is a proof for this point.

In this background, novel technologies to replace the mainstream charge-based electronics are the hot topics for both academics and industries. Beyond the electrical charge, the spin freedom of electron attracts a broad attention and is considered to have a bright future. By being investigated for several decades, the spintronics was born from the discovery of Giant Magnetoresistance (GMR) effect and then rapidly developed. Figure 1.2 summarilizes the highlighted milestones for the development of the spintronics. The devices based on spintronics show the performance advantages in many aspects. The first one is low power. This is due to the non-volatility, which means that the information can be maintained without electrical power. With this feature, the system can be powered off in the idle state, which reduces greatly the standby power. Furthermore, spintronics can also allow reducing dynamic power that is normally caused by the large data traffic in the conventional Von-Neumann architecture. The possibility of spintronic devices to be 3D integrated above CMOS circuits at the back-end process can promise to significantly shorten the distance between memory and logic chip. Besides the power efficiency improvement, the potential advantages in terms of scalability and latency make the spintronic devices be used for various logic and memory applications. For example, spintronics has revolutionized ultra-high density Hard Disk Drives (HDDs) since the last 20 years.

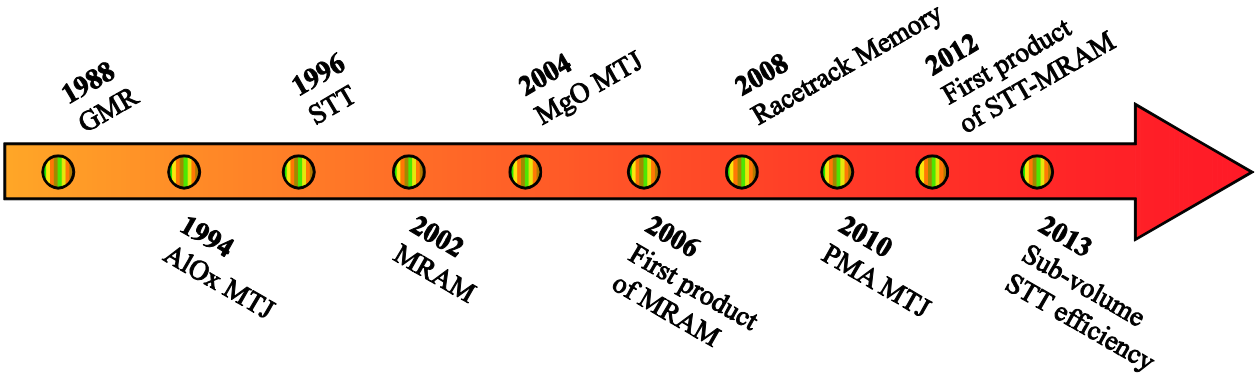


Figure 1.2 Milestones for the development of the spintronics.

Magnetic tunnel junction (MTJ), one of the most important spintronic devices, is the basic element of magnetoresistance random access memory (MRAM) which becomes a most promising candidate for the next generation of universal non-volatile memory. As a result of the tunnel magnetoresistance (TMR) effect, the MTJ resistance depends on the relative magnetization orientations of two ferromagnetic layers in MTJ. Much of the academic and industrial research efforts are presently focused on developing efficient strategies for switching magnetization in MTJs. One promising method relies on using spin transfer torque (STT), which involves low threshold currents and well-understood mechanisms. Furthermore, only a bi-directional current is needed in this approache, which simplifies greatly the CMOS switching circuits and thereby allows for higher density than the other approaches. However, some unexpected effects have been discovered using this approach in small MTJs (e.g. lateral size of 40 nm), such as erroneous state switching with reading currents and short retention times. These problems are mainly related to the in-plane magnetic anisotropy, which can not provide a sufficiently high energy barrier to ensure thermal stability. This issue limits greatly the potential for future miniaturization of MTJs. One compelling solution addressing this issue involves the perpendicular magnetic anisotropy (PMA) in certain materials (e.g. CoFeB/MgO), because it allows high energy barrier to be attained for small-size structure (< 40 nm) while maintaining the possibility of fast-speed operation, high TMR ratios and low threshold currents.

Racetrack memory is an emerging spintronic concept based on current-induced domain wall (CIDW) motion in magnetic nanowire. Combining with MTJs as write and read heads, CMOS integrability and fast data access speed can be achieved. In this concept, the data are stored via the magnetizations of magnetic domains separated by domain walls (DW). Due to STT mechanism, the DWs can be propagated consecutively in a direction by a spin current, which makes the racetrack memory possible to be widely applied for logic and memory designs. Considering the assets of PMA, the distance between the adjacent DWs can be extremely small, hence racetrack memory based on PMA materials is expected to provide ultra-high density.

Thanks to the diverse advantages demonstrated by spintronics as wall as various milestone breakthroughs of its related materials and techniques, hybrid spintronics/CMOS logic and memory circuits open a novel route to manipulate information more efficiently. Taking advantages of spintronic devices, the emerging circuits or systems can also realize low power,

high density and high speed. For the past decade, many spintronics based logic and memory circuits and their prototypes have been designed and presented. From the relatively mature spin valve for HDDs to recently commercialized STT-MRAM, from magnetic full adder (MFA) for magnetic processors to magnetic content addressable memory (CAM) for internet router and search engines, spintronics or concepts based on it has seeped into a majority of the advanced logic and memory systems.

However, spintronics is still on the way. There inevitably exist a lot of obstacles to challenge its feasibility and application potential. For example, although STT-MRAM has recently been commercialized, its density cannot be comparable to that of mainstream memories (e.g. dynamic random access memory (DRAM)). This density issue has become the main bottleneck to hinder the further development of STT-MRAM, and many approaches have been proposed to overcome it. Among them, multi-level cell (MLC) is a feasible scheme. Similarly, the high threshold current for propagating DWs limits the density of racetrack memory. As a large capacity racetrack memory requires a long magnetic nanowire, the high resistance of magnetic nanowire makes it difficult to provide an enough high propagation current. The optimizations at the levels of circuit and material have been proposed, however, none of them can offer a not only feasible but also reliable performance. Indeed, CIDW motions can be triggered below the threshold current due to the Walk breakdown, which has recently been discovered. This interesting counterintuitive phenomenon may be an opportunity to achieve a practically large capacity racetrack memory.

1.2 Contributions

In view of the foregoing research background and motivation, this thesis is focused on the integration of spintronic devices. From compact modeling to circuit design and optimization, the contributions of this thesis have been made at a series of levels. The main research contributions of this thesis are as follows:

- Compact modeling of PMA STT MTJ is firstly developed. By studying and synthesizing the related physical models, STT static, dynamic and stochastic behaviors in PMA material structures are integrated. This compact model can contribute to test and explore the performance of individual cell or hybrid MTJ/CMOS circuits based on PMA STT MTJ.

- Compact modeling of PMA racetrack memory is also developed. The physical model describing CIDW motion in PMA magnetic nanowire is embedded. By combining with PMA STT MTJ compact models as write and read heads, the full system of PMA racetrack memory is carried out. This compact model can be used extensively in various emerging logic and memory circuits, which strongly stimulates the development and application of racetrack memory.

- 1-bit MFA based on PMA STT MTJ and multi-bit MFA based on PMA racetrack memory are designed. Their performance advantages in terms of power, speed and density compared with the conventional designs indicate the prospectives for applying spintronic devices in computing logic applications.

- A novel design of CAM based on PMA racetrack memory is proposed. Taking advantage of 3D integration and sharing of CMOS writing and sensing circuits, this emerging design can achieve unprecedentedly high density and low power.

- A PMA STT-MRAM based on MLC is designed to envisage the density issue. This design utilizes innovatively the STT stochastic behavior, which is normally considered a negative side, to realize ultra-fast multi-level function. This design avails as well to improve the neuromorphic systems.

- PMA racetrack memory with magnetic field assistance is proposed to overcome the high threshold current issue. As mentioned above, this issue limits seriously the capacity potential of racetrack memory. Based on the recent observation involving DWs triggered by a low current under an external magnetic field, this design provides a new alternative for capacity improvement of racetrack memory and futher benefits for its feasibility and promotion.

1.3 Organization of the thesis

This thesis is organized into six chapters, including the introduction and the conclusion.

In Chapter 2, we introduce the state of the art concerning the work of this thesis, which includes the origine of the spintronics, the spintronic devices and the spintronic logic and memory circuits. The parts involving spintronic devices and circuits emphasize the MTJ and the racetrack memory that are the highlights of this thesis.

In Chapter 3, we present the compact modeling of PMA STT MTJ and PMA racetrack memory. The physical bases of them are elucidated in details, for example, STT static, dynamic and stochastic models for PMA STT MTJ as well as one-dimensional (1D) DW model for racetrack memory. The implementation of compact modeling in Verilog-A language is demonstrated. Diverse simulations (e.g. Direct Current (DC) simulation, transient simulation and statistical Monte-Carlo simulation) are carried out to validate the functionality of these compact models.

In Chapter 4, we focus on the hybrid spintronics/CMOS circuit design with emphysis of PMA STT MTJ and PMA racetrack memory based logic and memory circuits. Two basic circuits, i.e. writing and sensing circuits, are firstly presented. Their operations and performances are involved. By using the compact models of PMA STT MTJ and PMA racetrack memory, MFA and CAM are respectively proposed and studied. In comparison with the conventional circuits, the performance of these spintronic applications in terms of power, speed and density are discussed.

In Chapter 5, two design optimizations envisaged to overcome the density issues: MLC for STT-MRAM and magnetic-field assistance for racetrack memory, are successively proposed. From the structural implementation to the performance analyses, these optimized designs are thoroughly investigated and assessed.

In Chapter 6, the thesis will be globally concluded. Perspectives and future reseach directions will also be outlined.

The hierarchy of the work in this thesis can be depicted by Figure 1.3, which involves from the state variable to the architecture.

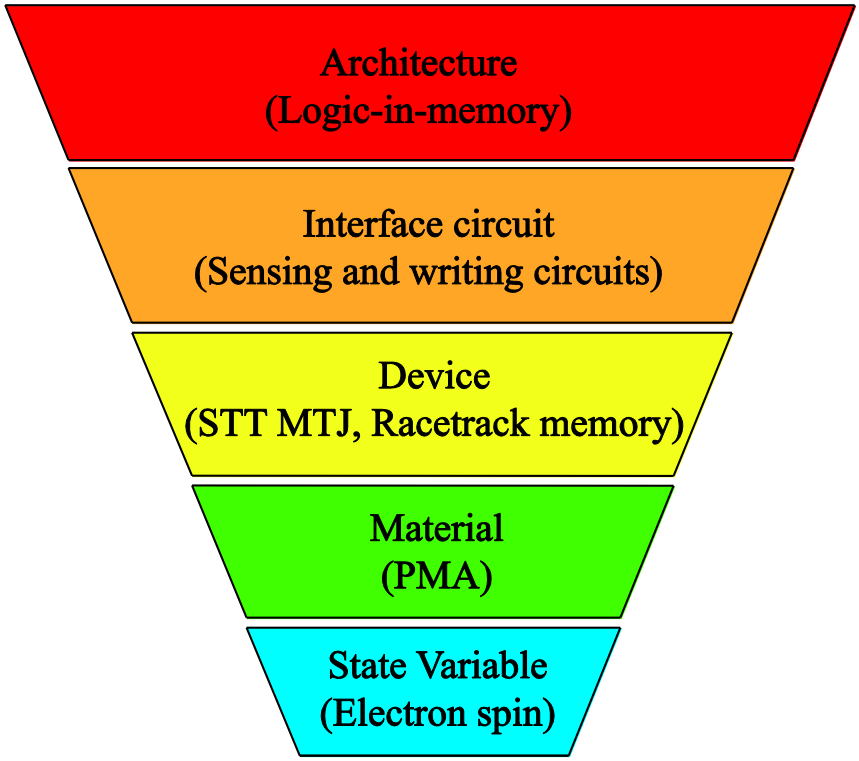


Figure 1.3 Hierarchical organization of the work in the thesis.

Chapter 2 State of the art

2.1 Spintronics

Spintronics (spin transport electronics or spin electronics) is an emerging interdiscipline whose principal idea is to control the electrons' spin degrees of freedom for electronics. It is considered a potentially promising technology to replace the mainstream charge-based electronics [1].

The spin of the Electron has been known by public for almost 100 years. Although its theoretical and experimental studies were developed by many scientists and it was proven to play the key role in several areas of condensed matter physics, it was rarely involved with the conventional electronics based on charge of electrons [2]. The story had a twist in 1988 due to the Giant MagnetoResistance (GMR) effect that was respectively discovered by Albert Fert in France and Peter Grünberg in Germany [3-4]. GMR effect was observed in thin ferromagnetic structures, which is because: in paramagnets, the numbers of up-spin (majority) and down-spin (minority) electrons are the same, or it is to say that their spin direction is random; in ferromagnets, with the difference between the numbers of up-spin and down-spin electrons, the magnetism can influence the electrical transport, in turn the process can be reversed.

This outstanding research discovery is afterwards considered a milestone in physics, especially in condensed matter domain. Because of this, the Nobel Prize in Physics was awarded to both of these physicists in 2007. This discovery is also deemed as the "birth" of spintronics.

A fundamental understanding of spintronics can be explained by three main aspects: spin polarization, spin relaxation and magnetoresistance [5]. Firstly, in electrical spin injection, spin polarization is defined by the imbalance of spin population. As a current passing through a magnetic material can be polarized by the local magnetic moment, it will lead to the nonequilibrium spin accumulations [6]. Secondly, when the current carrying the spin-polarized electrons reaches and passes another magnetic material, spin relaxation will take place, which drives the spin population back to equilibrium. The length for situation wherein the spin keeps its direction is called the spin diffusion length [7]. A material parameter named damping constant is usually used to describe the speed of the spin relaxation. Thirdly, similar to the electrical resistance, there is a magnetoresistance for magnetic material. As it is found to be sensible to the nonequilibrium spin population, its changes are thus measured to detect the spin states.

Benefiting from this phenomenon, spintronic devices, not merely based on GMR effect, offer the possibility to integrate with hybrid process.

In details, GMR effect can be described by two channel model as shown in Figure 2.1. In the simplest form, two ferromagnetic (FM) layers are separated by a non-magnetic (NM) metal layer. When the magnetization directions of two FM layers are parallel, the spin-up electrons can pass through the structure nearly without scattering, providing a relatively low resistance. Contrarily, in the anti-parallel configuration, both spin-up and spin-down electrons will suffer from the scattering, which leads to a higher resistance. In some multilayers structure, the difference between the resistances in parallel and anti-parallel states can reach 100%, which is the reason why it is named by giant magnetoresistance. This outstanding effect soon attracts intense research attention. The concept "spin valve" is one of the most important applications based on GMR effect, which has been widely used in the hard disk drives (HDDs) as read heads [8]. At the beginning, the "current in plane" (CIP) spin valve geometry was mainstream. The "current perpendicular to plane" (CPP) was afterwards found to be more convenient for the downscaling, however more difficult to fabricate [9]. From 1997 to 2003, HDD's area density has been increased by three orders of magnitude (from ~0.1 to 100 Gbit/in^2) [10].

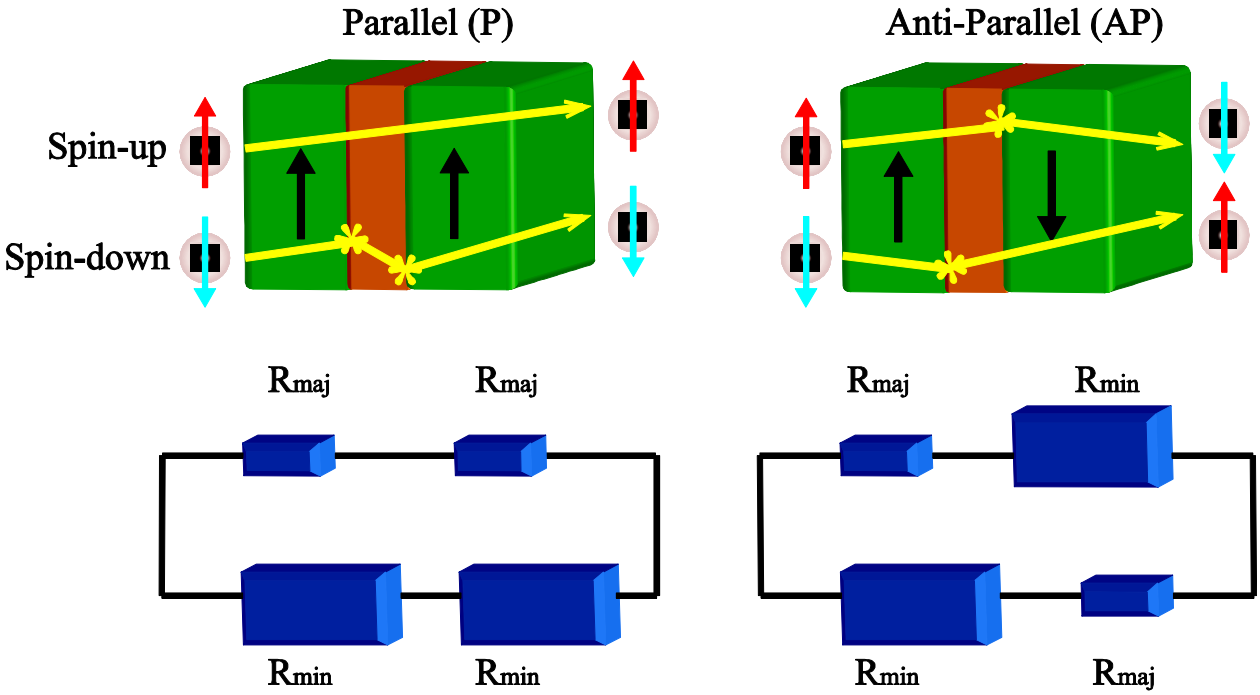


Figure 2.1 Two channel model to describe GMR effect.

2.2 Spintronic devices

Beyond the spin valve mentioned above, there are a number of devices based on spintronics. Among them, magnetic tunnel junction (MTJ) and magnetic domain wall (DW) nanowire are no doubt the most promising ones thanks to their mature technologies and high performances. In this thesis, we will mainly focus on these two emerging spintronic devices and investigate their integrations in hybrid logic and memory circuits.

2.2.1 Magnetic tunnel junction (MTJ)

2.2.1.1 MTJ structure

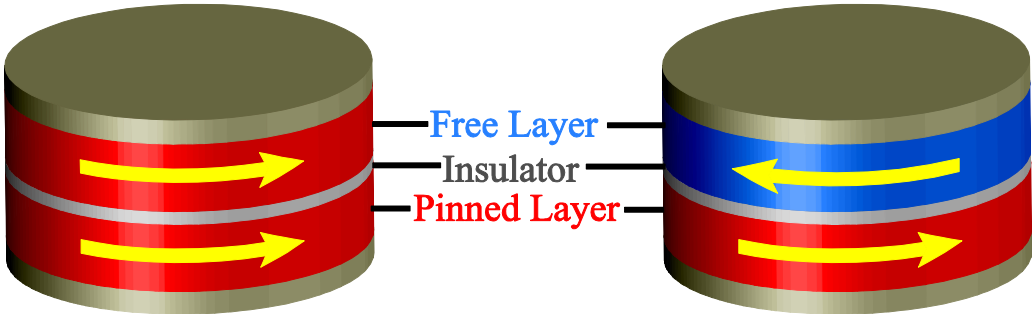


Figure 2.2 MTJ structure

By replacing the non-magnetic metal layer by a thin insulating (e.g. AlxOy and MgO) layer, the magnetic tunnel junction (MTJ) is created as shown in Figure 2.2 [11-13]. In the inchoate types of MTJ, two FM layers have different coercivities (defined as the magnetic field to switch the magnetization). Then in the recent and mainstreams types of MTJs, one of the FM layers is pinned layer whose magnetization is fixed in a specific direction and the other is free layer whose magnetization can be tuned. Nowadays, it is considered as the basic element of magnetoresistance random access memory (MRAM) and provides a great impetus for electronics development.

With respect to the spin valve based on GMR effect, MTJ is related to Tunnel MagnetoResistance (TMR) effect. The spin-polarized electrons pass from one FM to the other by a tunnel effect with the conservation of spin direction (see Figure 2.3). Actually, this effect was observed and proposed by Jullière as early as 1975 [14]. By analyzing the phenomenon of

tunneling conductance, he summarized a physical model to describe the change of conductance between parallel (P) and anti-parallel (AP) magnetizations in two FM layers. The related tunneling magnetoresistance ratio can be defined as

$$TMR = \frac{\Delta R}{R_P} = \frac{R_{AP} - R_P}{R_P} = \frac{G_P - G_{AP}}{G_{AP}} \tag{2.1}$$

where R_{AP} and R_P are the resistances of anti-parallel and parallel states of MTJ, which have the inverse relationship with the conductances G_{AP} and G_P. The expressions of conductance and spin polarization are given by

$$G_P = N_{M1}N_{M2} + N_{m1}N_{m2} \tag{2.2}$$

$$G_{AP} = N_{M1}N_{m2} + N_{m1}N_{M2} \tag{2.3}$$

$$P_i = \frac{(N_{Mi} - N_{mi})}{(N_{Mi} + N_{mi})} \tag{2.4}$$

where N_{Mi} and N_{mi} are the effective densities of states of majority and minority electrons at the Fermi energy in both magnetic layers. As a result, the TMR ratio is expressed in terms of the spin polarization by

$$TMR = \frac{2P_1P_2}{1 - P_1P_2} \tag{2.5}$$

Unfortunately, caused by the technical limitations, the experiment of Jullière was difficult to reproduce and the progress of TMR had been impeded for almost 20 years. The breakthrough occurred in 1994 by using amorphous Al_2O_3 as the tunneling barrier (insulating layer) to realize the room temperature magnetic tunneling transport [12-13]. By optimizing the material and fabrication condition, the TMR ratio of this structure can reach up to 70%. Although this value is already much larger than GMR in spin valve, it is still far away from the requirement for spintronic applications, for example, high density MRAM requires at least 150% TMR at room temperature. Another considerable leap of MTJ is using a single-crystal MgO tunnel barrier that can provide even larger TMR, which is sometime called the giant TMR effect [15-16]. So far, the

TMR ratio record of MgO based MTJ can reach as high as 600% at room temperature [17]. These results are of great importance not only to avoid the CMOS process mismatch and parameter variation, but also to miniaturize the sense amplifier circuit area [18].

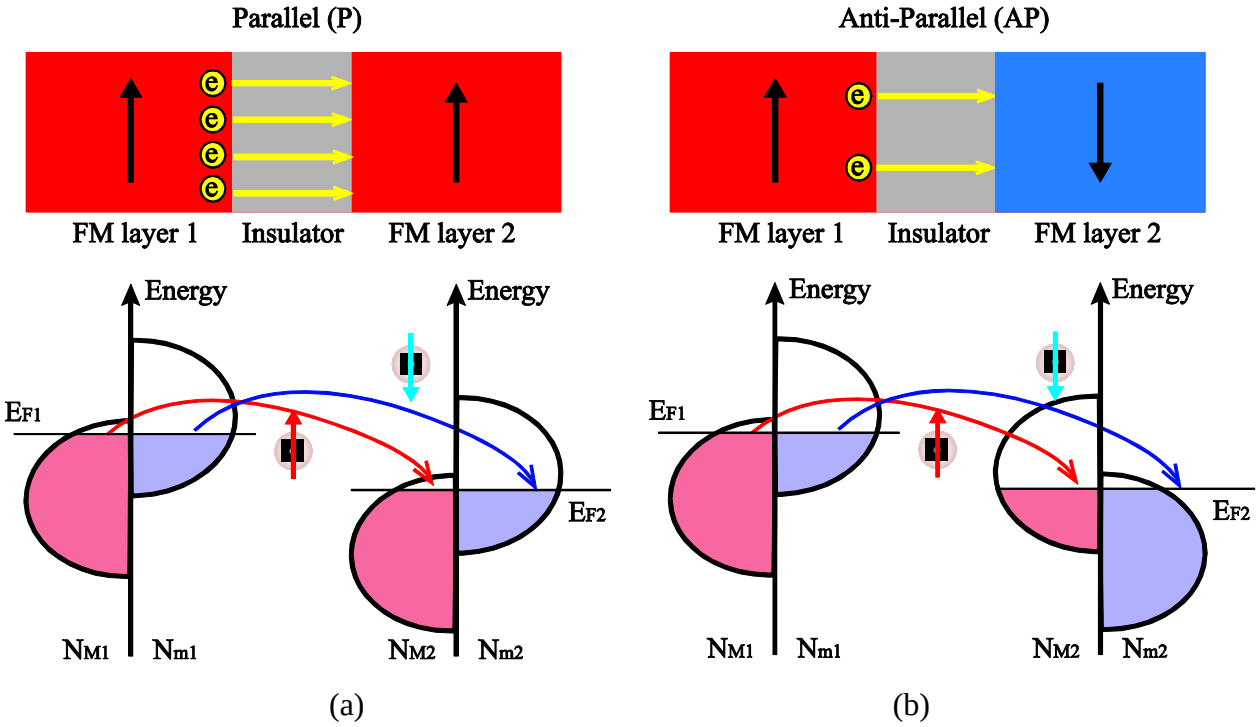


Figure 2.3 Schematic illustration of electron tunneling in MTJ. (a) Magnetizations in two FM layers are parallel (P state). (b) Magnetizations in two FM layers are anti-parallel (AP state).

2.2.1.2 Switching approaches for MTJ

2.2.1.2.1 Field-induced magnetic switching (FIMS)

Field-induced magnetic switching (FIMS) is a field-only switching approach as shown in Figure 2.4 [19]. Due to its typicality and maturity, the first generation of MRAM was built based on this switching approach. In such mechanism, the magnetic fields are generated by two orthogonal current lines, for example, Iw works as the word line and Ib works as the bit line to switch the magnetization of free layer. This structure has an advantage in sensing, which is totally independent with the writing lines. However, the write selectivity based on the combination of two perpendicular pulses of magnetic fields, i.e. Hb and Hw, may result in narrow write margin and half-selectivity issues. Furthermore, the high currents (>10 mA) to generate magnetic fields

yield considerable power consumption meanwhile the electromigration effect limits its scalability. These issues hinder its commercialization. Thanks to the toggle switching patterned by Freescale, MRAM based on this advanced switching method was commercialized in 2006 [20]. However this approach cannot yet overcome the drawbacks of speed, density and power caused by using magnetic field for switching.

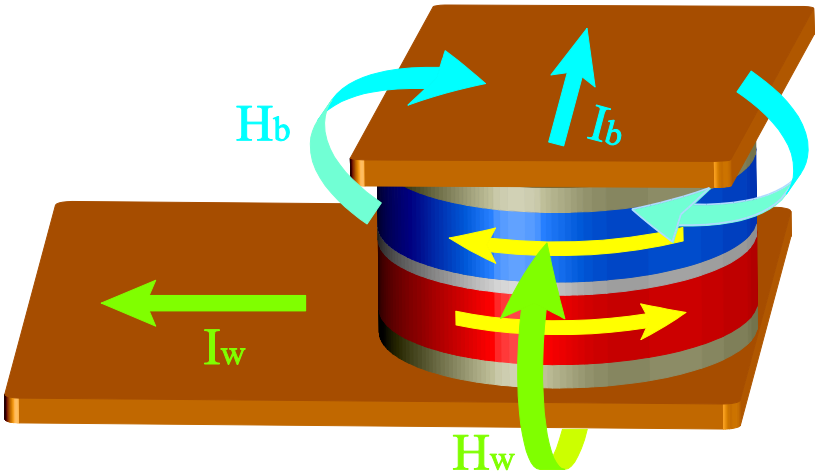


Figure 2.4 FIMS approach structure.

2.2.1.2.2 Thermally assisted switching (TAS)

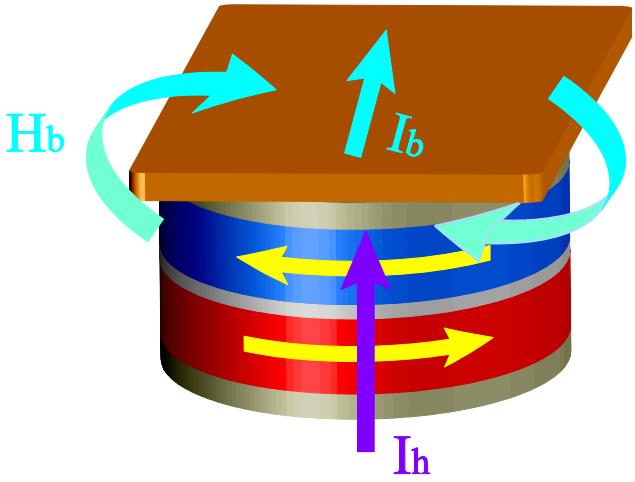


Figure 2.5 TAS approach structure.

To improve the performances of write selectivity, power consumption and thermal stability, thermally assisted switching (TAS) was proposed (see Figure 2.5) [21-22]. It is worthy noting that an additional anti-ferromagnetic (AFM) layer is normally added above the free layer to pin

the free layer at the standby temperature. Its basic principle is that a current flowing through MTJ heats the magnetic layers above their magnetic ordering temperature to reduce greatly the required switching field [23]. Similar to the FIMS structure, two orthogonal current lines are installed to achieve write selectivity; in contrast, one (Ih) is used to heat the MTJ and the other (Ib) is used to generate the switching field. This approach promises relatively lower power, higher density and higher thermal stability comparing with the pioneering FIMS approach; however, the mandatory heating and cooling processes lower the operation speed, which makes TAS approach not competitive in the high speed logic applications.

2.2.1.2.3 Spin transfer torque (STT)

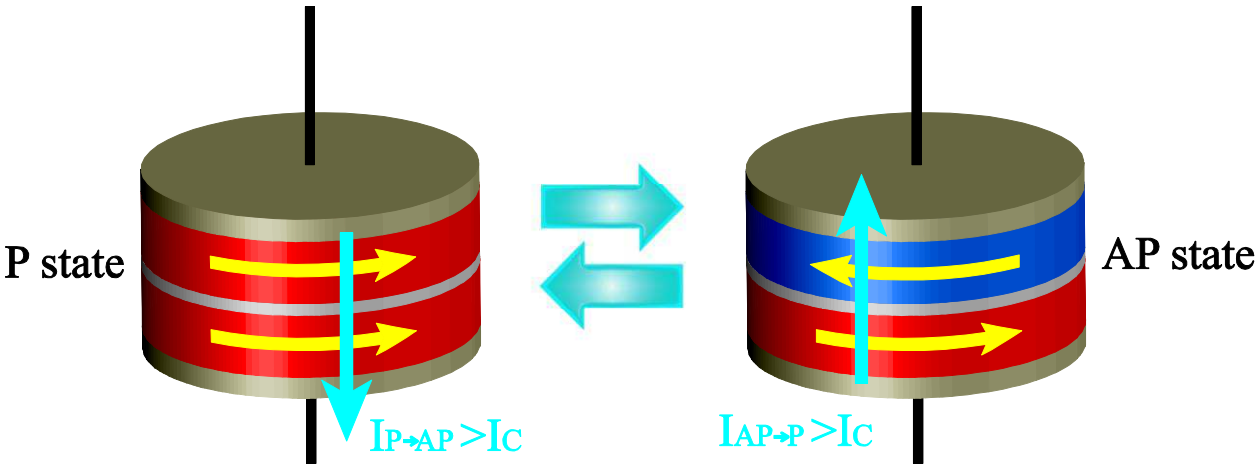


Figure 2.6 STT switching approach structure.

While the relative direction of magnetizations of two FM layers affects the flow of spin-polarized current in GMR and TMR structures, a reverse effect, called spin transfer torque (STT), was proposed independently by Berger and Slonczewski in 1996 [24-25]. They found that a spin-polarized current injected perpendicularly to the plane could equally influence the magnetizations. This interaction is attributed to angular momentum transferred from the polarized electrons to the local magnetization of the FM layer. Once the amount of electrons exceeding the threshold value (often represented by critical current or critical current density), the STT exerted by the current will switch the magnetization of the free layer of MTJ [26-27]. The STT switching approach was initially researched on GMR based spin valve [28-29], and then focused on the MTJ providing a significantly higher magnetoreisistance [30-31]. In MTJ, one FM layer acts as a polarizer for an electric current, which then transfers angular momentum by exerting a torque on the

magnetization of the other FM layer. This current-only approach simplifies greatly the switching process as it only requires a bi-directional current. Moreover, the magnitude of current for STT is normally less by an order than that for generating a large magnetic field. As a result, STT switching approach is widely considered the most promising one to be applied in the future MRAM applications, and we will concentrate to investigate the spintronic devices based on this switching approach in this thesis [32-33].

2.2.1.2.4 Thermally assisted spin transfer torque (TAS+STT)

Combining the aforementioned advantages of TAS approach and STT approach, thermally assisted spin transfer torque (TAS+STT) switching approach was proposed [34-35]. Like STT switching approach, only one polarized current flows through the MTJ to heat and switch simultaneously the free layer. When the temperature of MTJ is risen above the blocking temperature of the AFM layer associated to the free layer and the current exceeds the threshold value of STT, the magnetization of free layer will be switched. This combined approach can provide a good tradeoff among density, power and thermal stability, however, it still requires the supplementary time for cooling and power for heating, which is adverse to achieve high-speed low-power applications.

2.2.1.3 Perpendicular magnetic anisotropy (PMA)

To address the requirement of high-performance MTJ for the future logic and memory applications, there are usually five criteria to evaluate: small area, high TMR ratio, low STT switching current, capacity to withstand the standard semiconductor processing and high thermal stability. With the shrinking of size, the conventional MTJ with in-plane magnetic anisotropy becomes more and more difficult to satisfy these criteria. Recent material progress showed that the MTJ with perpendicular magnetic anisotropy (PMA) could offer lower switching critical current, higher switching speed and higher thermal stability compared with that with in-plane magnetic anisotropy [36]. These can be explained by the following theories.

The barrier energy and critical current of STT switching in the materials with in-plane magnetic anisotropy can be expressed as:

$$E_i = \frac{\mu_0 M_S \times Vol \times H_C}{2} \tag{2.6}$$

$$I_{C0i} = \alpha \frac{\gamma e}{\mu_B g}(\mu_0 M_S)(H_{ext} \pm H_{ani} \pm \frac{H_d}{2})Vol \tag{2.7}$$

where H_C is the coercive field, H_{ext} is the external field, H_{ani} is the in-plane uniaxial magnetic anisotropy field, H_d is the out-of-plane magnetic anisotropy induced by the demagnetization field, μ_0 is the permeability in the free space, M_S is the saturation magnetization, *Vol* is the volume of the free layer, μ_B is the Bohr magneton, γ is the gyromagnetic ratio, *e* is the electron charge, *m* is the electron mass.

On the other hand, the barrier energy and cirtical current in materilas with PMA are given as:

$$E_p = \frac{\mu_0 M_S \times Vol \times H_K}{2} \tag{2.8}$$

$$I_{C0p} = \alpha \frac{\gamma e}{\mu_B g}(\mu_0 M_S) H_K Vol \tag{2.9}$$

where H_K is the perpendicular magnetic anisotropy field.

By comparing Eq. 2.6 and Eq. 2.8, as H_K is higher than H_C, PMA allows obtaining relatively high barrier energy with a small size. By comparing Eq. 2.7 and Eq. 2.9, as H_K is much lower than H_d, the critical current for PMA materials can be significantly reduced.

From 2002, when the first MTJ with PMA was reported, this advantageous structure attracts a great deal of attentions from academics and industries [37]. A variety of material systems has been attempted, for example, rare-earth/transition metal alloys, multilayers and other alloy materials. However, they have not been able to truly realize low critical current and high thermal stability at the same time. This situation didn't change until the Ta/CoFeB/MgO structure was revealed in 2010 [38-39]. Figure 2.7 demonstrates the excellent performances of this structure. It takes advantages of CoFeB-MgO interface anisotropy to provide a good tradeoff among the area (40 nm), critical current (~50 µA), thermal stability (40 k_BT) and TMR ratio (>100%).

Thanks to the material and technical improvement of MTJ, especially MgO based MTJ with PMA, a lot of persistent and intensive efforts have been made for the past years to develop the

high performance MTJ based systems, such as MRAM [40-43] that will be introduced in the following sections.

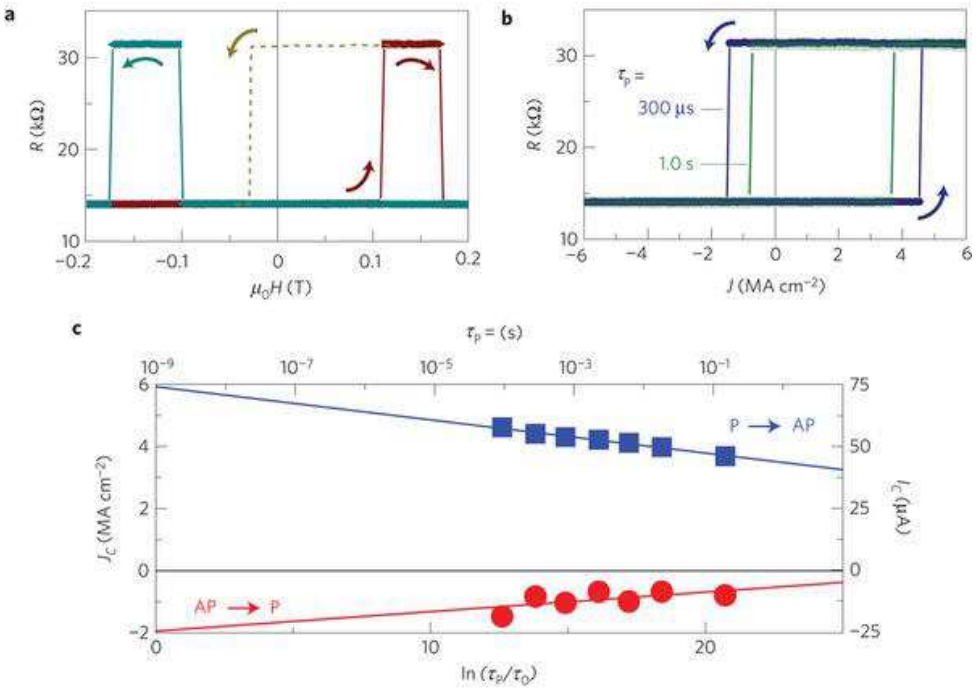


Figure 2.7 TMR and current-induced magnetization switching for Ta/CoFeB/MgO structure MTJ with PMA. (a) Perpendicular R-H curve. (b) Typical results of current-induced magnetization switching at current pulse duration of 300 µs and 1.0 s. (c) Critical current density as a function of pulse duration. [38]

2.2.1.4 Recent progress of MTJ

With the development of materials and techniques, there was a lot of progress for MTJ appearing during this thesis work. Here we mention briefly the most important and promising three of them, that are high spin torque efficiency in sub-volume regime, spin hall effect (SHE) and voltage-induced effect.

Along with the downscaling pace of MTJs beyond sub-volume limit (~40 nm), MTJ displays a relatively high thermal stability factor and low STT critical current. This so-called "high spin torque efficiency" is a strong stimulus for high density MTJ application. Spin torque efficiency is defined here as [188]:

$$\kappa = \frac{E}{I_{c0}} \tag{2.10}$$

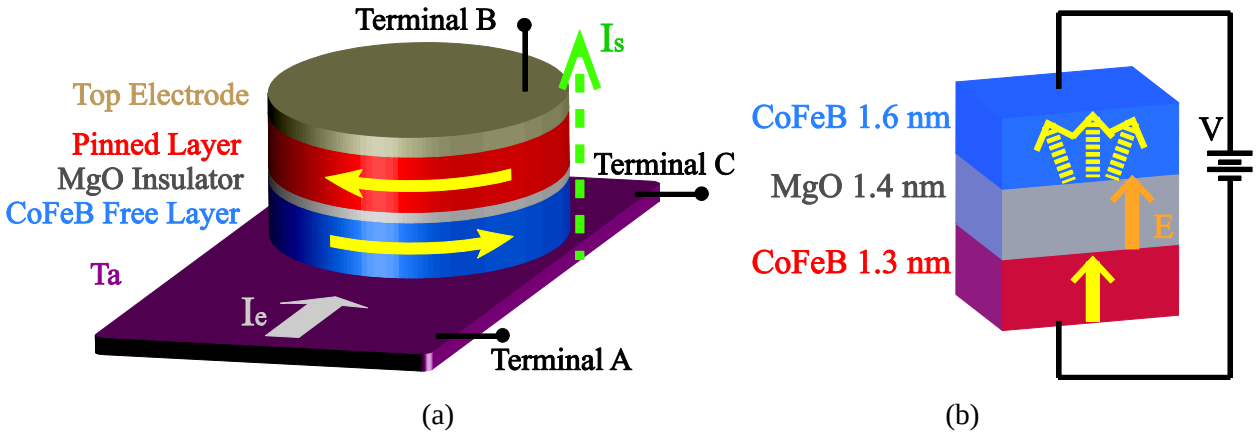
where E is barrier height (or thermal stability factor), and I_{c0} is the average critical current. Spin torque efficiency reflects the capability of spin polarized current to reverse the barrier height. Practically, when the lateral size of MTJ sacling down to the sub-volume limite, sub-volume activation effects make the leading term guiding the magnetization switch in devices, which is negligible when lateral diameter is larger than the limit. Thanks to this effect, the scaling gain (faster operation, higher density and improved spin torque efficiency) can be further continued, which benefits greatly for the miniaturization of MTJ.

Beyond the STT switching mechanism, it has been discovered that a spin-polarized current can be generated by the SHE [44]. Due to the spin-orbit coupling, the electrons with different spins deflect in different directions. However, this effect was usually too modest to limit its application. Recently, it was reported that a giant SHE in a high-resistivity from of tantalum (β-Ta) could generate a spin current strong enough to induce the switching of MTJ [45]. Based on this prominent phenomenon, a three-terminal SHE device was proposed as shown in Figure 2.8(a). The electric current flowing horizontally induces a spin current to pass vertically through the in-plane MTJ structure. As the spin polarization of spin current is governed by the direction of the electric current, the magnetization switching direction of MTJ depends thus on the sign of electric current.

Although this three-terminal device would cause area efficiency degradation compared with the conventional MTJ, it exhibits various assets in many aspects. For example, by optimizing the thickness of Ta layer, the switching current can be decreased by nearly one order of magnitude compared with STT switching mechanism. The reduction of switching current and lower resistance of Ta layer can also lead to an advantageous switching energy. In addition, two-terminal MTJ always suffers from the reliability issue due to different resistance requirements for writing and sensing operations: low resistance is expected for writing operation and high resistance is more suitable for sensing operation. Three-terminal device separates the writing and sensing operations structurally, solving properly the reliability challenges. Thanks to these assets, this SHE based MTJ structure is regarded to have great potential towards the future magnetic memory and non-volatile logic design.

From the aforementioned introduction, we can find that the magnetization switching of MTJ has generally been carried out by either magnetic field or spin-polarized current. As the generation of

magnetic field is also implemented by the electric current, we can say that both of the above switching mechanisms are based on current, and then consuming energy. Beyond the spin-polarized current, voltage-induced effects provide another possible alternative to manipulate the magnetization [46-47]. It was recently observed that the interfacial perpendicular anisotropy can be controlled by the applied bias voltage. This is due to spin-orbit coupling which makes the surface anisotropy be modulated by the accumulated charge near the interface. From the point of view of implementation, there are two voltage-based switching mechanisms that have recently been exhibited. The first one is applying an ultra-fast voltage pulse to result in temporal change of magnetic anisotropy [48]. A toggle magnetization switching process under a constant bias field is realized by controlling the pulse duration of voltage. In the second one, the coercivity of ferromagnetic layers can be modified by the voltage-controlled interfacial anisotropy [49]. Based on this phenomenon, electric-field-assisted switching in MTJ has been successfully achieved (see Figure 2.8(b)). These voltage-controlled or electric-field-induced implementations require no current or just a very small current to switch MTJ, which offers a new path towards ultra-low power MRAM and logic systems.



(a) (b)

Figure 2.8 (a) Three-terminal device based on the giant spin Hall effect in β-Ta/CoFeB. (b) Electric-field-controlled switching in a CoFeB/MgO/CoFeB MTJ with interfacial PMA.

2.2.2 Magnetic domain wall (DW) nanowire

Magnetic domain walls (DWs) are the transition regions separating the domains with distinct magnetization direction, in which the magnetization vectors indeed rotate continuously [50]. From the 1960s, as DWs were considered to have a great potential application for future logic

and memory devices, intensive studies have been carries out to deepen both experimental and theoretical understandings [51]. At first, "magnetic bubble" materials were particularly investigated to realize magnetic bubble memories, wherein the magnetic bubbles or domains can be created and propagated in only one direction under an external magnetic field in thin films [52]. Each of bubbles stands for one bit of data and they are separated by the boundaries of bubbles, which can be considered as a kind of DW. As the magnetic bubble is very small and can be propagated along a series of parallel tracks, bubble memory was considered as a promising technology to achieve "universal memory" that could provide high density and performance to meet all storage needs. Although most of the bubble material based devices failed to be commercialized, the substantial theories and experimental developments effectuated at that time, in particular, the basis of DW dynamics based on Landau-Lifshitz-Gilbert (LLG) equations, were playing a vital role up to now [53].

As for the withdrawing of magnetic bubble memory from the history, there are principally two reasons: one is the introduction of other more competitive technologies, for example, hard drive and other semiconductor memory; the other one is the limitations of itself. Since most of bubble-based devices uses magnetic field to control magnetic bubble motion, large space cost is required for the generation of magnetic field. In addition, the speed of magnetic bubble motion is relatively limited for the extensive application. Nevertheless magnetic bubble memory is the first concept applying the field-induced DWs to store information.

With the development of nanolithography techniques, well-defined shape rather than large film can be patterned. One of the most successful applications of field-induced DW motion in sub-micrometer magnetic nanowire was to realize magnetic logic [54]. These logic gates consumed ultra-low dynamic power if we didn't take into account the magnetic field generation, which operated as the clock of logic circuits. The logic state was non-volatile allowing zero standby power to keep the computing results. However these circuits also had critical shortcomings for practical applications due to magnetic field. For instance, their speed was low (e.g. <100 kHz) and the magnetic field generation with advanced CMOS circuits was a big challenge to manage the power dissipation. In this case, with respect to the limitation of magnetic field-induced DW motion, current-induced DW (CIDW) motion is more recently reported and will be introduced below.

2.2.2.1 Current-induced DW (CIDW) motion

The concept of current-induced DW (CIDW) motion was firstly proposed by Berger in 1978 [55]. The spin direction of electron is changed from the magnetization direction of one domain to that of the other domain when the electron crosses a DW. This change can be regarded as an angular momentum change. A reaction torque is induced on the DW to conserve angular momentum. And this torque transferred from the electrons leads to a direction change of the local magnetization moment, which results in DW motion. From 1984, Berger and his collaborators presented continuously a series of theories and experiments concerning the exchange interaction between electrons and DWs [56-59]. They demonstrated that the DW velocity could reach dozens of m/s by applying a current with current density in the order of 10^{10} A/m^2 and microsecond long pulse. However, due to the limited fabrication techniques, this early work was mainly based on wide and thick NiFe films. Some additional current-induced effects (e.g. Eddy current and self-field effect) might be introduced to drive DW motion, which made the analyses of CIDW motion be much more difficult.

In 1996, STT effect was first predicted theoretically by Slonczewski and Berger [24-25], in which a spin polarized current can exert torque on magnetic moment and reverse the magnetization direction. It is extensively used in manipulating magnetization and controlling DW dynamics. This effect was later confirmed experimentally and drew enhanced attentions on magnetic devices based on CIDW motion.

Thanks to the advances in nano-fabrication and measurement techniques, the nanowires with the size below 100 nm can be patterned [60-62]. This form has a lot of advantages. Firstly, it minimizes greatly the self-field effect generated by the current flowing through the system, which allows us to investigate CIDW motion more ideally. Secondly, the width of nanowire can be smaller than that of DWs, which allows a single DW created to across the nanowire.

Lots of materials have been used to fabricate the nanowires, for example, single layer permalloy [56] and ferromagnetic semiconductor GaMnAs [63], etc. In our work, we mainly study the CIDW motion in magnetic nanowire with PMA, which has been revealed to require much less current density for DW propagation than those with in-plane magnetic anisotropy [64]. So far, there are actually certain PMA materials demonstrating a steady and pure CIDW motion, for

example, $(Pt/Co)_m$ [65-66], $Pt/Co/AlO_x$ [67] (see Figure 2.9), Co/Ni [68-70] and CoFeB/MgO [71-73]. Among them, by exhibiting entirely interfacial PMA and large TMR ratio, CoFeB/MgO structure is promising to implement CIDW motion based devices.

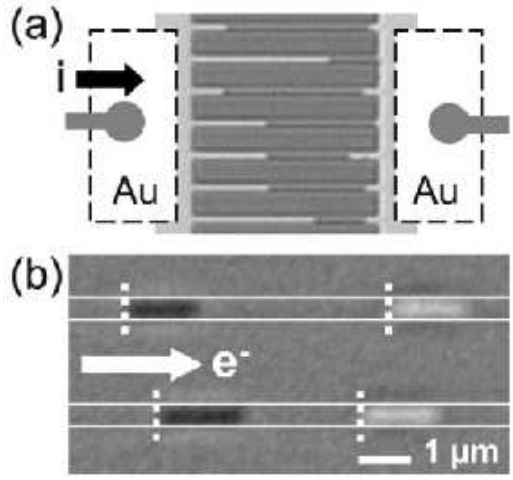


Figure 2.9 (a) Kerr microscope image of the $Pt/Co/AlO_x$ nanowire with schematically shown Au constacts. (b) Differential Kerr microscope image of CIDW motion in $Pt/Co/AlO_x$ nanowire. [67]

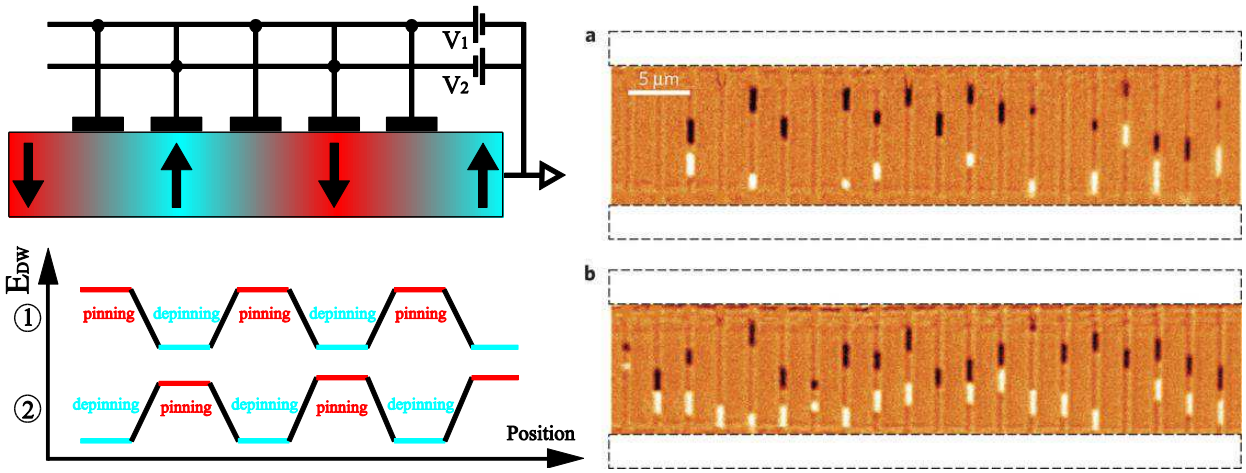
2.2.2.2 Recent progress of magnetic DW motion

Recently, various emerging phenomena of CIDW have been demonstrated. Among them, electric-field-controlled effect and spin-orbit torque (SOT) draw undoubtedly the widest attention.

Similar to the electric-field-controlled switching in MTJ, electric-field control of DW motion has also been observed in ultra-thin film with strong PMA, such as $Pt/Co/AlO_x$ [74-75] and $Pt/Co/GdO_x$ (see Figure 2.10(a))[76-79]. The control behavior is displayed by tuning the DW velocity via modulation of magnetic anisotropy. As it is a mechanism using gate voltage instead of current, ultra-low power can be expected. By enhancing or alleviating the DW velocity, pinning or depinning of DW can be operated. However, the experimental observations show that electric-field-controlled effect only impacts DW in its thermally activated creep regime and is still too weak for the case of high-speed DW motion, e.g. CIDW motion. The key challenge of this technology is to find a system with enhanced voltage-controlled magnetic anisotropy effect as well as large magnetoresistance. It is more recently reported that, in a specifically patterned structure, the pinning strength can reach as high as 650 Oe, enough to halt DWs with a speed of

at least 20 m/s, which encourages voltage-controlled pinning to be applied in the realistic fast spintronic logic and memory applications [79].

One of the characteristic features of CIDW motion based on STT mechanism is that the DW motion is along the electron flow (against the direction of the current). However, recent experimental measurements exhibited that CIDW motion in some ultra-thin asymmetric magnetic nanowires (e.g. Pt/Co/AlOx and Pt/CoFe/MgO) could be driven in the direction of the current, opposing to the electron flow [80-81]. In addition, this DW displacement can be extremely efficient, which means that a relatively low current density (e.g. ~10^{12} A/m^2) can propagate the DWs in a high velocity, more than twice of that with STT mechanism. These counterintuitive phenomena are difficult to be understood under STT mechanism system. In this case, an additional torque, so-called spin-orbit torque (SOT), due to strong spin-orbit coupling was proposed.



(a) (b)

Figure 2.10 (a) Example of electric-field-controlled DW motions in magnetic nanowire: By using the configuration of voltage ①, the DWs are pinned at the initial positions. Once changing to configuration of voltage ②, the DWs move to the final positions. (b) Kerr micrographs showing the DW motion induced by Rashba effect [80].

Although the origins of this SOT have not been completely elucidated by now, it could come from two effects: the Rashba spin-orbit coupling occurring at two different heavy-metal/ferromagnet and ferromagnet/oxide interfaces due to structure lacking inversion symmetry (see Figure 2.10(b)) [82-84]; the spin Hall effect (SHE) converting charge current to pure spin

current in the thin ferromagnet [85-87]. These two mechanisms can be considered to provide an in-plane effective magnetic field and a perpendicular spin current, with respect to the electric current direction. Furthermore, the DWs in these ultra-thin wires are found to be chiral, which is ascribed to the combination of SHE and Dzyaloshinskii-Moriya interaction (DMI) [88-89]. These emerging discoveries and interpretations stimulate the appearance of novel spintronic devices, for example, skyrmions based storage devices, which can overcome the bottlenecks of current DW-based devices and offers qualitative breakthroughs in terms of power consumption, speed, size and reliability. However, this progress is still on the road and masses of fundamental and technological investigation should be continuingly carried out.

2.2.3 Other spintronic devices

Besides the MTJ and the magnetic DW nanowire, there are several other emerging spintronic devices. They all feature non-volatility, high density, high speed and low power. Here we introduce briefly some of them: spin-transistor, nanomagnet, 3D magnetic ratchet, spin wave and memristor.

The concept of the spin-transistor has been predicted early in the 1990s [90], and it was experimentally developed recently thanks to the rapid progress of ferromagnetic study on spin-orbit interaction. Spin-transistors, including “spin-MOSFET” and “spin-FET” devices, benefit from a similar structure to MOS transistors, as shown in Figure 2.11(a). In the spin-FET and its related devices, the source and drain have the same spin alignment. The on/off switching operation can be executed by spin precession of the spin-polarized carriers injected in the channel through spin-orbit interaction, which is controlled by the gate voltage. Since spin-MOSFET requires no spin precession of spin-polarized electrons in the channel, the channel region of the spin-MOSFET requires a material with low spin-orbit coupling. One “magical” material for the spin transport channel is single layer graphene, which holds ultra-high mobility and very low spin-orbit interaction [91-92]. By using the spin-transistors, full spin computing systems can be realized so as to achieve ultra-low power operations.

Nowadays, integrated circuits are mainly built from transistors. However, as circuits become even denser, the difficulties for creating the high-performance transistors are more and more obvious. Nanomagnet could be a potential alternative to realize logic and memory functions

without conventional transistors. Although the primal concepts are functional only at cryogenic temperature, Cowburn and Welland proposed a magnetic nanostructure in 2000, in which information can be propagated across a cellular automata device at room temperature [93]. Five years later, Imre et al. reported an architecture that implemented complete logic functions [94]. Recently, the nanomagnet based devices with PMA illustrated the great benefit of the universal majority decision. Based on the nanomagnets logic gates, such as inverter and majority gate (see Figure 2.11(b)), more complex digital logic circuit (e.g. 1-bit full adder) can be designed and fabricated [95].

Moving from mainstream 2D structure towards 3D structure could allow performance enhancements, especially it alleviates greatly the pressure to the continually increasing storage density and shrinking of the lateral size. By accurately controlling the thickness of each magnetic layer and the exchange coupling between the layers, the 3D magnetic ratchet allows information in the form of a sharp magnetic kink soliton to be unidirectionally shifted from one magnetic layer to another [96]. As demonstrated in Figure 2.11(c), in a superlattice with numbers of layers, solitons are injected at the bottom or top. This "physical" shift register transfers data over near-atomic distances without requiring conventional transistors, and can thus execute true atomic-scale digital logic operation. Moreover, as solitons can selectively be manipulated with an external field or potentially by a spin polarized current, complex logic operation can be performed with a data shift register. This opens a route to realize 3D microchips for memory and logic computing. However, the route is still long, many fundamental technical challenges, such as dispersion in properties, dipole fields, data retention lifetime as well as reading and writing of solitons, which should be further investigated.

Spin wave can also be used for logic, by which the data are encoded in spin wave phase or amplitude (see Figure 2.11(d)) [97-98]. The process is performed by wave interference and integration with magneto-electric cells may enable non-volatile storage similar to all spin logic. The key advantage of spin wave based devices is that information transmission is accomplished without charge transfer, which can be highly beneficial for power efficiency. Moreover, it can also resolve the problem of interconnections and enhance architecture computing abilities [99]. However it consumes a lot of power for spin wave generation and data amplification. The spin

wave signal degrades rapidly during propagation. In addition, there are other challenges involving CMOS compatibility, low propagation speed and device scalability.

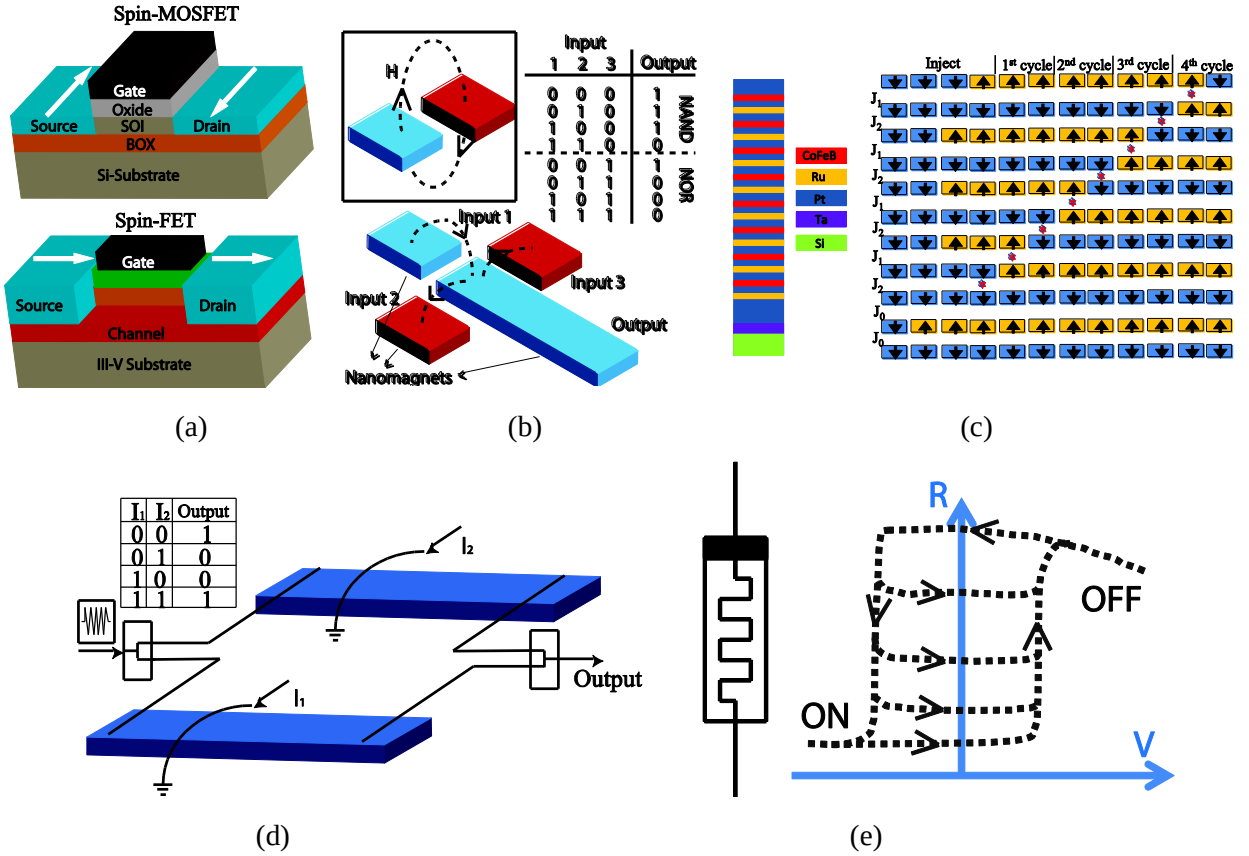


Figure 2.11 (a) Spin-MOSFET and Spin-FET. (b) Nanomagnet logic (NML). (c) 3D magnetic ratchet. (d) Spin wave logic. (e) Memristor. [90-104]

The fourth fundamental passive circuit elements (beyond the resistor, the capacitor and the inductor), called the memristor (i.e. memory+resistor), was envisioned by Leon Chua as early as 1971 (see Figure 2.11(e)) [100]. It fills the gap in the set of relationships of circuit elements and covers that between the magnetic flux and the charge. However, the device was not proven experimentally until 2008 by researchers from HP-labs, which brought an upsurge for the memristors and memristive systems [101]. In an ideal memristor, when current flows in one direction through the device, the resistance increases. If the current flows in the other direction, the resistance decreases. When the current is off, the memristor retains its resistance. This is the reason why the memristor can switch among multi-level resistance states. Although there is no uniform definition so far, the most generalized definition is that all 2-terminal non-volatile memory devices based on resistance switching effects could be considered as memristor.

According to that definition, many implementations have been carried out, e.g. TiO_2 memristor [101], ferroelectric memristor [102], and spintronic memristor [103], etc. Based on these devices, reconfigurable logic circuits and novel computing memory concepts attract more and more interest. One of the most promising applications is the neuromorphic system [104].

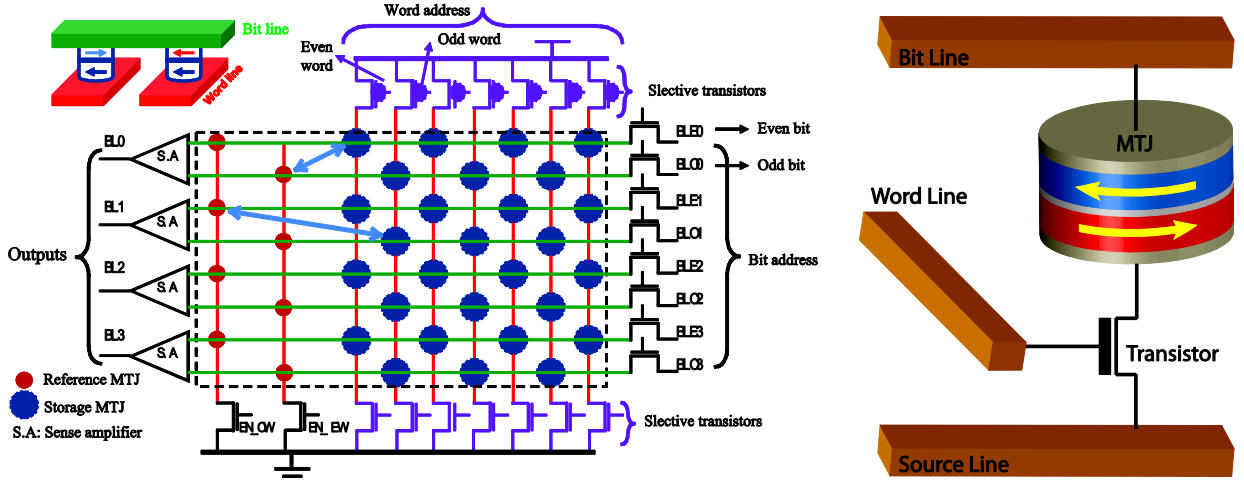
In summary, STT-MTJ and CIDW motion based devices currently bring out the most mature and promising features, which demonstrate a broader application base. That is the reason why this thesis focuses on the investigation of STT-MTJ and CIDW motion based devices.

2.3 Spintronics based hybrid logic and memory circuits

Based on the spintronic devices introduced above, the hybrid circuits are designed to execute the specific logic and memory functions. Here some hybrid circuits are reviewed with emphasis on Magnetoresistance Random Access Memory (MRAM), racetrack memory and logic-in-memory.

2.3.1 Magnetoresistance random access memory (MRAM)

The high resistance feature of MTJ compared to spin valves allows it more compatible for CMOS integration. Based on this basic element, MRAM is one of the most important spintronic applications and is believed to have a considerable impact on the future electronics beyond CMOS [105].



(a) (b)

Figure 2.12 (a) Schematic of the cross-point architecture for MRAM. (b) Schematic of the 1T/1MTJ structure for MRAM.

From the architectural point of view, two basic cell architectures have been proposed to realize MRAM [1]. As the cross-point architecture promises very high density, the architecture shown as Figure 2.12(a) was firstly proposed [1, 106-107]. In this cell architecture, the MTJs are connected to the crossing points of two orthogonal conducting arrays. By enabling one line of each array, the MTJ at the crossing point can be switched through an enough high magnetic field. However, this architecture suffers from several design technique challenges, for example, the unwanted sneak currents and low data access speed. In order to eliminate the unwanted current paths, one

transistor is usually added for each memory cell to build one transistor per cell (1T/1MTJ) architecture (see Figure 2.12(b)) [1, 107]. In this architecture, every MTJ is associated with a MOS transistor, whose gate is connected to the word line (WL), drain is connected to bit line (BL) through MTJ and source is connected to the source line (SL). This architecture is regarded the promising one to be widely applied for STT-MRAM. Although the density of the 1T/1MTJ cell architecture is less than that of the cross-point architecture due to the transistor required for each memory cell, it benefits for selecting the cell to implement the operations, not only reading but also writing.

According to the switching approaches effectuating on the MTJ, as mentioned above, there exits different generations of MRAM. FIMS-MRAM was the first generation and has already been commercialized. However, this approach requires high currents to generate magnetic field, which leads to high power consumption and low memory density. Since that, the academic and industrial efforts are spent to find more power-efficient, scalable, fast and reliable switching approaches. TAS-MRAM, STT-MRAM and TAS-STT-MRAM were then proposed. Table 2.1 summarizes the performance of different generations of MRAM.

Table 2.1 Comparison of various MRAM technologies

Technology	FIMS MRAM	TAS MRAM	STT MRAM	TAS-STT MRAM
Scalability	Poor (>60 nm)	Good (>40 nm)	Very good (<10 nm)	The best (<8 nm)
Min cell size	Large (~30 F^2)	Small (~10 F^2)	Very small (~6 F^2)	The best (~4F^2)
Endurance	10^{16}	10^{12}	10^{16}	10^{12}
Writability	Poor	Good	Very good	The best
Power	Very high	High	The best	Low
Latency	Long (>10 ns)	Very long (>20 ns)	The best (<5 ns)	short (<8 ns)

Among the diverse types of MRAM, STT-MRAM is considered the most promising technology and attracted the broadest attention of this domain [27, 32, 42, 108]. Aiming to achieve higher density, the size of MTJ is expected to continue scaling down. This trend leads to an increase of coercive field of MTJ for remaining stable in the presence of thermal fluctuations.

Correspondingly, in magnetic field-induced switching mechanism, the current generating magnetic field should also be increased, which is limited by the realistic techniques and applications. Contrarily, STT based current-induced switching mechanism is directly related to the energy barrier of magnetization switching, rather than the dimension of MTJ. Thanks to the high spin torque efficiency in the sub-volume regime discovered recently, it allows STT-MRAM to shrink potentially down to 10 nm [188]. Meanwhile, to ensure the scalability of STT-MRAM, high TMR ratio, good thermal stability and low switching current should be kept. The thermal stability is determined by the energy barrier between different configurations of magnetization, and the energy barrier is the product of the magnetic anisotropy and volume. In order to reduce the volume while keeping a high thermal stability, a higher magnetic anisotropy is thus required.

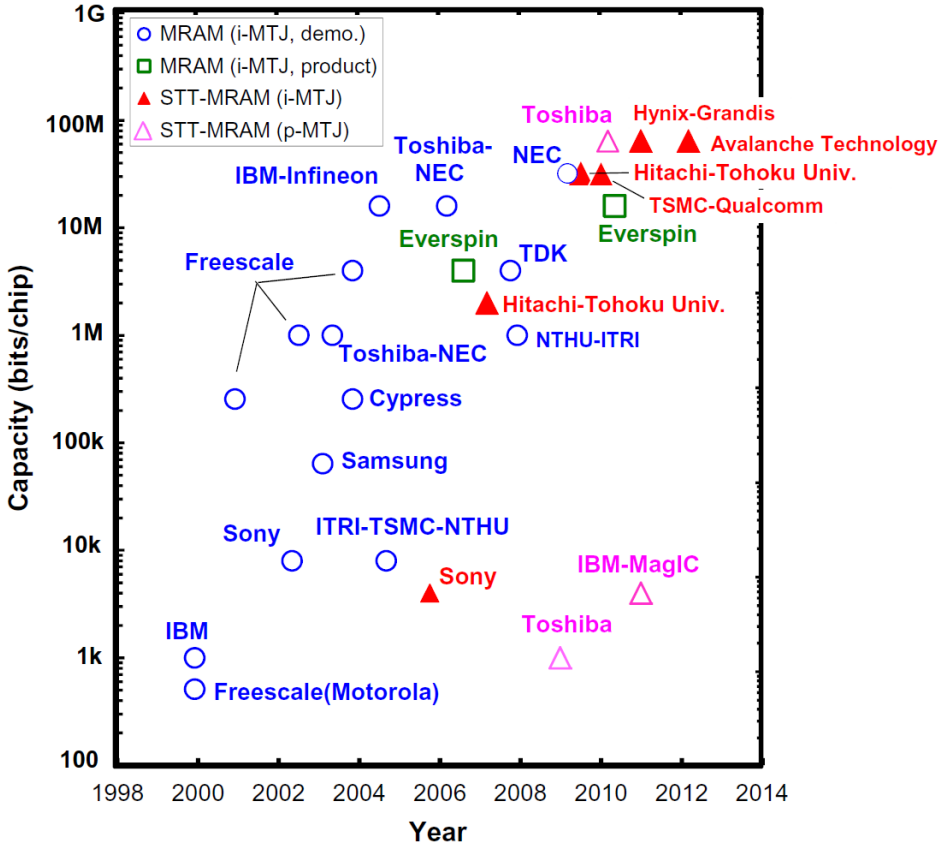


Figure 2.13 MRAM developments. MRAM stands for magnetic field switching MRAM, while STT-MRAM stands for those employing STT switching. [36]

In this situation, PMA introduced above come out, which provides a stronger magnetic anisotropy compared to in-plane magnetic anisotropy [36]. On the other hand, the threshold

switching current is proportional to the product of energy barrier and the Gilbert damping constant. The reduction of switching current requires a material with low Gilbert damping constant, which is a principal topic targeted by many material groups [109-111]. With the development of material and technique, numbers of prototypes and real products have been exhibited. Figure 2.13 summarizes some of the remarkable work relevant to MRAM development [112-118]. This encourages us to expert much more commercial products in the next years.

2.3.2 Racetrack memory

The observation of electrical CIDW motion in magnetic nanowires promises numerous perspectives [1, 65, 119] and the most interesting one is to build a novel ultra-dense non-volatile storage device, called "racetrack memory" (see Figure 2.14). The term "racetrack memory" was firstly proposed by Parkin in 2008 [120-121]. In the concept that he proposed, write head nucleates a local domain in the magnetic nanowire and a current pulse drives the domain to move sequentially from write head to read head. Data or magnetization direction is stored between two artificial potentials or constrictions, which pin the DW as no current pulse is applied. The distance between two constrictions can be extremely small to some nanometers and this allows an enormous storage (>GB) in a small die area. Compared with other non-volatile memory candidates [122], the scalability potential of racetrack memory is evident. By using MTJ as write and read heads, its operations, such as DW motion, domain nucleation and detection, can be addressed directly by CMOS circuits [123]. This hybrid integration makes racetrack memory promise high performance like high speed (>100 MHz) and low power beyond classical STT-MRAM. The nanowire can be built in 3D or 2D, the latter one is easier to be fabricated and become the mainstream solution for the current research on this topic. Based on in-plane magnetic anisotropy, the first racetrack memory prototype was presented in 2011 by IBM despite of its small capacity 256 bits [124]. However the intrinsic low energy barrier separating the two in-plane magnetization directions of storage layer leads to short data retention in advanced technology node (e.g. 22 nm) [40]. This drawback limits its use for high-density racetrack memory. PMA in some structures (e.g. CoFeB/MgO) providing a high energy barrier [38,125] were demonstrated and PMA MTJ become one of the most promising candidates to realize a read head. Advantageous domain wall nucleation current and speed with PMA MTJ were also observed recently [39] and this makes it be a better write head than in-plane MTJ.

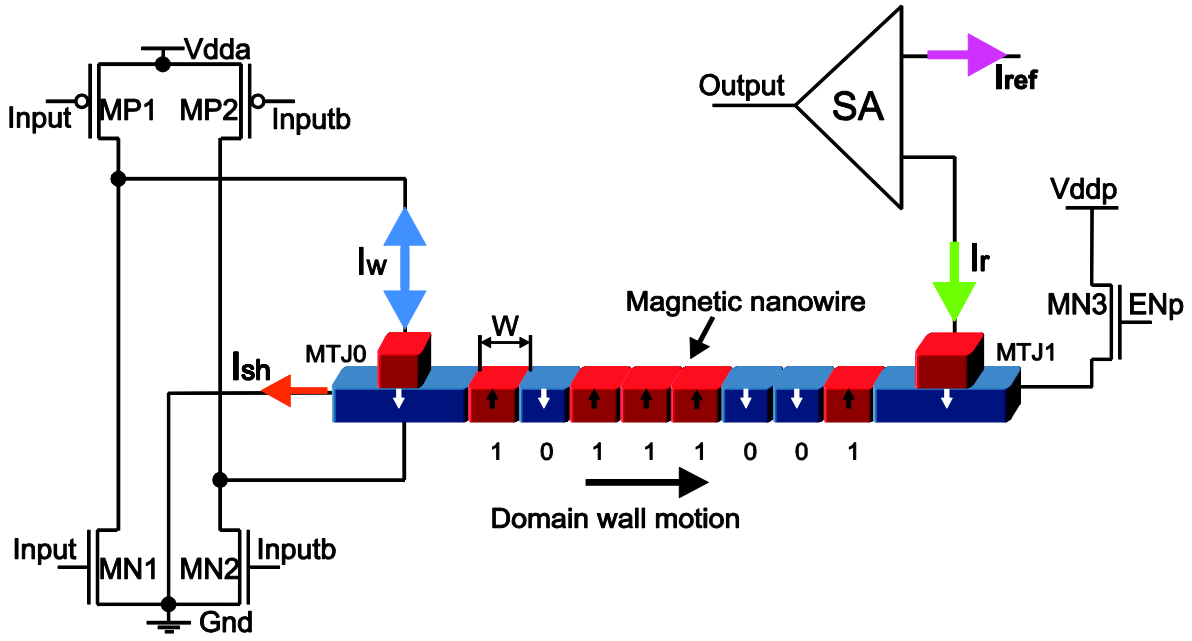


Figure 2.14 Racetrack memory based on CIDW motion, which is composed of one write head (MTJ0), one read head (MTJ1) and one magnetic nano-stripe. Iw nucleates data or magnetic domain in the magnetic stripe through STT approach, Ish induces DW motion along the magnetic stripe and Ir detects the magnetization direction through TMR effect.

The Cross-section structure of racetrack memory is shown in Figure 2.15, which includes mainly three parts: a magnetic stripe separated by constrictions to store data, two MTJs as write and read heads. The number of constrictions equals to the number of stored bits. It is noteworthy that the CMOS circuits dominate the whole area of this racetrack memory as the magnetic stripe is implemented at the back-end through 3D integration as MRAM.

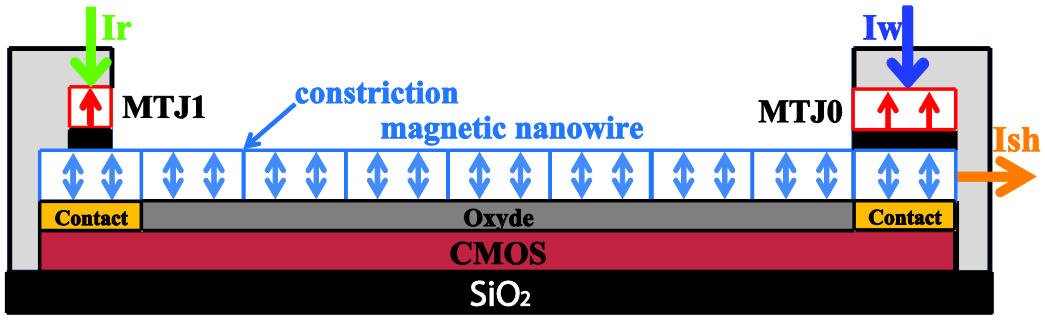


Figure 2.15 Cross-section structure of racetrack memory. At the back-end process, the magnetic nanowire is implemented above the CMOS/MTJ interfacing circuits, which generate Ir for reading, Iw for DW nucleation and Ip for DW propagation.

Figure 2.14 shows simultaneously one example of CMOS circuits to generate Iw and Ish, which are respectively bi-directional and uni-directional at the side of write head. Ir is driven by a sense amplifier [126] and it can convert the stored data from different magnetization directions to digital signal '0' or '1'. The writing and sensing circuits will be elucidated in details in the Chapter 4. In order to achieve the best write and read reliability, the width of write and read heads are different. For writing, a lower resistance of MTJ0 with larger width can reduce the rate of oxide barrier breakdown, which is one of the most significant constraints of the high-speed STT switching mechanism. On the contrary, high resistance of the MTJ1 with smaller width for reading can greatly improve the sensing performance [126].

2.3.3 Logic-in-memory

Logic-in-memory is an architecture in which memory elements are distributed among the arithmetic units for improving speed and energy performances with respect to conventional Von-Neumann architecture. The basic idea of this concept was proposed in 1969 by W. H. Kautz [127], however until the last years the appearance of non-volatile memories (e.g. MRAM) stimulates numerous efforts on this field. A lot of logic units, such as full-adder and flip-flop, have been designed or fabricated based on the logic-in-memory architecture.

As known that two elements are normally indispensable for the computer: logic unit (e.g. central processing unit (CPU)) executing arithmetic processes and memory (e.g. random access memory (RAM)) storing information. The Von-Neumann architecture where CPU and memory are separated is considered as the mainstream architecture and plays a dominant role for the computer. However, for the advanced computing system that is fabricated with deep sub-micron CMOS technology (e.g. 22nm) and built up with GHz CPU and GB memory, two main bottlenecks hinder its realization. Firstly, the processing frequency is limited by the data transfer distance between CPU and main memory. Secondly, performances are degraded due to the increasing static and dynamic power [128]. In this context, logic-in-memory concept, where logic units and memories are integrated together, is considered as promising solution to overcome the power and speed issues of Von-Neumann architecture. In this emerging concept, there is no need of large data traffic; the memory cell itself acts as storage and operator at the same time, which improves greatly the latency and energy efficiency.

A key challenge of logic-in-memory is to integrate appropriate memory technology. Traditional memories (e.g. DRAM) result in more complex circuits and larger energy consumption because the stored data require power supply to maintain [129-130]. This is also a radical reason why logic-in-memory concept has been slowly developed for several decades since it was proposed. Thanks to the experimental demonstrations of novel non-volatile memories such as STT-MRAM [1] and RRAM (OxRAM, PCRAM, CBRAM ...) [131], the stored data can remain while power is off. This can avoid the long interconnection with power supply and eliminate the standby power caused by leakage currents of CMOS. Moreover, most of non-volatile memories can be integrated in the back-end process above the CMOS circuits, hence higher density and lower latency can be expected for non-volatile logic-in-memory architectures.

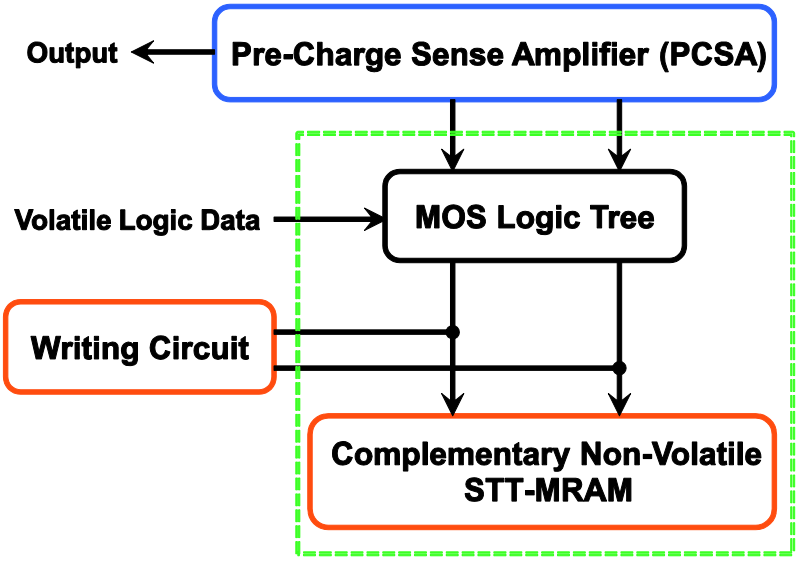


Figure 2.16 Generic logic-in-memory architecture based on STT-MRAM.

Figure 2.16 shows a generic logic-in-memory architecture based on STT-MRAM, which is composed of three parts: a sensing circuit (e.g. Pre-charge sense amplifier (PCSA)) evaluates the logic results on the outputs, a write logic block programs the STT-MRAM, and a logic control data block [136]. Considering every bit of MTJ costs a high programming energy (from 0.2 to 0.5 pJ/bit @40 nm) and relatively long switching delay (~ns), the logic data block contains simultaneously a MOS logic tree and MTJs in order to keep an area-power-efficient advantage. In this case, the logic volatile data can be driven by a high processing frequency, contrarily to analog non-volatile data that should be changed with a relatively low frequency, i.e. they are

more critical data or quasi-constant for computing. Depending on the MOS state in the logic tree and the STT-MRAM element state, the discharge currents are different in both branches and the current sense amplifier latches opposite logic values on outputs.

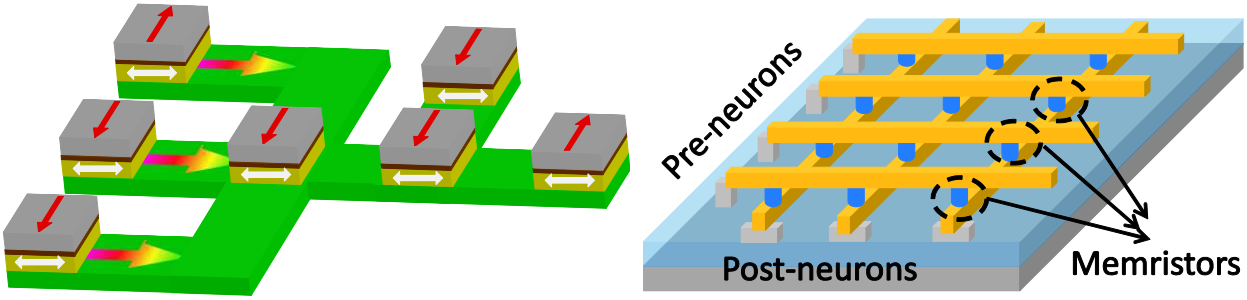
Benefiting from the progress of non-volatile technology, numbers of logic-in-memory prototypes have been recently reported. For instance, W. S. Zhao proposed a MRAM based non-volatile magnetic flip-flop for FPGA in 2006 [132]. In 2008, teams from NEC and Hitachi/Tohoku University presented the prototypes of a non-volatile latch and a STT-MRAM based magnetic full adder [133-134]. In 2010, the HP lab presented an advanced scheme in which logical computing was implemented in a unit consisting of only two memristors [135].

2.3.4 Other spintronic hybrid circuits

All spin logic (ASL) circuits [137] employ nanomagnets as digital spin capacitors to store data information and spin currents (through STT) to communicate, realizing logic gates based on a spin majority evaluation. Figure 2.17(a) shows an example to demonstrate the possible layout for constructing cascadable ASL logic gates. The magnetization directions of the nanomagnets can be switched between the stable states if enough torque is exerted on them. Information stored in the input magnet is used to generate a spin current that can be routed along a spin-coherent channel to the output magnet, determining its state based on the STT effect. The key features of ASL circuits are their compactness and completeness, because no MOS transistor is needed for the logic operations and all the logic functions can be constructed with a minimal set of Boolean logic gates. With such a design, a full spin computing system can be expected with extremely low switching power. However this is still a theoretical prospect currently and many issues, such as reliability and clock control, remain unresolved.

Spintronics may also allow the emergence of radically novel computing paradigms in electronics. In particular, for several years, researchers have been designing “neuromorphic” circuits that work analogously to the brain (see Figure 2.17(b)) [138]. Such circuits could allow a form of intelligent and ultra-low power computing (the brain can solve problems inaccessible to supercomputers with only 20 W). However, as fabricating neuromorphic system with pure CMOS has severe limitations, they require massive and ideally non-volatile memory for their “synapses”. Several groups have thus proposed to use spintronic memristors as synapses [139], in

particular relying on DW motion [140]. The associated computing units (neurons) may be realized by CMOS, but in some situations may also be implemented by multiple input spin valves. One group has shown the potential of this approach on real life applications using computer simulations [140]. An interesting point is that neuromorphic applications can tolerate more errors than traditional logic. This could allow using spintronic devices with lower currents than the other applications, which promises low power.



(a) (b)

Figure 2.17 (a) Example of all spin logic circuits: the output is determined by the spin currents injected into the channel by the inputs. (b) Schematic of a neuromorphic with memristor synapses in a crossbar configuration.

2.4 On-going big projects

As mentioned above, spintronis has attracted more and more attention from the academics and the industries. Many projects on this hot topic are carried out all over the world. Here we present some big ones, among them, MAGWIRE and MARS are two projects that support the work of this thesis.

2.4.1 MAGWIRE

Fast, high capacity, low form factor and low power non-volatile memories are already an important part of all electronics systems, representing a growing market segment, and should increase their importance in the future en route towards the "Storage everywhere" society. The MAGWIRE europeen project aims at investigating the potential of a racetrack memory beyond the 32 nm technology node by proposing innovative solutions in the fields of materials with perpendicular anisotropy, DW spin structure engineering, fabrication processes, architecture and CMOS integration. Its ultimate goal is to implement the integrated 2D racetrack memory in the standard CMOS 45 nm technology node to fully benefit from the cost/scalability economics reflected by Moore's law. The work of this thesis contributes to the design, simulation and evaluation for implementation of the racetrack memory, which is the indispensable preparation for realizing the final integrated demonstrator. MAGWIRE project gathers four leading academic experts, the University of Paris Sud in France, EPFL in Switzerland, the Cambridge Cavendish Laboratory in the United Kingdom, and the CNR in Italy, a major semiconductor manufacturer MICRON in Italy and a equipment manufacturer SINGULUS in Germany. This project is expected to generate a strong impact in terms of mass-storage applications beyond the limits of NAND Flash technology. It will be an essential step in the development of all future IC applications and in particular for mobile devices.

2.4.2 MARS

The project MARS (MRAM based Architecture for Reliable and low power Systems) was lanched by the French National Research Agency (ANR) in 2012. It is headed by the LIRMM and its partners comprise EADS, SPINTEC lab, IEF, LRI and CEA-LIST. This project aims at exploring the contribution of the technology MRAM to embedded processors architectures, and

identifying and quantifying the levels and systems architecture benefits especially in terms of reliability, energy consumption and performances. There are four main goals of this project: 1) Innovative hybrid CMOS/MRAM cells for processor architecutre; 2) overall consideration from technologies to system architecture (performance/reliability/energy); 3) Exploration and validation on embedded processor core; 4) generic methodologies for other non-volatile memory technologies. There are some milestones that have been achieved, for example, the first open source library of MRAM compact models (Spinlib) has been created, including generic STT and TAS compact models for electrical simulations. The compact models of PMA STT MTJ and PMA racetrack memory developed in this thesis complete this library, in particular the integration of stochastic behavior benefits for the reliability study of this project.

2.4.3 Other projects

Beyond the two projects introduced above, there appear numbers of spintronic projects implemented by almost all the leading semiconductor companies, microelectronic laboratories and universities. The subjects concern from the fundamentals of spintronics to spintronics based computing and memory applications. For example, HYMAGINE project "Hybrid CMOS/Magnetic components and systems for energy efficient, non-volatile, reprogrammable integrated electronics" led by SPINTEC lab in France is to explore the potentialities of this hybrid CMOS/magnetic integrated technology for microelectronic circuits; "Research and Development of Ultra-low Power Spintronics-based VLSIs" is a research project launched by the Centre for Spintronics Integrated Systels (CSIS) at Tohoku University in Japan, which aims at demonstrating the innovation of the integration of spintronic devices into logic VLSIs; Normally-Off Computing Project headed by Toshiba Corporation in Japan is a project aiming to realize normally-off computing in order to achieve higher performance per power for wide varieties of cimputing systems. Besides these projects, some large serach centres for the spintronics are also established recently. For example, C-SPIN (Centre for Spintronic Materials, Interfaces, and Novel Architectures) is a centre that brings together top researchers from across the USA to develop technologies for spin-based computing and memory systems. This centre is divided into five research themes: perpendicular magnetic materials, spin channel materials, spintronic interface engineering, spin devices and interconnects, and spintronic circuits and architectures. Our work in this thesis is very relative to the latter two themes.

Chapter 3 Compact modeling of PMA STT MTJ and PMA Racetrack Memory

3.1 Introduction

In order to design and optimize memory and logic circuits based on spintronic devices, the capacity to extract performance criteria, such as speed, area, reliability and power, from hybrid spintronics/CMOS simulations is indispensable. Thereby the SPICE-compatible compact models of spintronic devices are necessary. As presented in Chapter 2, PMA STT MTJ and CIDW-based device are the most remarkable and promising two types of spintronic devices under the current research background. Particularly, as racetrack memory is a typical application based on CIDW motion which is also integrated with MTJs, we thus decided to develop the compact models of PMA STT MTJ and PMA racetrack memory. A number of realistic material parameters and physical models have been integrated into the models to achieve good agreement with experimental measurements. Concretely speaking, for the compact model of MTJ, it includes the oxide barrier tunnel resistance model, the bias-voltage-dependent TMR model as well as the STT switching static, dynamic and stochastic models. Although the STT switching mechanism is a complex physical phenomenon, these physical models that we integrated in the compact model can be considered as the most important ones to describe the macroscopic behaviors of STT switching, such as the critical current, the switching time and the stochastic behavior. On the other hand, we used one-dimensional (1D) model and probability model to describe the DW motion in the magnetic nanowire. Then by combining MTJs as read and write heads, the compact model of racetrack memory can be developed.

With comparison of various SPICE modeling languages, such as Verilog-A, C and VHDL-AMS, we decided to programme these models with Verilog-A language that is compatible with standard CMOS CAD tools (e.g. Cadence platform) and provides an easy parameter interface. Designers can change the variables (e.g. geometrical parameters: length, width and thickness of the free layer as well as the TMR ratio) through the parameter interface to adapt to their specific requirements and sample's realistic parameters.

After the compact modeling, their functionalities should be verified. We performed these validations through diverse simulations in Cadence environment: DC simulation was used to validate the static behavior; transient simulation was used to validate the dynamic behavior and statistical Monte-Carlo simulation was used to validate the stochastic behavior.

3.2 Physical models of PMA STT MTJ

As CoFeB/MgO structure exhibits the best TMR ratio and switching performance, a compact model of PMA STT MTJ based on CoFeB/MgO stack is developed in this Chapter (see Figure 3.1). In this compact model, the related physical models presenting the static, dynamic and stochastic behaviors of are required to be integrated electrically: first, the physical models give the resistances of the MTJ depending on its magnetic configuration and its bias voltage; second, they define the current thresholds required to switch between both configurations; and finally, they take into account the switching delays, including stochastic fluctuations.

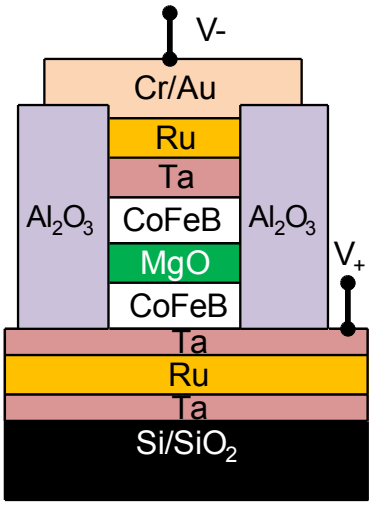


Figure 3.1 An example of PMA STT MTJ based on CoFeB/MgO stack.

3.2.1 Oxide barrier tunnel resistance model

The physical model of the tunnel junction conductance was introduced by Brinkman in 1970 [141]. In the original model, the conductance value is bias voltage dependent and is mainly determined by the oxide barrier height and the interfacial effect between oxide barrier and the ferromagnetic layers:

$$\frac{G(V)}{G(0)} = 1 - \left(\frac{A_0 \Delta\varphi}{16\bar{\varphi}^{3/2}}\right) eV + \left(\frac{9}{128}\frac{A_0^2}{\bar{\varphi}}\right)(eV)^2 \tag{3.1}$$

$$G(0) = (3.16\times10^{10}\bar{\varphi}^{1/2}/t_{ox})\exp(-0.125 t_{ox}\bar{\varphi}^{1/2}) \tag{3.2}$$

$$\Delta\varphi = \varphi_2 - \varphi_1 \tag{3.3}$$

$$A_0 = 4(2m)^{1/2} t_{ox} / 3\bar{h} \tag{3.4}$$

where e is the electron charge, m is the electron mass, V is the bias voltage, $\bar{\varphi}$ is the potential barrier height (e.g. 0.4 eV for crystalline MgO [15]), t_{ox} is the thickness of oxide barrier and $\bar{h}$ is Planck 's constant.

Considering the oxide barrier is symmetric, $\Delta\varphi$ in Eq. 3.3 is equal to 0. In order to integrate this model into the compact model, simplified equations obtained from the above equations are used to calculate the parallel state resistance of the CoFeB/MgO PMA MTJ [142],

$$R(0) = \frac{t_{ox}}{F \times \bar{\varphi}^{1/2} \times Area} \times \exp(1.025 \times t_{ox} \times \bar{\varphi}^{1/2}) \tag{3.5}$$

$$R(V) = \frac{R(0)}{1 + (\frac{t_{ox}^2 \times e^2 \times m}{4 \times \bar{h}^2 \times \bar{\varphi}}) \times V^2} \tag{3.6}$$

where *Area* is the MTJ area, *F* is a factor calculated from the resistance-area product (*RA*) value of MTJ, which depends on the material composition of the three thin layers. For instance, if *RA* is defined as 10 $\Omega\mu m^2$, which gives F = 332.2 with Eq. 3.5. Meanwhile, it is noteworthy that there is no evident dependence between the resistance and bias voltage for the most advanced MgO based MTJ [15]. Thus, we determine *R(0)* directly as the resistance in the parallel state:

$$R_P = R(0) \tag{3.7}$$

3.2.2 Bias-voltage-dependent TMR model

TMR effect is a key factor for the sensing operation of spintronic memory and logic circuits. For instance, the error rate caused by the mismatch variation of CMOS transistors will be greatly increased if the TMR ratio is relatively low [126]. Thereby the high TMR ratio is strongly expected to ensure a reliable sensing, which is in particular important for logic chip where there are no error correction circuits [143]. However it was found that TMR ratio decreases with bias

voltage V [15]. In order to describe this behavior, the following equation extracted from the theory shown in [144] is included:

$$TMR_{real} = \frac{TMR(0)}{1+\frac{V^2}{V_h^2}} \tag{3.8}$$

where TMR_{real} is the real value of TMR ratio during simulation, $TMR(0)$ is the TMR ratio with zero bias voltage, V_h is the bias voltage as $TMR_{real} = 0.5 \times TMR(0)$. Based on Eqs. 3.7-3.8, the resistance of MTJ in anti-parallel state, R_{AP}, can be defined with the Eq. 3.9.

$$R_{AP} = R_P \times (1 + TMR_{real}) \tag{3.9}$$

3.2.3 Physical models of spin transfer torque (STT)

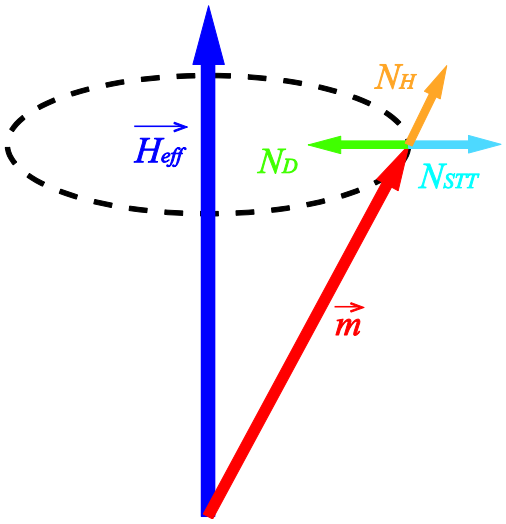


Figure 3.2 Diagram of the LLG equation. N_D is the damping term, N_{STT} is the STT term and N_H is the field precession term.

As mentioned in the Chapter 2, a spin-polarized current passing through a magnetic layer can lead to a change of angular momentum that is strong enough for reorienting the magnetization of the magnetic layer. This was predicted independently by Slonczewski and Berger in 1996 and is so-called spin transfer torque (STT) [24-25]. The effect of STT on the free layer of MTJ can be

described by the extra STT term in the Landau-Lifshitz-Gilbert (LLG) equation [145-146] given by

$$\frac{d\vec{m}}{dt} = -\gamma\vec{m}\times\vec{H}_{eff} + \alpha\vec{m}\times\frac{d\vec{m}}{dt} - \beta J(\vec{m}\times\vec{m}\times\vec{M}) \tag{3.10}$$

where $\vec{m}$ and $\vec{M}$ are the unit vectors of the free and pinned layers' magnetizations, α is the damping constant [147], γ is the gyromagnetic ratio, β is the STT coefficient depending on both the spin polarization and the geometric configuration of the spin torque efficiency. $\vec{H}_{eff}$ is the effective field that includes the external field, the anisotropy field, the magnetostatic field, the Oersted field and the exchange coupling field.

In this equation, the first term on the right is to describe the precession of the field-induced magnetization. The second term describes the intrinsic damping process that results in a decrease of the precessional angle as a function of time. The last term on the right is the STT term whose vector direction is opposite to the damping direction. In the current-induced system, the magnetization switching on the free layer can be considered the competition between the damping term and the STT term (see Figure 3.2). When the current density is small, the STT term is weaker than the damping term, then the magnetization dynamics maintain in an equilibrium state. In contrast, if the current density is high enough to make STT term stronger than damping term, the magnetization can be excited to larger precessional angles and further be switched. The critical current is defined as the threshold current to distinguish these two regimes, which is described by STT switching static model presented in the following section.

3.2.3.1 STT switching static model

The threshold for excitations driven by STT is given by the critical current. The static behavior to describe STT switching in PMA MTJ is mainly based on the calculation of threshold or critical current I_{C0}, which can be expressed by the Eqs. 2.8-2.9 mentioned in the Chapter 2 [38].

Note that the spin accumulation effects are neglected in this compact model and the spin polarization efficiency factor g is firstly obtained with the following equation to describe the

asymmetric current case [148]. It provides the best agreement with the experimental results illustrated in [38],

$$g = g_{SV} \pm g_{Tunnel} \tag{3.11}$$

where the sign depends on the free-layer alignment. g_{SV} and g_{Tunnel} are respectively the spin polarization efficiency in a spin valve and tunnel junction nanopillars. They are both predicted by Slonczewski,

$$g_{SV} = \left[-4 + (P^{-1/2} + P^{1/2})^3 (3 + \cos\theta)/4\right]^{-1} \tag{3.12}$$

$$g_{Tunnel} = (P/2)/(1 + P^2 \cos\theta) \tag{3.13}$$

where P is the spin polarization percentage of the tunnel current, θ is the angle between the magnetization of the free and the pinned layers [24, 149].

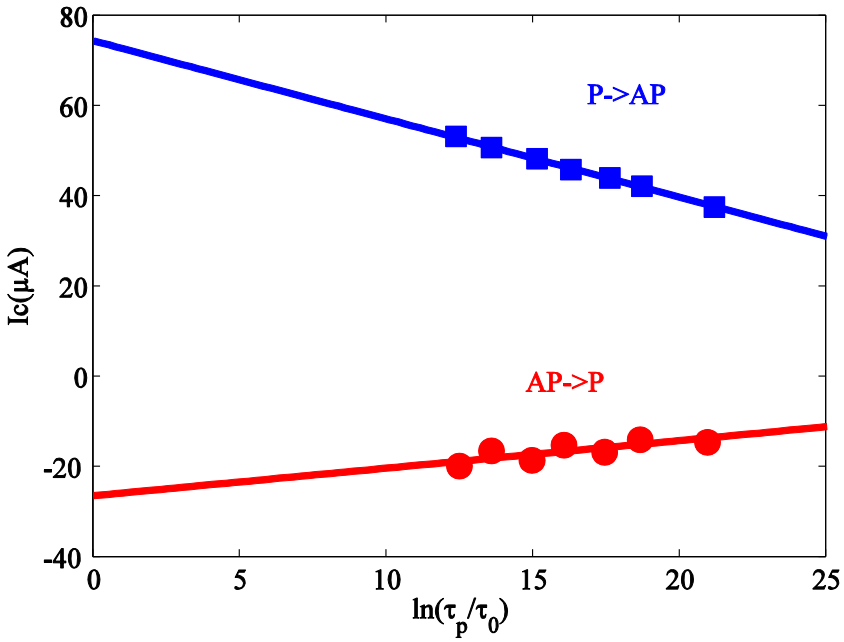


Figure 3.3 Verification of the static model with measured data reproduced by [38]. Solid lines represent the STT static model, the red points and blue squares represent the experimental results.

The good agreement between the physical model and experimental measurement has been verified by using MATLAB. Figure 3.3 shows the verification of static model with the measured

data reproduced by Ohno group [38]. The blue and red solid lines represent the STT switching static model for parallel to anti-parallel process and anti-parallel to parallel process, respectively. The blue squares and red points represent the experimental results. From the Figure 3.3, the overlaps between the lines and the squares (or points) show the good agreement and the feasibility of this physical model to describe the STT switching static behavior.

On the other hand, more recent experimental progress of IBM shows that an MTJ involving symmetric electrodes provides a single spin polarization efficiency factor *g* for both state change processes (anti-parallel state to parallel state process or parallel state to anti-parallel state process) of MTJ [43], which allows the same critical current for both parallel and anti-parallel states. In this mechanism, *g* is only related to TMR ratio and described as follows:

$$g = (TMR(TMR+2))^{1/2} / 2(TMR+1) \tag{3.14}$$

3.2.3.2 STT switching dynamic model

The dynamic switching behavior of STT in PMA MTJ shows the dependence of switching current on switching duration. It is considered to be a complex process as it can be categorized into two regimes depending on the relative magnitude between switching current (*I*) and critical current (I_{C0}, calculated by Eq. 2.8 for static behavior): thermally assisted ($I<0.8\,I_{C0}$) and precessional ($I>I_{C0}$) switching regimes. Thermally assisted regime can be described by Néel-Brown model and Precessional regime can be described by Sun model [150-151]. Note that there are no clear experimental results and theories related to the range from $0.8\,I_{C0}$ to I_{C0}, we thus neglect this range and consider no effect occurs in this range.

For each model, the relationship between current and duration follows different laws. For practical applications, the two regimes have their own specific interest: the thermally assisted regime corresponds to low current density but slower switching, which is usually used for the sensing operation; the precessional regime corresponds to fast switching (sub 3ns) but high current density, which is usually used for the writing operation. These two regimes will be respectively introduced as follows.

3.2.3.2.1 Thermally assisted switching regime (Néel-Brown model)

In the sub-threshold condition where the current remains below the critical current ($I<0.8I_{co}$), the switching can still occur thanks to thermal activation above the voltage/current-dependent barrier. In this case, the switching behavior can be described by Néel-Brown model [152]:

$$\frac{d\,\mathrm{Pr}(t)}{(1-\mathrm{Pr}(t))dt} = \frac{1}{\tau_1} \tag{3.15}$$

$$\tau_1 = \tau_0 \exp\left(\frac{E}{k_B T}\left(1-\frac{I}{I_{c0}}\right)\right) \tag{3.16}$$

where τ_0 is the attempt period, Pr(t) is the switching probability. Eq. 3.15 can be transformed to a simple formula:

$$t = -\tau_1 \ln(1-\mathrm{Pr}(t)) \tag{3.17}$$

These equations demonstrate that the STT dynamic switching behavior is probabilistic or stochastic. However, from Eq. 3.17, it can convert this stochastic behavior to be deterministic by determining the switching probability. That means ones should apply a specifically long current pulse to get the determined switching probability. This assumption would greatly simplify the description and analyses of the thermally assisted regime. Meanwhile, the stochastic effect is still the key point for this regime, which will be described and integrated in the following sections.

3.2.3.2.2 Precessional switching regime (Sun model)

In the case that the switching current is near or exceeding the critical one, the STT excitation becomes more obvious and deterministic. The high current pulse drives the magnetization to process, then after reaching the switching time, a magnetization reversal will occur suddenly and quickly [39]. Considering a small thermal fluctuation in this regime with a relatively high thermal stability, the average switching time is given by

$$\frac{1}{<\tau>}=\left[\frac{2}{C+\ln(\frac{\pi^2\xi}{4})}\right]\frac{\mu_B P_{pin}}{em_m(1+P_{pin}P_{free})}(I-I_{c0}) \tag{3.18}$$

where C is the Euler's constant, $\xi = E/k_{BT}$ is the activation energy in units of k_{BT}, P_{pin}, P_{free} are the tunneling spin polarizations of the pinned and free layers, we assume here that $P_{pin}=P_{free}=P$ in this compact model, m_m is the magnetic moment of free layer. Figure 3.4 shows the good agreement of this dynamic model with the experimental data extracted from [39]. From this figure, the increase of I and decrease of I_{c0} both contribute to scale down the switching latency. Considering the high currents are always ensured by the large-size transistors, this physical model also implies the alternatives to optimize the tradeoff between the overall area and the speed of hybrid spintronic/CMOS circuits.

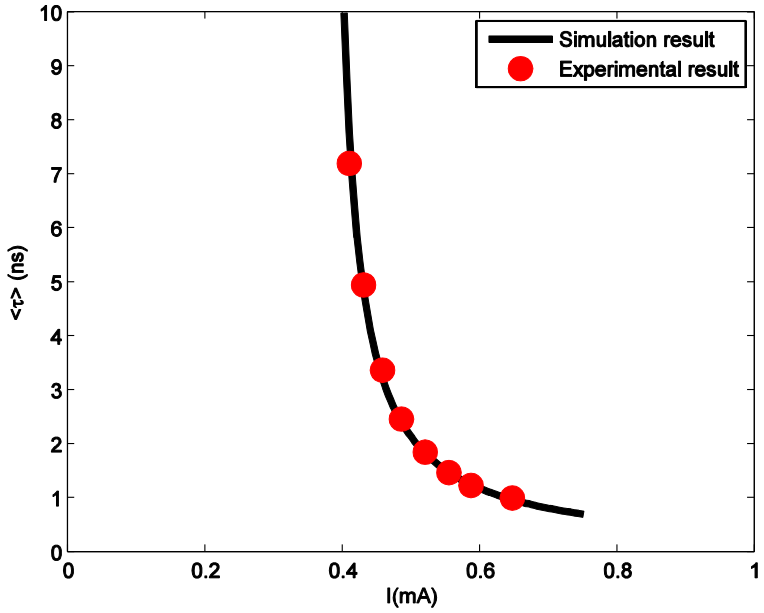


Figure 3.4 Comparison of the dynamic model with measured data, the size of MTJ is set to 105nm to adapt the experimental setup.

3.2.3.3 STT switching stochastic model

Recently, a lot of experimental and theoretical results have shown that, although STT switching

may allow sub-nanosecond switching duration, the switching process of STT is intrinsically stochastic, which results from the unavoidable thermal fluctuations of magnetization (see Figure 3.5) [153-155]. They are responsible for large fluctuation in the switching duration, which can be proven by the Eqs. 3.15-3.18 describing the dynamic behavior. Moreover, the stochastic behavior can also be divided into two regimes: thermally assisted ($I<0.8\,I_{C0}$) and precessional ($I>I_{C0}$) switching regimes.

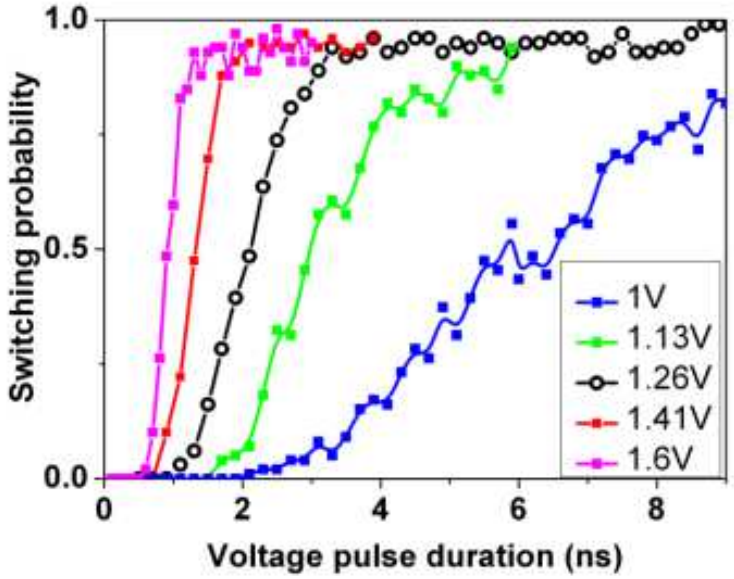


Figure 3.5 Experimental measurements of STT stochastic switching behaviors, high writing current drives faster speed and higher switching probability [153].

For the thermally assisted regime, we can transform Eq. 3.15 to another form:

$$\Pr(t) = 1 - \exp(-t/\tau_1) \tag{3.19}$$

It describes the probability density function (PDF) of the switching duration for this regime, which follows an exponential distribution with characteristic time τ_1 decreasing with the current density.

In the super-threshold region described as precessional regime, the stochastic switching is triggered by a thermal fluctuation which creates an initial angle between the current spin-

polarization and the magnetization of magnetic layer. The switching duration then follows a specific exponential-like distribution centered on the average switching delay time calculated by Eq. 3.18 [156].

From the above expressions, it shows that, in both regions, increasing the switching probability requires to increase either the write current or the current pulse duration. It could also be of great benefit for tolerating the high mismatch and process variations [157-158].

3.2.4 Hierarchy of the physical models

As shown in Figure 3.6, the physical models and the parameters are inextricably linked. From the viewpoint of physical models, tunnel barrier resistance model and TMR model are combined to obtain the resistances of MTJ; spin polarization efficiency model completes STT static model and the critical current deduced by static model is the key factor for dynamic and stochastic models. In this way, the resistance and current characters of MTJ have both been described. For the viewpoint of parameters, fourteen constants including the general constants (e.g. e and m) and the material constants (e.g. α and P) have been integrated. Six variables in terms of geometry, TMR, saturation field and cocervity field are set as inputs for designers to define the characteristics of this PMA STT MTJ. This integration hierarchy offers the great flexibility and facility for the following hybrid spintronics/CMOS circuit designs.

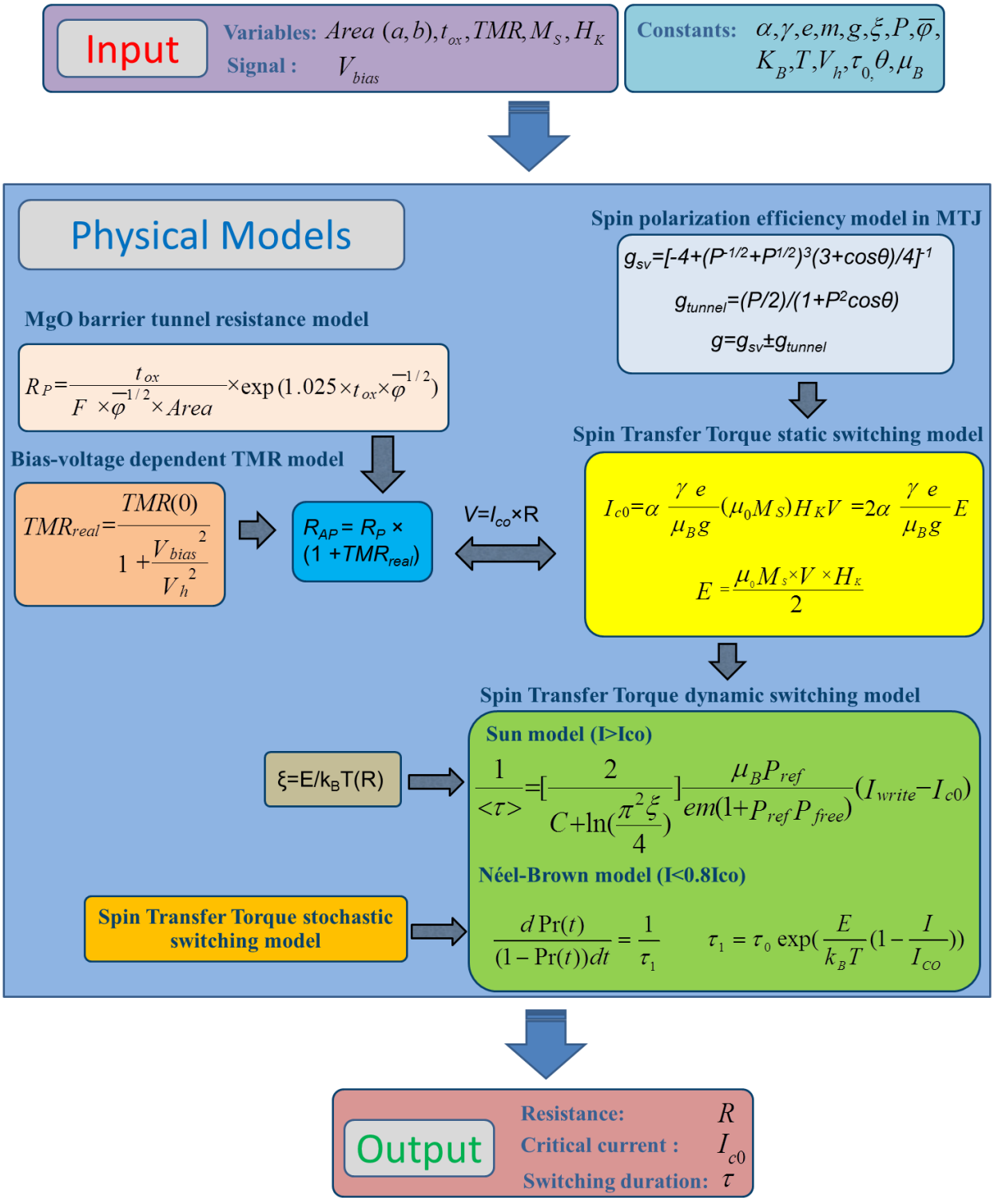


Figure 3.6 Diagram of integration of physical models into PMA STT MTJ

3.3 Physical models of PMA racetrack memory

Racetrack memory is a CIDW-based system. Its functional bases are the nucleation, detection and propagation of DW. Among them, the nucleation and detection can be executed by the MTJ write and read heads. Correspondingly the physical models for these two parts are the same with those of MTJ elucidated above.

Besides the nucleation and detection of DW, the DW propagation or motion in the magnetic nanowire is of key importance for the racetrack memory. When current density exceeds the threshold, the DW motion will be triggered. As the velocity is a crucial factor to describe the DW motion, we firstly introduce the DW motion velocity model. Then the stochastic effect is taken into account and the DW motion probability model will be addressed.

3.3.1 One-dimensional (1D) model

For the racetrack memory, the speed performance is governed by the velocity of DW motion. The physical model to calculate DW velocity is indispensable for the compact modeling of racetrack memory. According to the previous literatures, the dependence of DW velocities on current and magnetic field can be described by the one dimensional (1D) model. This 1D model is deduced from the LLG equation in a 1D system [159], which can be described as:

$$\dot{\phi}_0 + \alpha \dot{X} / \lambda = \gamma H + \beta u / \lambda + f_{pin} \tag{3.20}$$

$$\dot{X} - \alpha \lambda \dot{\phi}_0 = v_\perp \sin 2\phi_0 + u \tag{3.21}$$

where X is the position of a DW, and ϕ_0 is the angle that the DW magnetization forms with the easy plane. λ is the width of DW, α is the Gilbert damping constant, β is the dissipative correction to the STT, H is the external field, γ is the gyromagnetic ratio, f_{pin} is the pinning force. The velocity constant $v_\perp$ comes from the hard-axis magnetic anisotropy $K_\perp$ ($\sim K_\perp \lambda / \hbar$). u is spin current velocity. These two equations can describe a lot of qualitative features of DW motion driven by the field and the current. The field acts as a "force" to drive ϕ_0, the current acts as a "torque" to drive X. In addition, as the "torque" is also contributed from the hard-axis

magnetic anisotropy, the state of ϕ_0 can determine whether there is intrinsic pinning or pure STT. Considering only the process after depinning, Eq. 3.20-3.21 can be solved analytically and described in the forms of the influence of field and current on the velocity:

$$V = V_H + V_j \tag{3.22}$$

The velocity is the vector sum of field-induced (V_H) and current-induced velocities (V_j). Above the Walker breakdown field, the field-induced velocity contribution is given by

$$V_H = \alpha^2 \mu H \left\{ 1 - \frac{1}{1+\alpha^2} \sqrt{1 - \left(\frac{H_W}{H} \right)^2} \right\} \tag{3.23}$$

where the mobility $\mu = \gamma\lambda/\alpha$, H_w is the Walker breakdown field.

The general racetrack memory is based on CIDW motion, which means there is normally no magnetic field. Hence, the dependence of DW velocity on current is the key point for our compact model. Regarding the relationship between α, the damping constant, and β, the nonadiabatic coefficient, the dependence can be categorized into three cases. Before introducing these three cases, we should indicate the definition of the spin current velocity [160], which is given by Eq. 3.24.

$$u = \frac{\mu_B P j_p}{e M_S} \tag{3.24}$$

where j_p is the propagation current density. Figure 3.7 shows the dependence of DW velocity on current according to different configurations of α and β, which depends on the material of the magnetic nanowire.

When $\beta > \alpha$,

$$u_{WB} = \frac{1}{2} \gamma H_K \Delta \frac{\alpha}{\beta - \alpha} \tag{3.25}$$

$$< v >= \frac{\beta}{\alpha} u \qquad (u < u_{WB}) \qquad (3.26)$$

$$< v >= \frac{\beta}{\alpha} u - \frac{\sqrt{(1-\frac{\beta}{\alpha})^2 u^2 - (\frac{1}{2}\gamma \Delta H_K)^2}}{1+\alpha^2} \qquad (u > u_{WB}) \qquad (3.27)$$

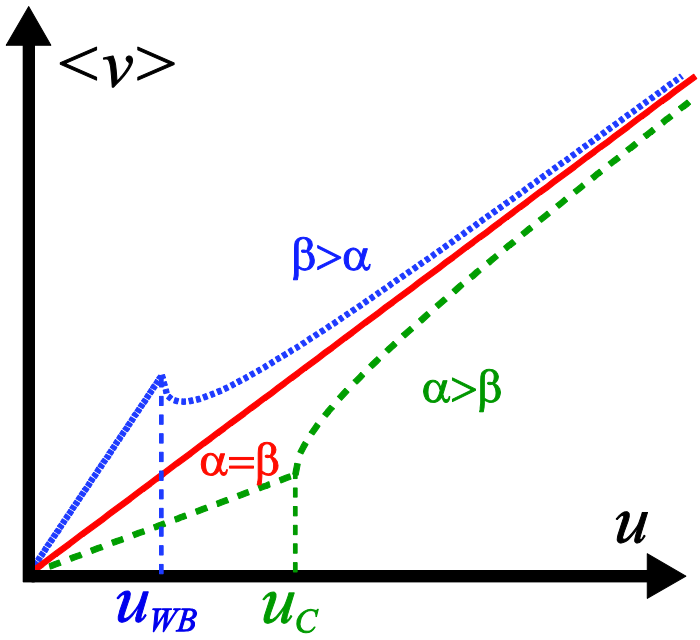


Figure 3.7 Dependence of DW velocity on current described by 1D model.

When $\alpha > \beta$,

$$u_C = \frac{1}{2}\gamma H_K \Delta \frac{\alpha}{\alpha - \beta} \qquad (3.28)$$

$$< v >= \frac{\beta}{\alpha} u \qquad (u < u_C) \qquad (3.29)$$

$$< v >= \frac{\beta}{\alpha} u + \frac{\sqrt{(1-\frac{\beta}{\alpha})^2 u^2 - (\frac{1}{2}\gamma \Delta H_K)^2}}{1+\alpha^2} \qquad (u > u_C) \qquad (3.30)$$

When $\alpha = \beta$,

$$< v >= u \tag{3.31}$$

where u_{WB} is the Walker breakdown velocity, u_C is the critical velocity corresponding to the critical current density of DW motion. In order to achieve a high speed racetrack memory, the current density should be more or far more than the critical one. In both cases, when applying a much higher current, the DW velocity approaches to spin current velocity. In our compact model, we take this assumption into account, which means we use the spin current velocity to directly represent DW velocity. Thus,

$$V_j = u = \frac{\mu_B P j_p}{e M_S} \tag{3.32}$$

In addition, we suppose that the DWs are definitely pinned when the current density is lower than the critical one, the velocity is thus kept to zero in this condition. We verified this physical model by comparing with the micromagnetic simulations done by Ohno group (see Figure 3.8) [71]. From the figure, a current density of ~2 x 10^8 A/cm^2 can trigger a DW motion in 50 m/s, which is beneficial for the logic and embedded memory circuits.

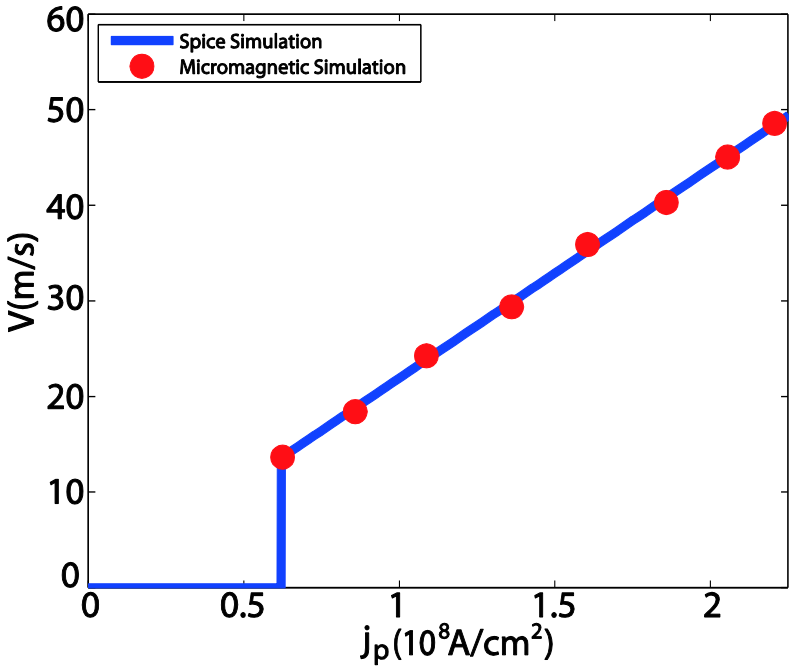


Figure 3.8 Good agreements with micromagnetic simulation for DW motion velocity as a function of current density.

By considering the distance W between two adjacent constrictions (see Figure 2.14), we can calculate the necessary pulse duration for current to move one storage element by the Eq. 3.33. For example, when W is 40 nm, the DW velocity is 50 m/s, the pulse duration can be as small as 0.8 ns. If neglecting the nucleation process, the frequency of racetrack memory can thus be as high as 1 GHz. If considering the nucleation process time (e.g. 1-2 ns), the frequency can still be 500 MHz.

$$D = W / V_j \tag{3.33}$$

3.3.2 DW motion stochasticity

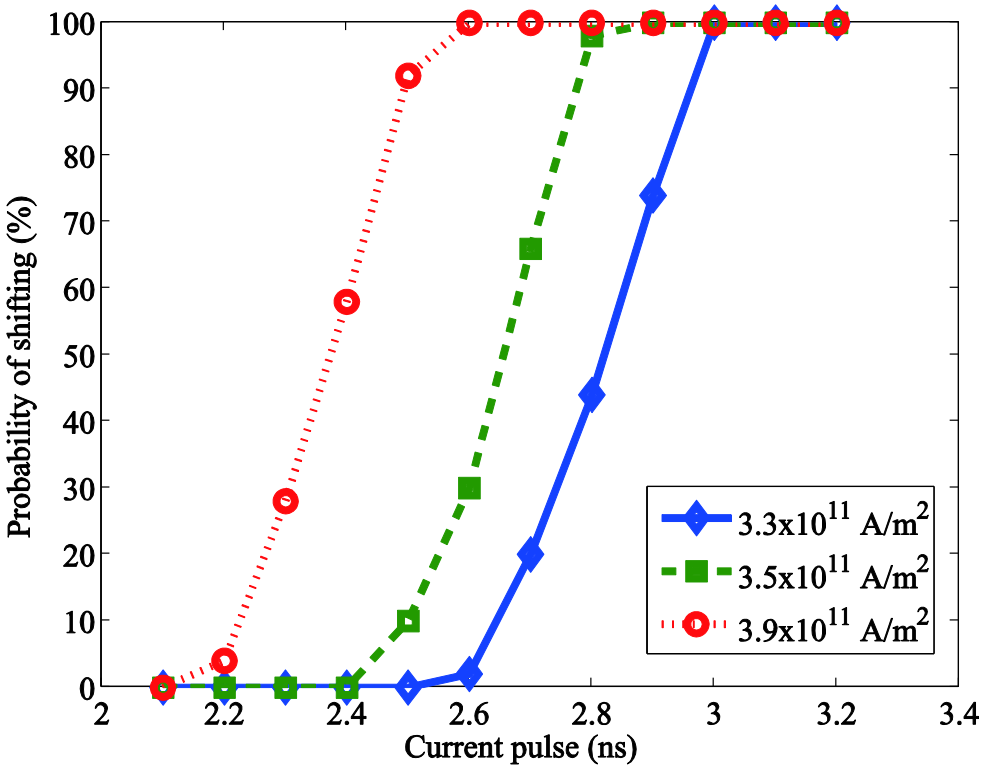


Figure 3.9 Dependence of cumulative probability of DW motion versus shifting current pulse duration and magnitude.

Caused by the thermal activation, stochastic nature has been found for DW motion in diverse structures and materials. With the reduction of the applied current or field by optimizing the techniques and the materials, the stochasticity of DW motion will be further enhanced [161]. DW velocity and displacement are susceptible to stochastic effect, which exerts a considerable influence to the feasibility and reliability of DW-based devices, not only racetrack memory.

However, as there have not been some coherent experimental results or physical theories concerning the pure CIDW in PMA materials, we refer to the measurements of DW motions in spin valve induced simultaneously by current and field. Once the more coherent results are discovered, we will update it in the next version of model. Under this condition, after depinning, the DW motion velocity is found to follow a Gaussian-like specific distribution centered with the value calculated by Eq. 3.32 [162]. We analyze the dependence of cumulative probability of DW motion versus different current pulse durations and magnitudes in Figure 3.9. It illustrates a coherent functionality of stochastic behavior where we can also find that higher and longer current pulse yields a more probable DW motion.

3.4 Compact modeling

3.4.1 Modeling language: Verilog-A

Modeling language is a crucial point for modeling. Especially, if we want to create an easy, efficient, accurate, fast and compatible compact model, the choice of modeling language is vitally important. Modeling language is the interface between the physical models of the component or system and the electrical simulators. Four types of languages are commonly used to model most of analog or digital components and systems, which are programming languages (e.g. C and FORTRAN), MATLAB, VHDL-AMS and Verilog-A. The programming languages have their advantages, such as fast simulating speed and direct access to simulator; however it must face the derivative problem and has no standard interface, which limit its applications in general macro-modeling. MATLAB is good at fitting data, but it cannot run directly in any analog simulator. VHDL-AMS is the first analog behavioral modeling language. Although it is able to run in some simulatiors, such as AMS designer of Cadence and ADVance MS of Mentor graphics, etc, whereas they are only AMS simulators and there is no clear definition of VHDL-A. Furthermore, VHDL-AMS is a much more verbose language and is harder to simulate quickly.

Verilog-A, the all-analog subset of Verilog-AMS language developed by Accellera, is an industry standard modeling language for programming most of behavioral models of analog systems [163]. Verilog-A can run not only in the same AMS simulators as VHDL-AMS, but also in Spectre [164], Eldo [165], ADS [166] and internal simulators of foundries such as STMicroelectronics, IBM and TSMC, etc. It exhibits various good features, such as capability of differential-algebraic equations, conservative or signal-flow systems, mixed disciplines; feasibility of parameterization, hierarchy, analog operators and analog events, etc. Moreover, Verilog-A models can be shared, promising global standardization. Large high-level systems including mixed-discipline and non-electrical systems could be quickly investigated with deeper design exploration. Another reason to choose Verilog-A as the modeling language for the compact models is that it is easy to understand, both characterization engineers and designers can comprehend it easily, which improves the continuity of this work and simplifies the update and development of future model versions.

3.4.2 Parameters of compact models

In order to verify functionality of the compact models and use them for further performance analyses, numbers of parameters coming from the physical experiments are embedded in the models. There are mainly three types of parameters including general constants, device technology parameters and device parameters. The technology parameters depend mainly on the material composition. The device parameters depend mainly on the process and mask design. These two types of parameters can be modified by the designers to optimize the circuit performance according to different magnetic process and material composition. We give default values of the parameters in Table 3.1. In this table, we also define the value variation range for some parameters as they cannot be changed unlimitedly according to the mainstream technology, for example, the thickness of oxide layer is confined between 0.6 nm to 1.2 nm.

Table 3.1 Parameters in the compact models of PMA STT MTJ and PMA racetrack memory

General Constants

Parameter	Description	Unit	Default value
k_B	Boltzmann constant	J/K	1.38×10^{-23}
e	Elementary charge	C	1.60×10^{-19}
μ_B	Bohr magneton constant	J/Oe	9.27×10^{-28}
C	Euler's constant		0.577
T	Ambient temperature	K	300
m	Electron mass	Kg	1.60×10^{-19}
$\bar{h}$	Plank's constant		1.0545×10^{-34}
μ_0	Permeability of free space	H/m	1.25663×10^{-6}

Technology Parameters

Parameter	Description	Unit	Default value
α	Gilbert damping coefficient		0.027
β	Nonadiabatic coefficient		0.027
γ	Gyromagnetic constant	Hz/Oe	1.76×10^{7}
P	Electron polarization percentage		0.52
H_K	Out of plane magnetic anisotropy	Oe	1734
M_S	Saturation field in the free layer	Oe	15800
$\bar{\varphi}$	Oxide layer energy barrier height	eV	0.4
V_h	Voltage bias when the TMR(real) is 0.5TMR(0)	V	0.5

RA	Resistance area product	$\Omega \cdot \mu m^2$	5 (5-15)
τ_0	Attempt period	ns	1
Δ	DW width	nm	10
θ	Angle of magnetization direction between the free layer and reference layer	°	0 or 90

Device Parameters

Parameter	Description	Unit	Default value
a	Length of MTJ	nm	40-65
b	Width of MTJ	nm	40-65
t_{ox}	Thickness of the oxide layer	nm	0.85 (0.6-1.2)
$TMR(0)$	TMR with zero volt bias voltage		120% (50%-600%)
w	distance between two adjacent constrictions	nm	40
t_{fl}	Thickness of the free layer	nm	1.3 (0.8-2)
Vol	Volume of the free layer	nm^3	a x b x t_{fl}

3.4.3 Component Description Format (CDF)

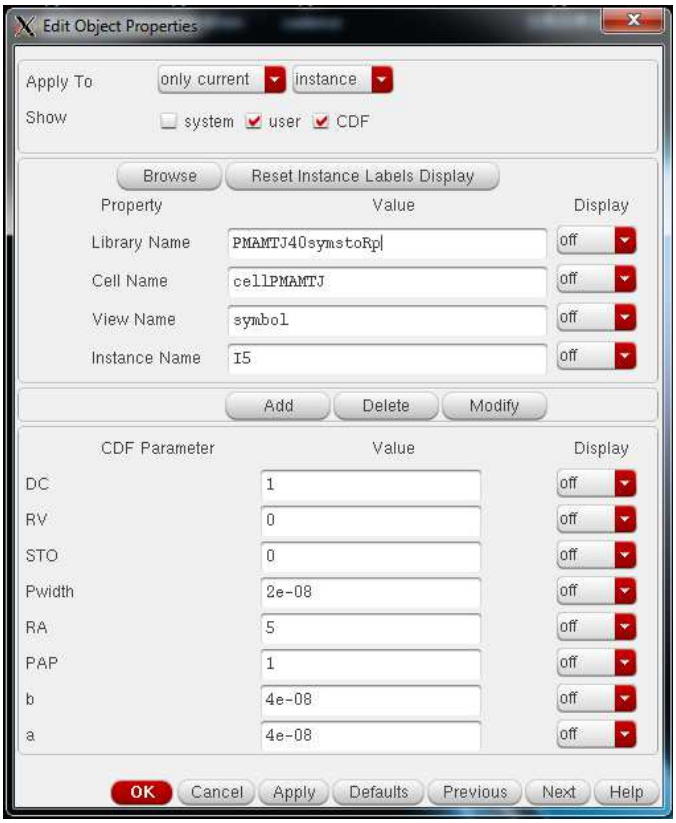


Figure 3.10 Component Description Format (CDF) in Cadence

In Cadence design system, we highlight the Component Description Format (CDF) function that is used to describe the parameters and their attributes of individual component or libraries of component. It allows the application independent on cellviews, and provides a graphical user interface (i.e. the Edit Component CDF form) for entering and editing component information.

Thanks to its favorable features, we initialize and define the specific parameters through CDF. For instance, by entering "0" or "1" in the column "PAP" in category "Property", we can modify the initial state of MTJ to parallel or antiparallel (see Figure 3.10). We can also calibrate "a" and "b" to adapt to the actual geometrical configurations. Furthermore, by using CDF tools we can modify multiple components individually, which facilitates implementation of more complex hybrid circuits.

3.4.4 Compact model cells for circuit design

3.4.4.1 PMA STT MTJ

Figure 3.11 shows the symbol of PMA STT MTJ compact model created in Cadence design system. This symbol contains three pins: a virtual output pin "State" is used to test the state of MTJ (parallel or anti-parallel). Its output must be one of the two discrete voltage-levels: level '0' indicates the parallel state; level '1' indicates the anti-parallel state. Another two pins "T1, T2" are the real pins of the junction. As there are pinned and free layers, these two pins are asymmetric: a positive current entering the pin "T1" can make the state change from parallel to anti-parallel; a positive current passing from "T2" to "T1" can switch the state back.

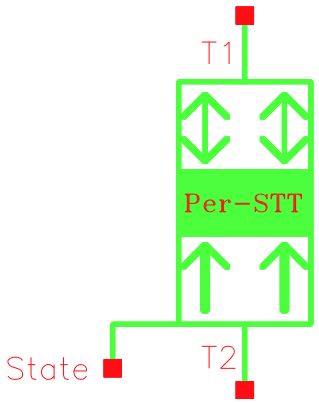
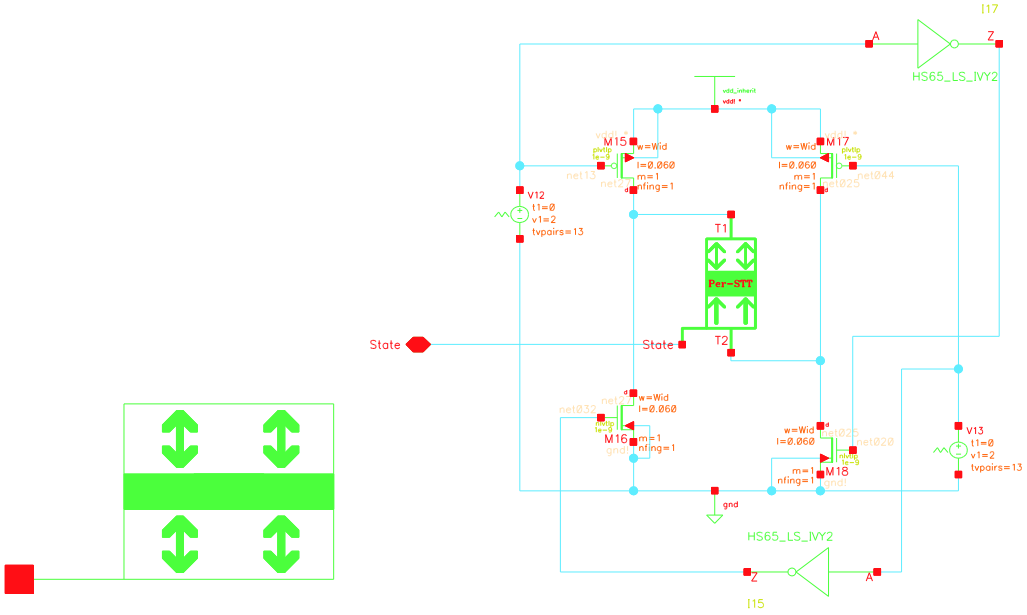


Figure 3.11 Symbol of the model PMA STT MTJ

3.4.4.2 PMA racetrack memory



(a) (b)

Figure 3.12 (a) Symbol of the write head of racetrack memory; (b) Schematic of internal DW nucleation circuit integrated in the write head symbol.

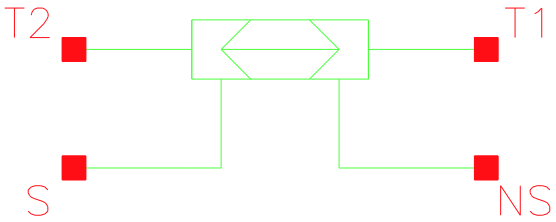


Figure 3.13 Symbol of magnetic domain in magnetic nanowire.

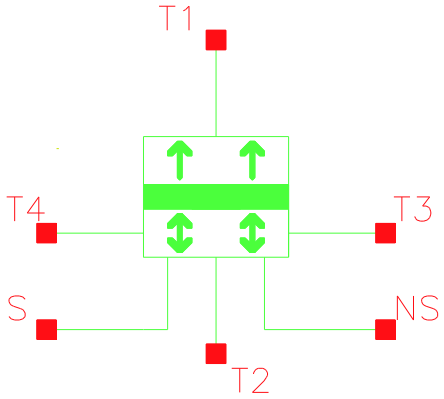


Figure 3.14 Symbol of the read head of racetrack memory.

Figure 3.12(a) shows the packaged symbol of write head with DW nucleation circuit, as shown in Figure 3.12(b), for PMA racetrack memory compact model in Cadence design system. As MTJ is used as write head, the DW nucleation circuit is similar to the writing circuit for individual MTJ cell. A bi-directional current can be generated to nucleate a local magnetic domain. The pin in the write head symbol outputs the initial state for racetrack memory that is ready to be propagated to the following part of magnetic nanowire. Figure 3.13 shows the compact model symbol of magnetic nanowire, which is the most important part for racetrack memory. Pins "T1, T2" are used to input a current and determine if it is enough for DW propagation. "NS" and "S" are used to receive the previous state and transfer it to the following part. Figure 3.14 demonstrates the symbol of read head. Except the pins similar to the three pins of MTJ, there are three extra pins: "T3, T4" are used to connect the DW propagation circuit; "NS" is used to detect the state of previous domain.

3.5 Validations of compact models

Once the compact models are created, the behavioral accuracy has to be proven via diverse simulations, such as DC, transient and statistical Monte-Carlo simulations. By using the compact model and STMicroelectronics CMOS design kit, PMA STT MTJ and PMA racetrack memory have been respectively validated. DC simulation is used to simulate the models or circuits with a constant current or voltage source. In this case, current or voltage is independent of time, which means it doesn't depend on the past state. This simulation is able to provide the useful information on basic and proper DC operating characteristics of the model. Transient simulation is dependent on time, which records and displays the behavior of model in discrete time. In this simulation, the simulator selects automatically the discrete interval to solve the differential and non-linear equations in the model and the designers can determine the interval or the number of the points according to the specific requirements. Statistical Monte-Carlo simulation is based on Monte-Carlo method, which is widely used from economics to nuclear physics. Especially, Monte-Carlo simulation used for integrated circuits yields the random variation of process and mismatch parameters. By effectuating enormous times of simulations and comparing their variations, the stochastic behavior of the model can be demonstrated and investigated. Therefore this simulation will be carried out to analyze the reliability feature for the hybrid spintronics/CMOS circuits in the following chapters.

3.5.1 Validation of PMA STT MTJ

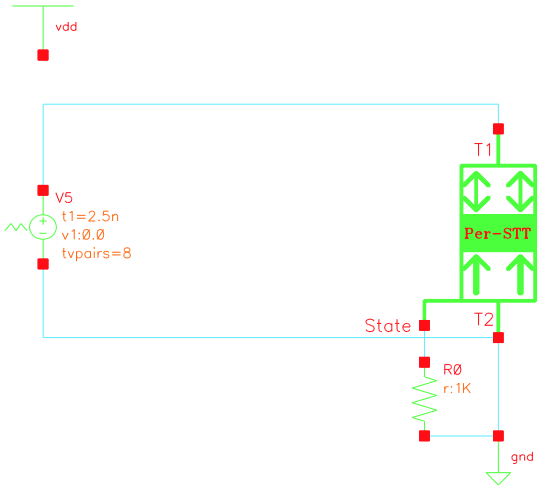


Figure 3.15 Schematic for the simulations of PMA STT MTJ.

Figure 3.15 shows a simple schematic to simulate an individual cell of PMA STT MTJ. It is worthy to note that, as the analogical behavior programed by Verilog-A is generally volatile, we use an external resistance (R0 in Figure 3.15) connected to the MTJ to save the non-volatile state. By detecting the voltage charging on the resistance, we can identify the state ('0' for parallel and '1' for anti-parallel) of MTJ.

3.5.1.1 DC simulation of PMA STT MTJ

DC simulation for the circuit shown in Figure 3.15 was firstly performed to verify the model functionality and the agreement of static behavior between physical models and experimental measurements. From the simulation results (see Figure 3.16), the critical current required to switch magnetization from parallel (P) to antiparallel (AP) state is about 72 μA, while for the reverse process the critical value is about 28 μA. These results confirm the strong switching asymmetry between two states shown in the experimental measurements of CoFeB/MgO PMA MTJs [38]. It is due to different spin polarization efficiency factor g in P and AP states (see Eqs. 3.11-3.13). It also describes the asymmetric voltage dependence for the resistances of MTJ, which is caused by the reduction of TMR ratio under a bias-voltage (see Eqs. 3.8-3.9) [15-17].

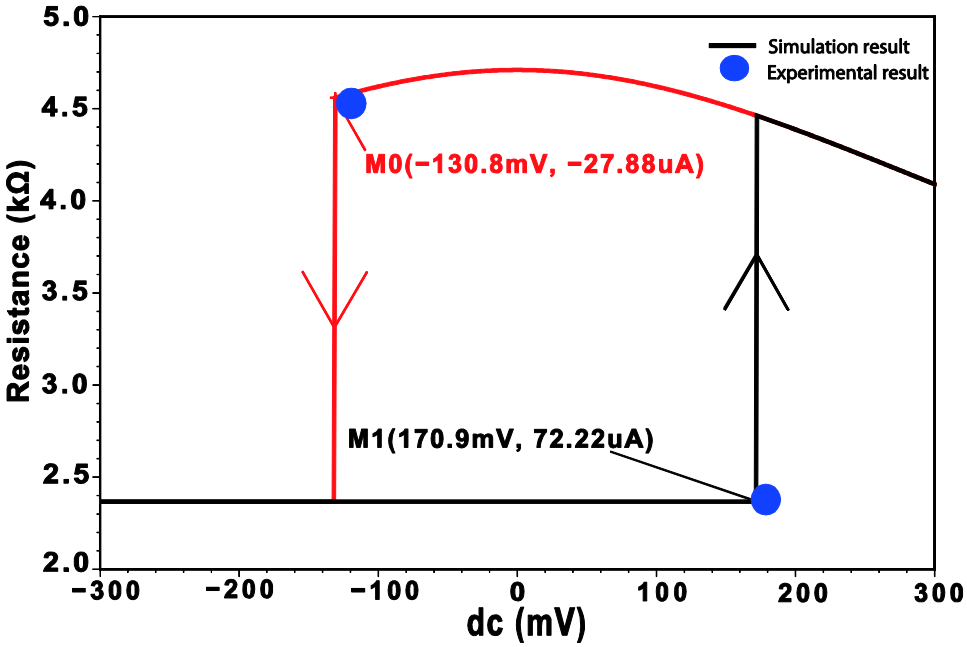


Figure 3.16 DC simulation of PMA STT MTJ (the red and black curves describe respectively the state switching from AP to P and from P to AP).

3.5.1.2 Transient simulation of PMA STT MTJ

Transient simulation, shown in Figure 3.17, was then performed to demonstrate the dynamic behavior in precessional regime. It verifies the agreement between physical models and experimental measurements extracted from [39]. It shows that the switching delay is inversely proportional to the writing current as described in Eq. 3.18. The static breakdown voltage of this MTJ is set to 1.5V, which leads to a maximum current of 633.75 µA that can be generated with the default configuration (see Table 3.1). In this case, the switching duration can be lowered down to ~0.5 ns, which potentially allows for a ~2 GHz operating frequency. This study confirms the potential application of PMA STT MTJ as a base for logic and memory chips. On one hand for logic computing, high currents can be sent to ensure fast speed. On the other hand for memory applications, small currents are used to provide high densities. This is because that the generation of high currents normally requires large-size transistors, which degardes the density effciency.

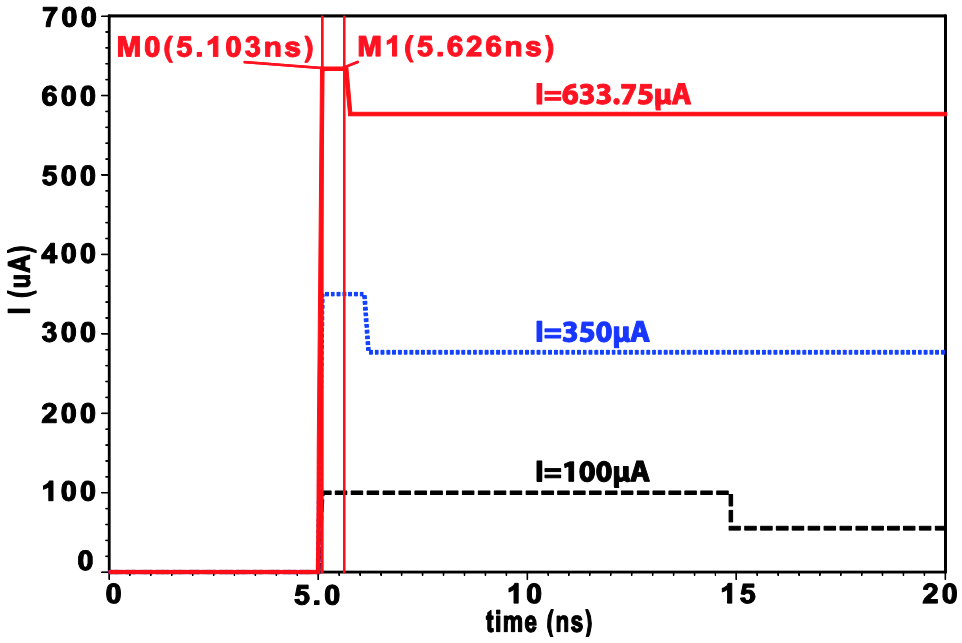


Figure 3.17 Transient simulation of the PMA STT MTJ demonstrates the integration of dynamic model and helps to study the tradeoff between die area and switching speed.

3.5.1.3 Statistical Monte-Carlo simulation of PMA STT MTJ

In order to integrate the stochastic behavior into our model, we used the random statistical functions provided by Verilog-A [163]. Concretely speaking, we utilize the function

"$rdist_normal" to generate a normal distribution and the function "$rdist_exponential" to generate an exponential distribution. To verify the functionality of this STT stochastic behavior, writing and sensing operations of a single MTJ cell are simulated respectively.

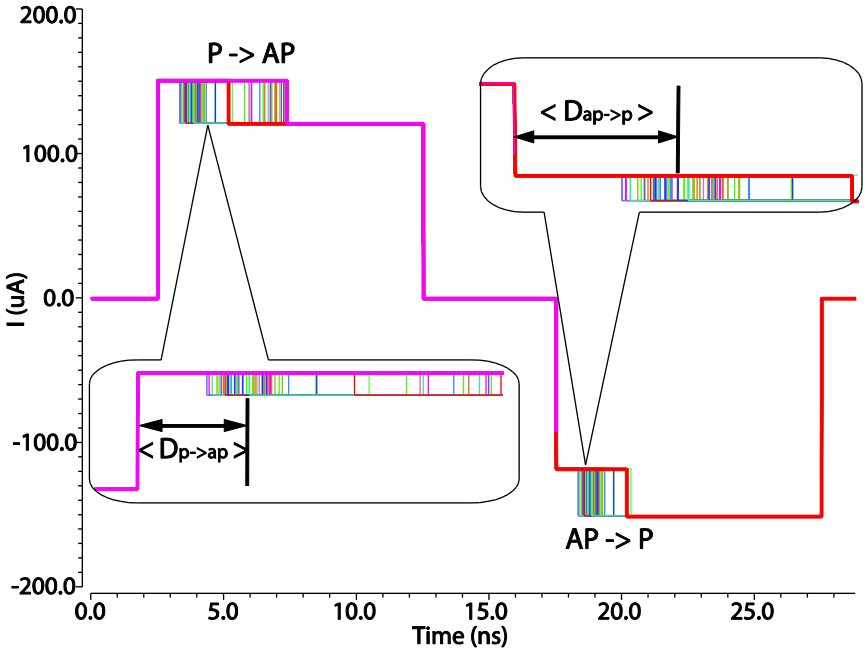


Figure 3.18 100 complete writing operation simulations (P to AP and AP back to P).

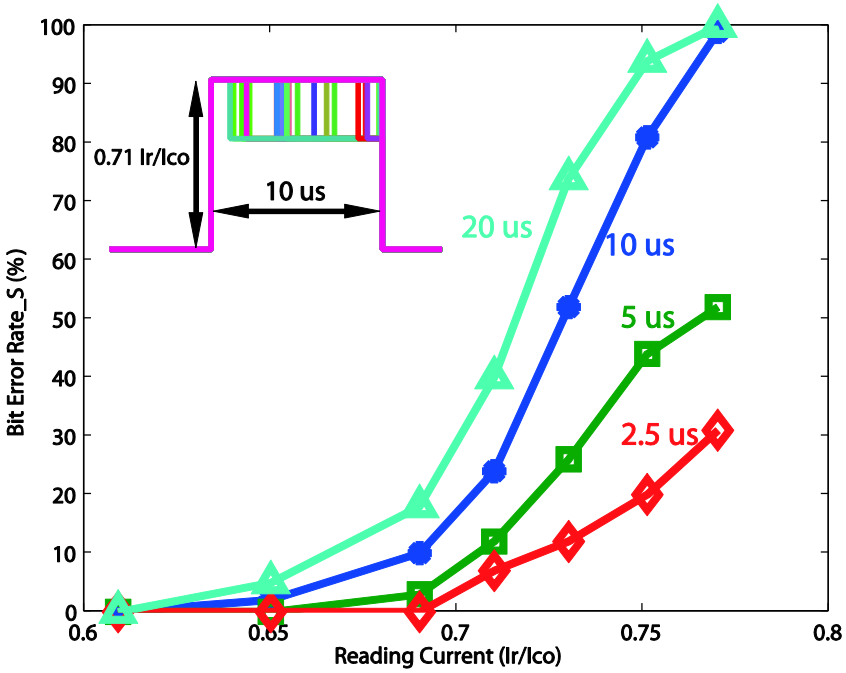


Figure 3.19 Dependence of sensing bit error rate (BER_S) versus sensing current for different switching duration pulses.

Figure 3.18 shows a statistical Monte-Carlo simulation of 100 writing operations. As expected the switching probability follows a random distribution around the average switching delay time as calculated by Eq. 3.18. The dependence of sensing Bit Error Rate (BER_S) versus sensing current for different duration pulses is illustrated in Figure 3.19. The BER_S grows exponentially with respect to the sensing current, which is consistent with the switching probability theory described by Eq. 3.19. In addition, for fixed amplitude of sensing current, increasing the pulse duration yields an increase in switching probability meaning an increase of probability of undesired writings during sensing.

Furthermore, in certain reliability issues studied in the following chapter, we also consider the variations of MTJ's key parameters, for example, MTJ TMR ratio, oxide barrier and free layer thickness variations. Sometimes, they are more important than stochastic effect, especially in deep-submicron technology below 40 nm. Figure 3.20 shows a Monte-Carlo simulation to illustrate the MTJ variation with 1% deviations of TMR ratio, oxide barrier and free layer thickness.

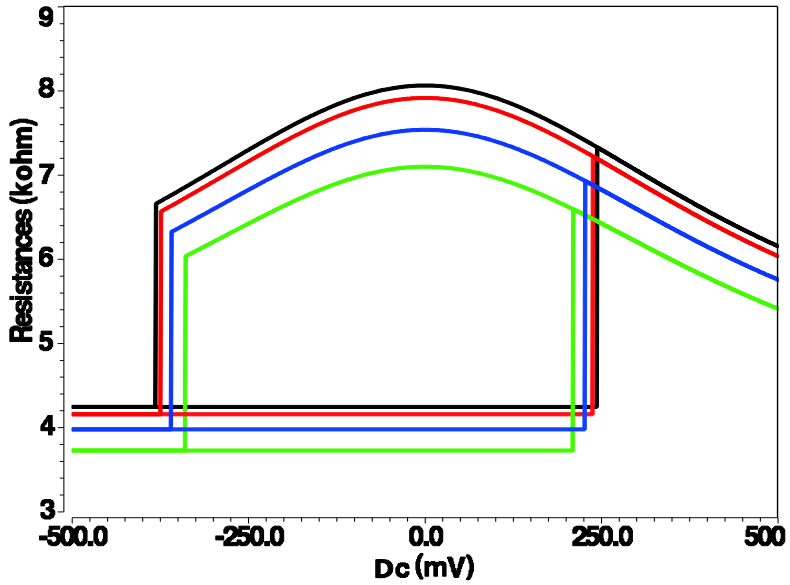


Figure 3.20 Statistical Monte-Carlo simulation of the PMA STT MTJ with 1% variation of TMR ratio, oxide barrier and free layer thickness.

3.5.2 Validation of PMA racetrack memory

Figure 3.21 shows a simulation schematic of 8-bit PMA racetrack memory, which includes write head, read head, magnetic nanowire and generation circuit for DW propagation current. In order to demonstrate the details of DW motion, we use particularly multiple magnetic nanowires connected with each other. In this case, each magnetic nanowire can be considered as an element of magnetic domains, which are separated by the constrictions. If two adjacent elements have opposite directions of magnetization, we can imagine that there is a DW at the place of the joint between them. From the point of view of information storage, each element of magnetic domains is considered one bit of data. Transient and statistical Monte-Carlo simulations have been carried out to validate the functionality of the compact model of PMA racetrack memory.

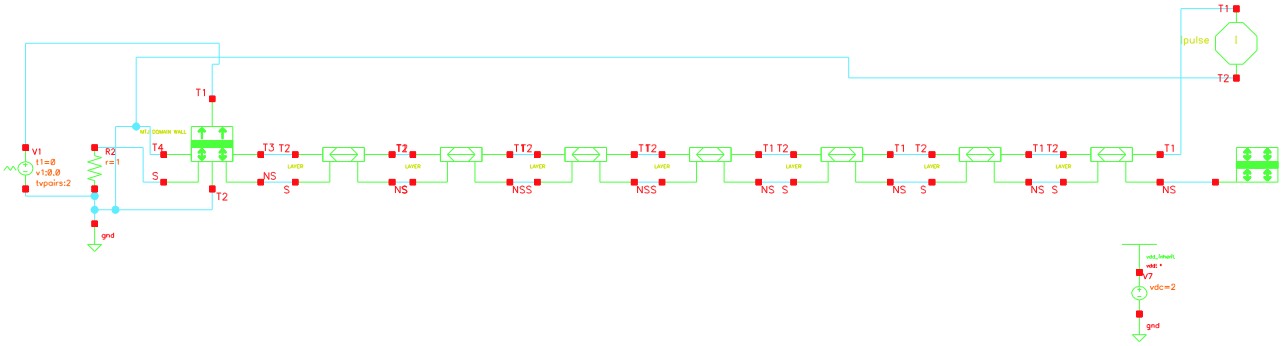


Figure 3.21 Simulation schematic of PMA racetrack memory.

3.5.2.1 Transient simulation of PMA racetrack memory

Figure 3.22 shows the transient simulation results driven by 50 MHz DW propagation current pulse (I_p) with magnitude of ~68.43 μA and duration of 2.5 ns. Initially, the states of all the storage elements in the magnetic nanowire are set to '0' and we plan to store an arbitrary logic pattern "…00010100111000…". The pattern is firstly sent to write head MTJ0 (see Figure 3.21) and the DW nucleation current I_w ~140 μA is activated during each data transition between logic '1' and '0' (see Figure 3.22(a)). I_p should be set to '0' during the data nucleation to avoid leakage currents. After the data nucleation, I_p is activated to induce DW motion in the magnetic nanowire and we can detect the same pattern at the read head MTJ1 after eight pulses (see Figure 3.22(c-d)). This simulation shows the correct operations of this PMA racetrack memory with low

I_w and I_p, allowing small die area. However the operating frequency, 50 MHz is relatively low for advanced logic circuits.

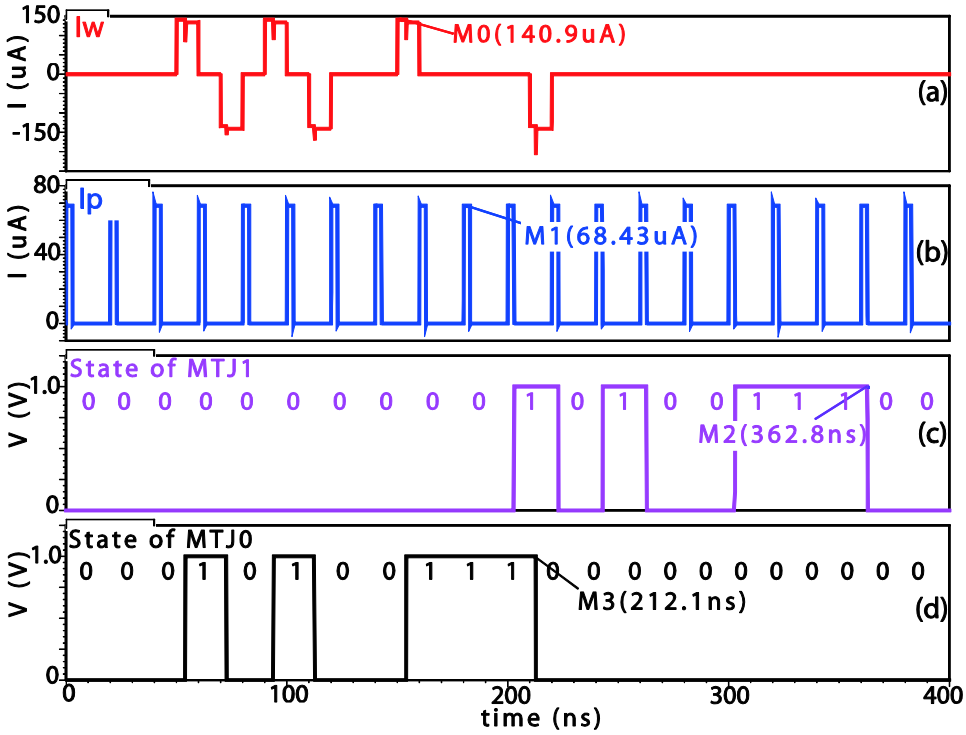


Figure 3.22 Transient simulation of PMA racetrack memory driven by 50MHz I_p (65 nm). (a) Current pulse I_w for switching the state of MTJ0. (b) Current pulse I_p valued 68.43μA with 2.5 ns duration at each 20 ns period. (c) State of read head MTJ1, the eighth storage element in the magnetic nanowire. (d) State of write head MTJ0, following I_p pulses, data is stored in the magnetic nanowire bit by bit.

We then increase the amplitudes of the driving current pulses. Figure 3.23 shows the transient simulation of this PMA racetrack memory with 500 MHz square wave I_p pulse. To ensure both the DW nucleation (~1 ns) and motion (~1 ns) in one cycle, current pulses I_w ~414 μA and I_p ~176 μA are respectively required. Note that it is difficult to use a 50% duty cycle square wave to drive the racetrack memory with in-plane anisotropy as the duration of DW nucleation is much longer than that of DW motion in that case. As square wave is generated more easily and used usually as the driving signal (e.g. CLK) in the integrated circuits, racetrack memory driven by square wave shows a bright prospect of wide application.

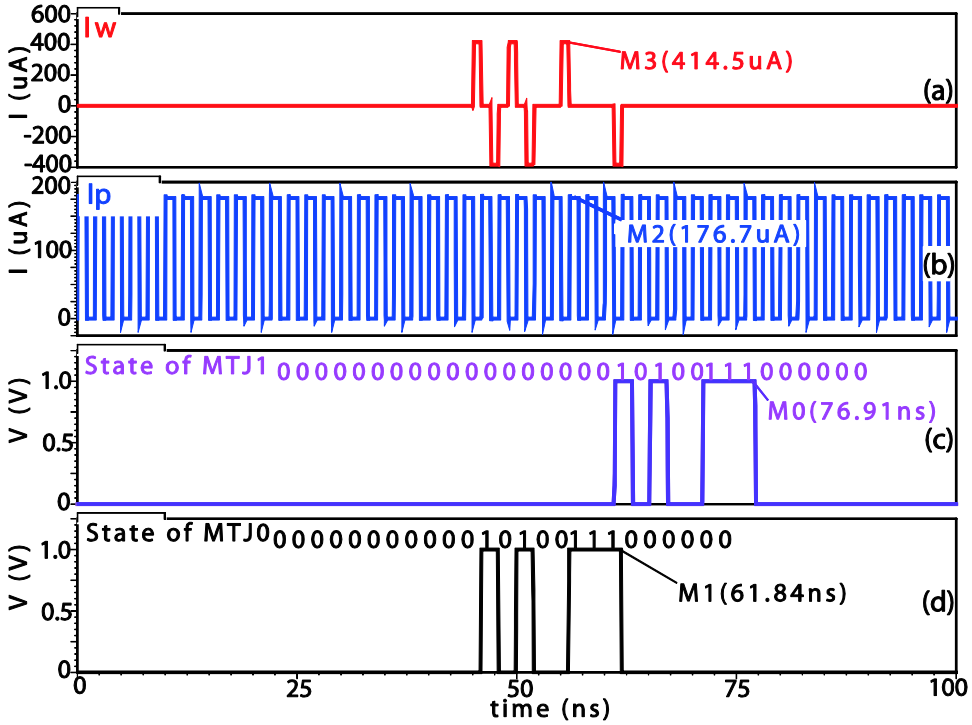


Figure 3.23 Transient simulation of PMA racetrack memory driven by 500MHz square wave I_p (65 nm). (a) Current pulse I_w for switching the state of MTJ0. (b) Current pulse I_p valued 176μA. (c) State of read head MTJ1, the eighth storage element in the magnetic nanowire. (d) State of write head MTJ0, following I_p pulses, data is stored serially in the magnetic nanowire.

Nevertheless, the higher currents allowing fast speed leads to larger die area. Figure 3.24 shows the performance tradeoff of this racetrack memory in terms of power, speed and area. The X-axis corresponds only to the size of five MOS transistors for DW nucleation and propagation (see Figure 2.14) as the sensing circuit is kept the same whatever the speed. F is the feature size of technology node. The Square wave I_p pulse is used for all the simulations and the data storage speed is shown to be linearly increased up to 500 MHz with 40 F^2 die area. Assuming the area of sensing circuit ~20 F^2 [126], one can briefly calculate cell area for our PMA racetrack memory, which is about 60/8=7.5 F^2/bit. If there are 64 constrictions [68] in the magnetic nanowire, the cell area would be nearly 1 F^2/bit. The energy dissipation per data storage operation E_{RM} can be described by Eq. 3.34:

$$E_{RM} = V_w \times I_w \times D_w + V_p \times I_p \times D_p \tag{3.34}$$

where V_w, V_p are the power supplies of current sources for DW nucleation and propagation; D_w, D_p are the pulse durations of I_w and I_p, which are both equal to half of Tp, the period of the square-wave current pulse I_p. E_{RM} can be advantageously lower than ~1 pJ, but it will be only slightly reduced for large die area, as shown in Figure 3.24. This is due to the nearly linear relationship between both DW nucleation and propagation currents with their durations (see Eqs. 3.18 and 3.33).

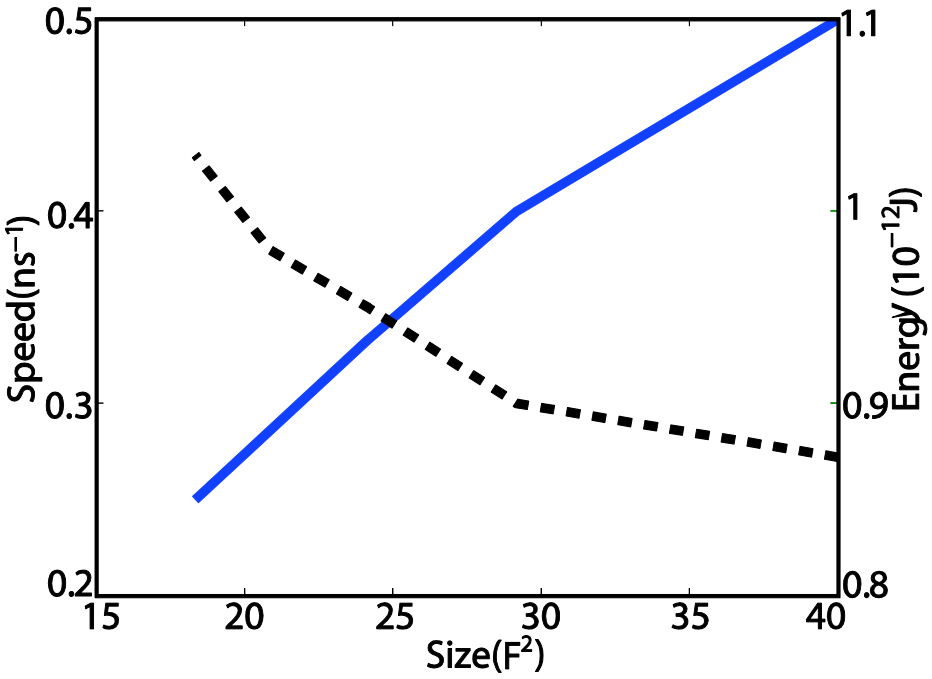


Figure 3.24 Dependence of I_p current pulse frequency (blue solid line) and energy dissipation (black dotted line) versus CMOS die area for one 8-bit racetrack memory.

The potential power efficiency of this PMA racetrack memory following feature size F minimization is also studied by using the compact model (see Figure 3.25). F presents important impact on dissipated energy as the small size of racetrack memory reduces linearly the DW nucleation and propagation currents while keeping the same delay (i.e. D_w and D_p) or operating speed. Two frequencies 500 MHz (solid line) and 250 MHz (dotted line) have been simulated and Figure 3.25 shows that about 75%~90% E_{RM} can be saved with the shrinking of technology node from 65 nm to 15 nm.

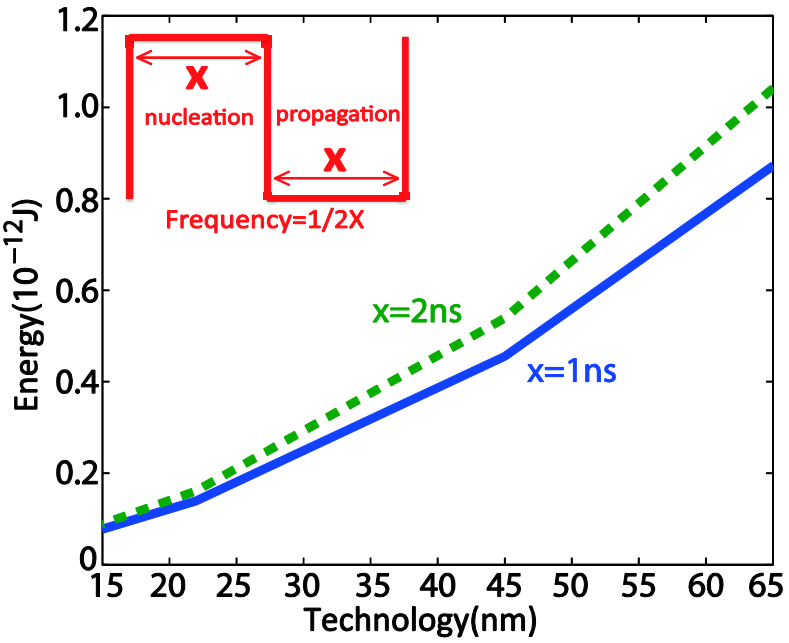


Figure 3.25 Dependence of energy dissipation and operating frequency versus racetrack memory technology node.

3.5.2.2 Statistical Monte-Carlo simulation of PMA racetrack memory

In order to illustrate the influence of stochastic effect on the PMA racetrack memory, we analyze full system of racetrack memory that involves magnetic nanowire, writing circuit, reading circuit and DW propagation current generation circuit. Considering 40 nm as the distance between two adjacent DWs, Monte-Carlo simulations for 7 bits, 8 bits and 9 bits have been carried out respectively (see Figure 3.26). We fix the voltage supply at 2 V and keep all the other geometrical parameters constant. The simulation results show the tradeoff relation among capacity, speed and reliability: 1) racetrack memory with higher capacity of storage is less reliable than that with lower density with respect to the same frequency; 2) reliability can be compensated at the expense of speed performance, which means that racetrack memory with higher capacity requires larger current pulse duration to achieve the same reliable operation. This is because longer magnetic nanowire leads to lower current density whereas it can store more bits of data, and lower current density degrades the reliability. Furthermore, even though there is the same current density passing through different lengths of nanowire, larger number of bits logically has a higher possibility to perform incorrectly than smaller number of bits.

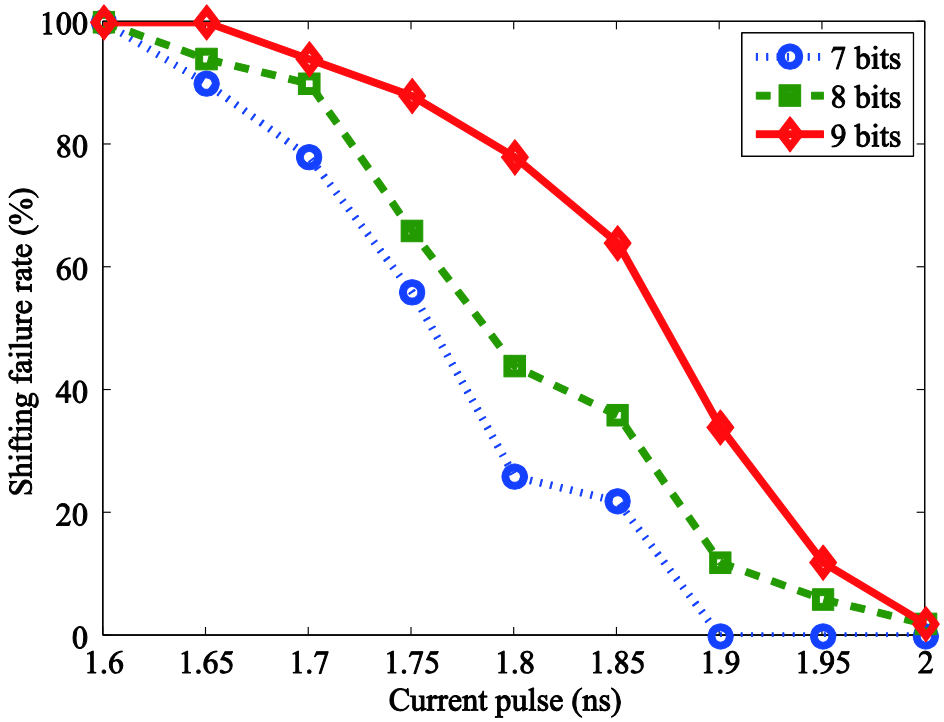


Figure 3.26 Dependence of shifting failure rate versus current pulse duration and racetrack memory capacity.

3.5.3 Layout implementation of PMA racetrack memory

Beyond the simulation validation of the compact model, we implemented a layout validation (see Figure 3.27) of racetrack memory including all the peripheral circuits. This implementation can also provide the performance on area. Magnetic layers counting nanowire and thin films of MTJ are above the CMOS part, which is up to Metal 2 level. Metal 3 is the last process to connect the circular MTJ nanopillars with peripheral circuits. In this example, the length of each nanowire is 3.8 µm, which can store at maximum 95 bits in 40 nm node. The size of this four word circuit is ~3.25 × 6.58 μm^2, and cell area is ~ 0.11 μm^2 which is equivalent to 7.7 F^2.

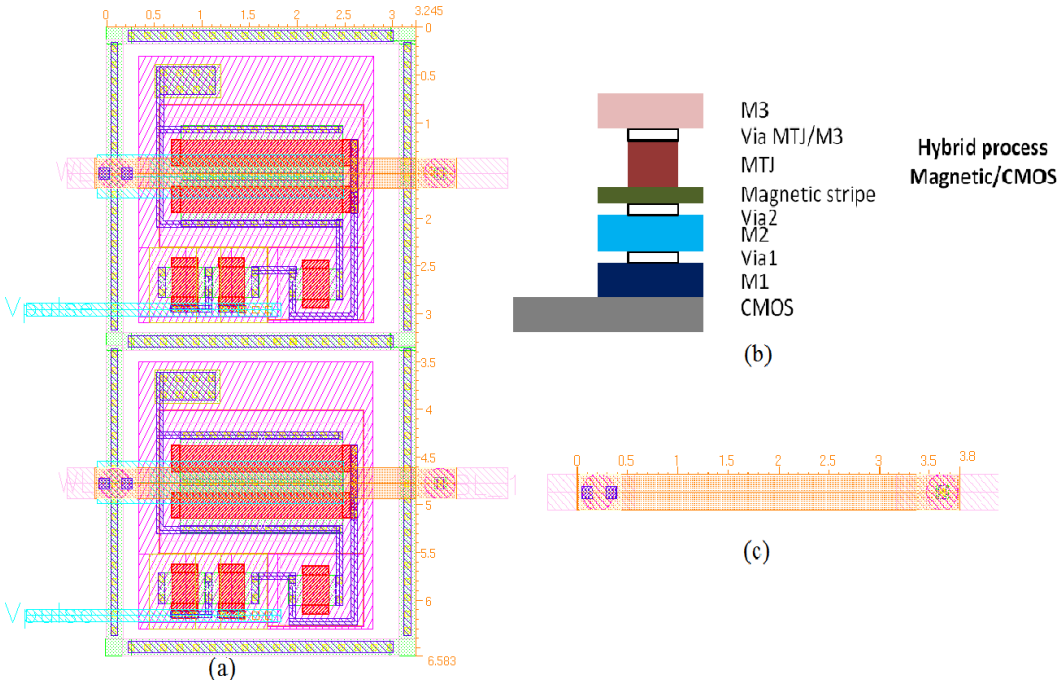


Figure 3.27 (a) 2 words of racetrack memory: each magnetic nanowire length reaches 3.8 µm which correspond to 95 bits per nanowire at maximum. (b) Vertical implementation of magnetic layers above CMOS process. (c) Zoom on the racetrack memory mask including two circular PMA MTJs on the two edges and a magnetic nanowire.

3.6 Conclusion

In this chapter, we presented the compact modeling of PMA STT MTJ and PMA racetrack memory. MTJ is the basic element of MRAM; racetrack memory is an emerging spintronic concept based on CIDW motion. They are considered the promising technologies for future logic and memory applications. The compact models of them are necessary for IC design of hybrid spintronics/CMOS systems.

In each modeling, numbers of physical theories (e.g. oxide barrier resistance model, STT static model, STT dynamic model, STT stochastic model and DW motion model) and material parameters have been integrated into the models to achieve excellent agreement with experimental results. In addition, the implementation using the Verilog-A language, which is compatible with standard CMOS computer-aided design tools (e.g. Cadence platform), provides an easy interface and allows this model to be easily extended to other device structures. Single cell simulations (e.g. DC, transient and Monte-Carlo simulations) were also performed to validate the static, dynamic and stochastic behaviors of PMA STT MTJ and PMA racetrack memory respectively.

These compact models can be very useful for hybrid spintronics/CMOS circuit designs by allowing for direct performance analysis, for instance, area, speed and energy. In the following chapters, we will use these compact models and CMOS technology design kit to investigate and analyze more complex hybrid logic and memory circuits.

Chapter 4 Hybrid spintronics/CMOS circuit design

4.1 Introduction

Beyond the simulations for validating the individual compact model as shown in the above chapter, our ultimate object is to design and analyze more complex hybrid spintronics/CMOS logic and memory circuits, for example, magnetic full adder (MFA) and content addressable memory (CAM). By applying the compact models and CMOS design kit, power, speed and area performances of the hybrid circuits can be analyzed through mixed simulations to obtain the best design for specific applications. In order to demonstrate the compatibility of these compact models for different technology node, 40 nm and 65 nm CMOS design kits have been used.

In this chapter, various examples will be presented. At first, sensing and writing circuits are respectively addressed. They are basic elements for detecting and switching the states of MTJ in all the magnetic logic and memory circuits. With respect to the sensing circuit built on sense amplifier allowing low power and low area, the writing circuit dominates power and area of entire hybrid logic and memory circuit. Reliability issues of them are also studied by taking the stochastic behavior into account.

A magnetic processor based on spintronic devices is actually expected to overcome the issue of increasing standby power and dynamic power dedicated to leakage current and data moving. MFA, a basic element to build low-power high-density arithmetic/logic unit for magnetic processor, is investigated as an example of magnetic logic circuits. We propose the MFAs based on PMA STT MTJ and PMA racetrack memory and compare their performances with those of CMOS based full adder.

The CAM is a special computer memory, which is expected to provide fast data access and high density. It is widely used in mobile, internet routers and processors that require ultra-high search speed. The mainstream CAMs suffer from high power and large area as its conventional structure is composed of numerous large-capacity static random access memory (SRAM) blocks in order to provide fast data access. We present a design of NOR-type CAM based on DW motion in PMA racetrack memory. The CMOS switching and sensing circuits are globally shared to optimize the cell area; the complementary dual nanowire allows the local sensing and a fast data search while keeping a low power.

4.2 Sensing circuit

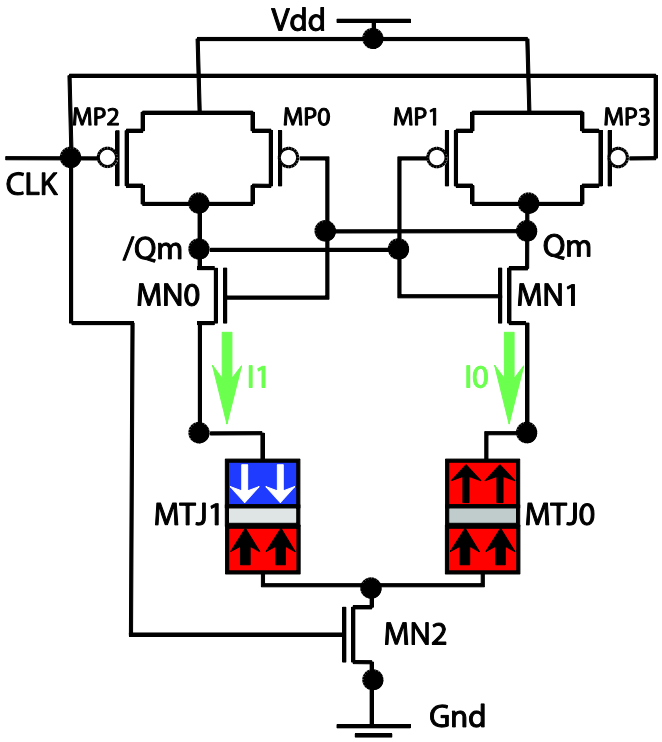


Figure 4.1 Pre-charge sense amplifier (PCSA) for MTJ state detection and amplification to logic level.

Due to the TMR effect, MTJ presents the property of resistance difference for different states. This resistance property allows MTJ to be compatible with CMOS sense amplifier circuit that detects the MTJ's configuration and amplifies them to logic level. Among various sense amplifiers, pre-charge sense amplifier (PCSA) is proposed to provide not only the best tradeoff between sensing reliability and power efficiency, but also high-speed performance [126]. Thereby we focus on PCSA and apply it for the hybrid logic and memory circuits investigated in this thesis.

4.2.1 Structure of sensing circuit

The PCSA circuit (see Figure 4.1) consists of a pre-charge sub-circuit (MP2-3), a discharge sub-circuit (MN2) and a pair of inverters (MN0-1 and MP0-1), which act as an amplifier. Its two branches are normally connected to a couple of MTJs with complementary states. It operates in 2 phases: "Pre-charge" and "Evaluation". During the first phase, "CLK" is set to '0' and the outputs ("Qm" and "/Qm") are pulled-up to "Vdd" or logic '1' through MP2-3 while MN2

remains off. During the second phase, "CLK" becomes '1', MP2-3 are turned off and MN2 on. Due to the resistance difference between the two branches, discharge currents are different. The lower resistance branch will be pulled-down to reach more quickly the threshold voltage of the transistor (MP0 or MP1), at that time, the other branch will be pulled up to "Vdd" or logic '1' and this low-resistance branch will continue to drop to "Gnd" or logic '0'.

4.2.2 Sensing operation and performance

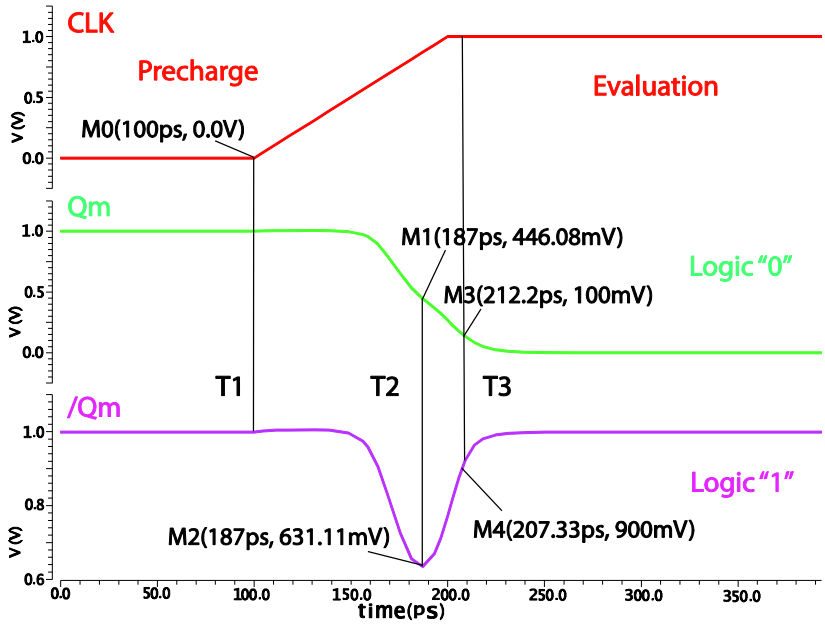


Figure 4.2 PCSA sensing operation in the case of MTJ0 with "parallel" state and MTJ1 with "anti-parallel" state.

Figure 4.2 shows a sensing operation of PCSA in the case of MTJ0 with "parallel" state and MTJ1 with "anti-parallel" state. Before the moment "T1", it is "Precharge" phase. Both outputs are pulled up to 1 V. Then the "Evaluation" phase starts from "T1", two branches begin to discharge after a small delay considering the rising time of "CLK" signal. At the time "T2", the branch "Qm" reaches the threshold firstly and this branch will continue to decrease to '0'. At the same time, MP0 begin to work and recharge the complementary branch "/Qm" back to "1". This sensing operation is so speedy. From the figure, we can find the whole process costs a sensing delay less than 100 ps. From the point of view of consumption energy, a sensing operation can only cost as low as 10 fJ. This high-speed and low-power feature makes PCSA suitable for the logic applications.

4.2.3 Reliability of sensing operation

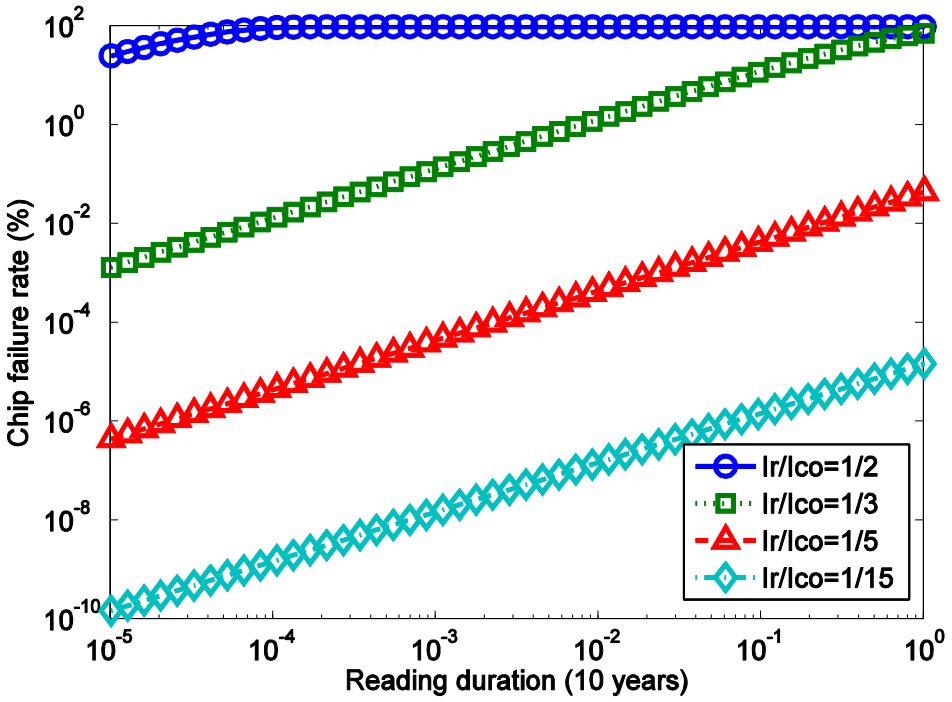


Figure 4.3 Dependence of chip failure rate on reading duration for different reading current.

We use PCSA circuit in the hybrid MTJ/CMOS design for the other reason: the read disturbance induced by sensing operations can be significantly decreased [167]. It is important for embedded STT-MRAM as it is an intrinsic nature and difficult to correct in logic circuit where complex error correction circuit (ECC) is prevent to ensure fast computing speed (e.g. 1 GHz). The read disturbance can be regarded as the unexpected switches during the sensing operation. As the sensing current is usually much lower than the critical current, the switch probability can be described by Néel-Brown model. If there are N bits of MTJs in the chip, the chip failure rate F_{chip} can be calculated by Eq. 4.1.

$$F_{chip} = 1 - \exp\left[-N \frac{\tau}{\tau_0} \exp\left(-\Delta\left(1 - \frac{I_r}{I_{C0}} \right) \right) \right] \tag{4.1}$$

where N is the number of bits per word, I_r is sensing current, I_{C0} is the critical current, τ is the read duration and τ_0 is the attempt period. As shown in Figure 4.3, lower I_r and shorter τ can

reduce greatly the chip failure rate for the STT-MRAM with the same thermal stability factor Δ =40.

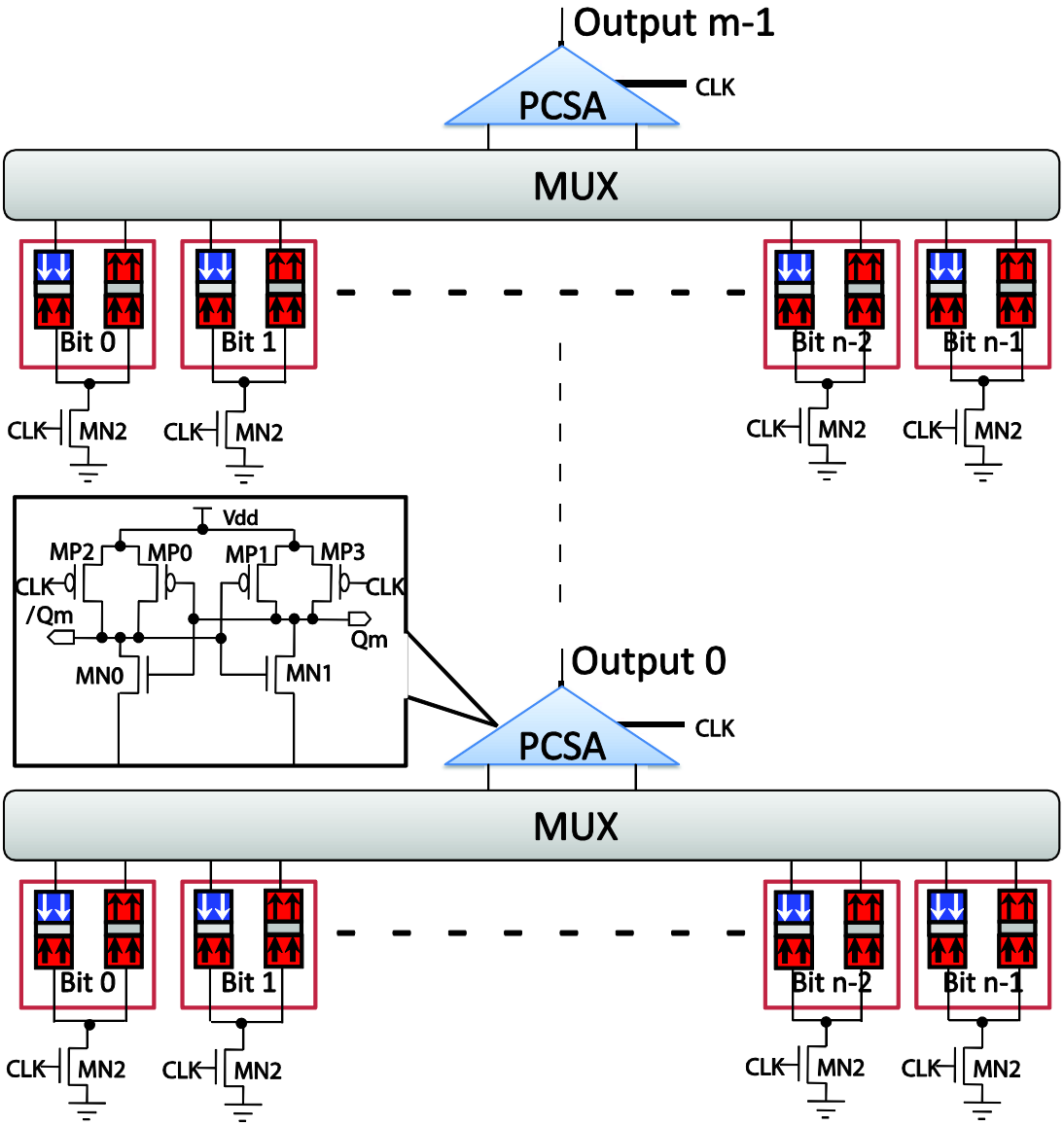


Figure 4.4 Schematic for 16k-bits PCSA sensing circuit (m=16, n=1k).

In reality, numbers of words of memories (e.g. 1k) normally share a sense amplifier. As shown in Figure 4.4, a 16k-bit (1k words of 16 bits) PCSA sensing circuit has been investigated. This enormous parallel structure leads to a huge capacitance, which drives the current pulse through the MTJ. As a result, an evaluation phase lasts almost 10 ns. By taking the effect of stochastic behavior into account, Monte-Carlo simulations after 1 µs of sensing duration (i.e. 1000-time

sensing operations) has been performed (see Figure 4.5). We found that the 33 errors occurred among 100 simulations. They are caused by either mismatch and process variations of CMOS part or STT stochastic behaviors of MTJ, or sometimes by both of them.

To identify the impact from each of them, we also performed Monte-Carlo simulations for sensing circuit with only mismatch and process variations. We found that the read disturbance was ~11%. Compared with the result presented in Figure 4.5, we can conclude that the stochastic behavior of MTJ greatly increases the error probability for a long-pulse current, and that this PCSA is not suitable for very large memory systems.

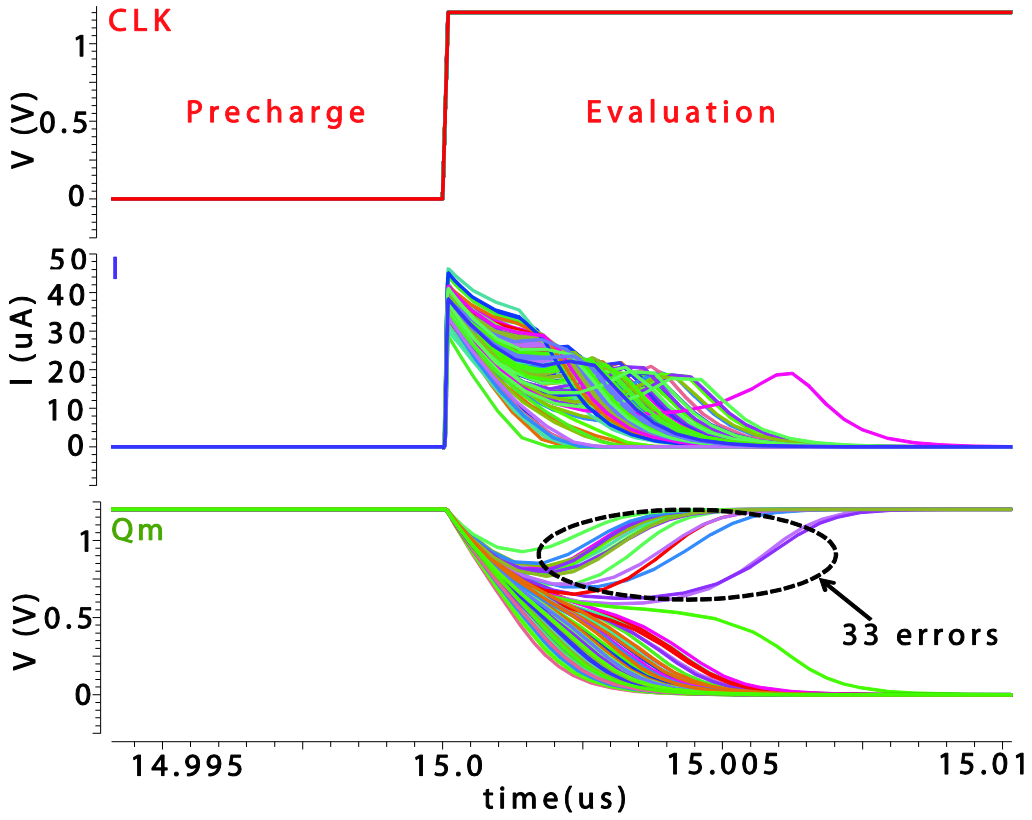


Figure 4.5 Monte-Carlo simulation of a 16k-bit PCSA circuit.

4.3 Writing circuit

Accroding to STT switching mechanism, a bi-directional current is required to switch the magnetization in free layer of MTJ. In order to achieve high-speed logic design, high current is required to ensure the speed. In contrast to the low power and low area of the sensing circuit, the writing circuit for PMA STT MTJ occupies the main area and power of the whole circuit. As a result, the study on writing circuit is of importance to hybrid MTJ/CMOS circuit design.

4.3.1 Structure of writing circuit

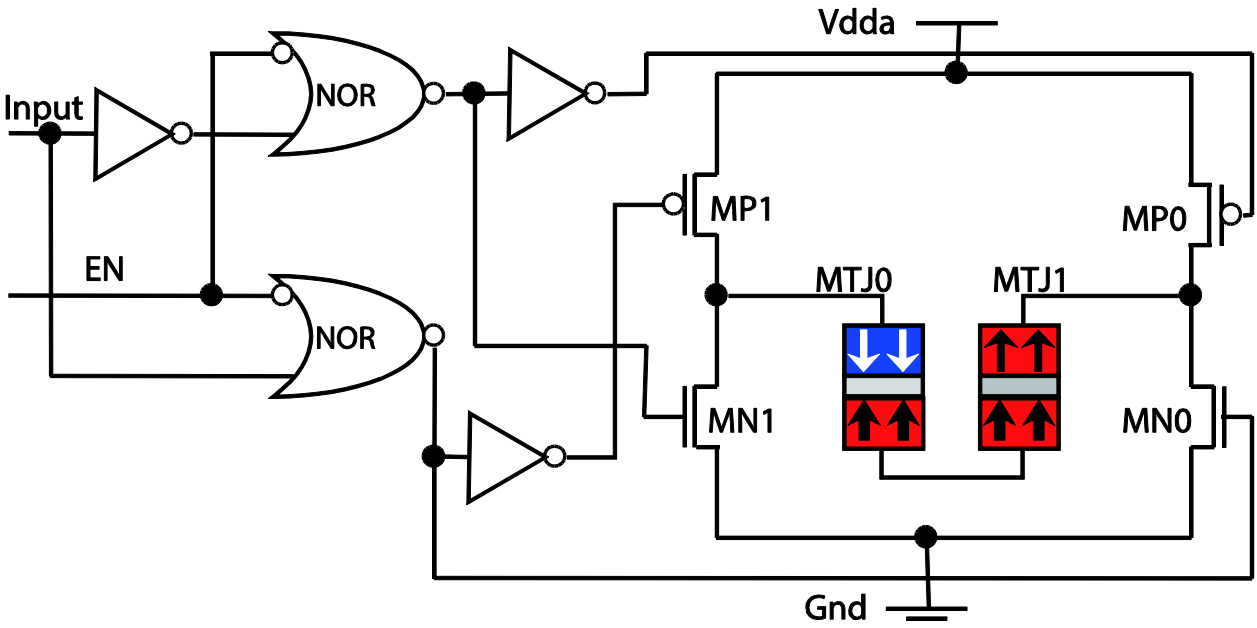


Figure 4.6 Full writing schematic for STT writing approach, which is composed of two modified inverters and logic control circuits.

In concert with the sensing circuit for a couple of MTJs with complementary states, a writing circuit to generate the bi-directional current for switching a couple of MTJs is designed as Figure 4.6. Two NMOS (MN0-1) and two PMOS (MP0-1) transistors construct the main circuit. Each time one NMOS and one PMOS are always left open and the others closed, which creates a path to make the current pass from “Vdda” to “Gnd”. Through two NOR and three NOT logic gates, the signals “Input” and “EN” control respectively the current direction and activation. Normally, it requires a “Vdda” higher than “Vdd” for logic operations to avoid the area overhead in the write circuit.

In order to generate the maximum current flowing through the couple of MTJs, both the transistors (one PMOS and one NMOS) should operate in their linear region above the threshold voltage V_{TH} to obtain the relatively lower resistances. In this case, they should satisfy the conditions: $V_{DS} << 2(V_{GS} - V_{TH})$ for NMOS and $V_{DS} >> 2(V_{GS} - V_{TH})$ for PMOS. Their resistances, R_{on} and R_{op}, can be approximately expressed by Eqs. 4.2-4.3, and the generated current can be obtained through the Eq. 4.4,

$$R_{on} = \frac{1}{\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{TH})} \tag{4.2}$$

$$R_{op} = \frac{1}{\mu_p C_{ox} \frac{W}{L}(V_{SG} - |V_{TH}|)} \tag{4.3}$$

$$I_{write} = \frac{Vdda}{R_p + R_{ap} + R_{on} + R_{op}} \tag{4.4}$$

where μ_n is the electron mobility, μ_p is the hole mobility, C_{ox} is the gate oxide capacitance per unit area, W is the channel width, L is the channel length, V_{GS} is the gate-source voltage.

4.3.2 Writing operation and performance

By using the PMA STT MTJ compact model, we simulate a writing operation including anti-parallel to parallel switching and parallel to anti-parallel switching (see Figure 4.7). We can find that the writing operation is not activated until the signal "EN" is set to '1'. The states of the couple of MTJs remain always opposite and the switching direction follows the signal "Input".

From Eqs. 4.2-4.4, we find that the most efficient method to improve the current value is by increasing W, but this leads to significant area overhead. Figure 4.8 shows a study of area, speed and energy performance for this circuit. Here, only the area of four transistors (MN0-1, MP0-1) has been taken into account as the area of logic control circuit is the same for different simulation and is often in the minimal size. A strong dependence between area and speed can be found, especially when the area is smaller than 0.2 μm^2. The speed improvement becomes less

significant for larger areas and saturates at ~1.1 GHz, which is different from the 2 GHz obtained with a single cell simulation in the previous chapter. There are two reasons for explaining this: first, "Vdda" is set to 2 V as 2.2V is the breakdown limit for CMOS technology node [168]; second, there are a couple of MTJs, so the bias-voltage for each one cannot be larger than 1 V as there is also bias on the transistors in the circuit.

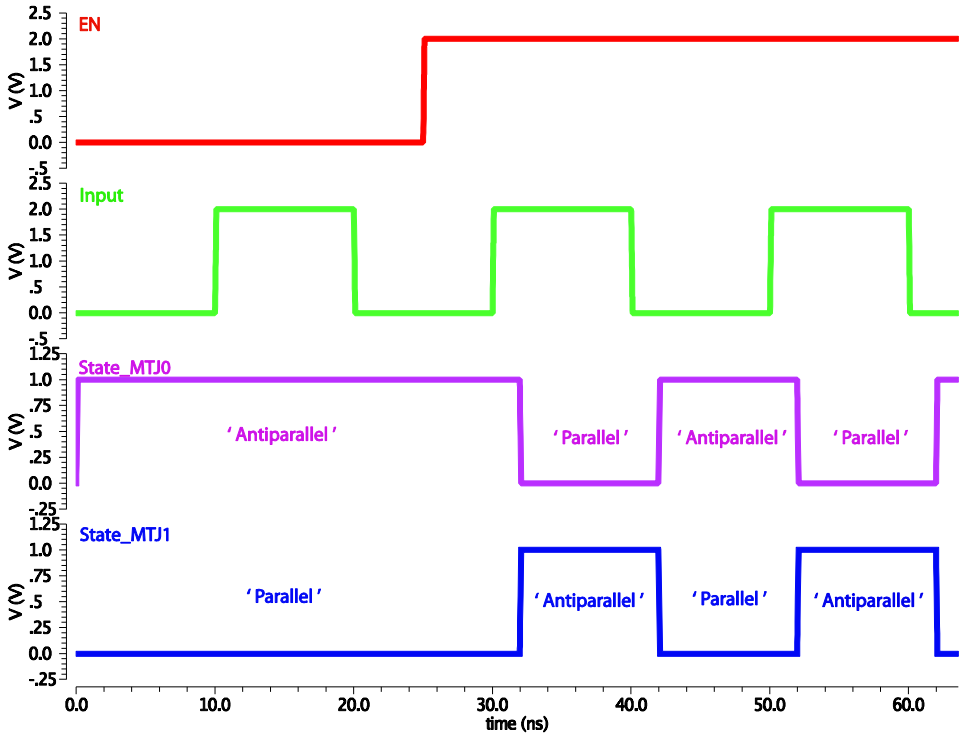


Figure 4.7 Writing operation for a couple of MTJs with complementary states. The signals "Input" and "EN" control respectively the current direction and activation.

The energy of each switching operation has been calculated with Eq. 4.5 based on the simulation results. We also find a turning point, ~0.1 μm^2, below which the energy will be increased rapidly with a smaller area due to the extremely long switching duration as the current I_{write} approaching to the threshold I_{c0}. Contrarily, the energy is nearly the same for whatever the size larger than ~0.1 μm^2. This is firstly because that the writing current and speed approach to be saturated. Even if the writing current can increase continuously, from Eq. 3.18, the current is inversely proportional to the switching duration when the current is much higher than the critical one. Therefore the energy will inevitably be saturated for a high writing current.

$$E_{operation} = Vdda \times I_{write} \times Duration \tag{4.5}$$

The region around the crossing point of the two curves (point 'A' in Figure 4.8) can be localized. It can be considered as a good tradeoff among the area (~0.096 μm^2 or 30 F^2), power (1 pJ) and speed (~500 MHz) performance of this switching circuit, and be suitable to build up both logic and memory chip. This simulation can also help to analyze the circuits with special requirements like 800 MHz operating frequency.

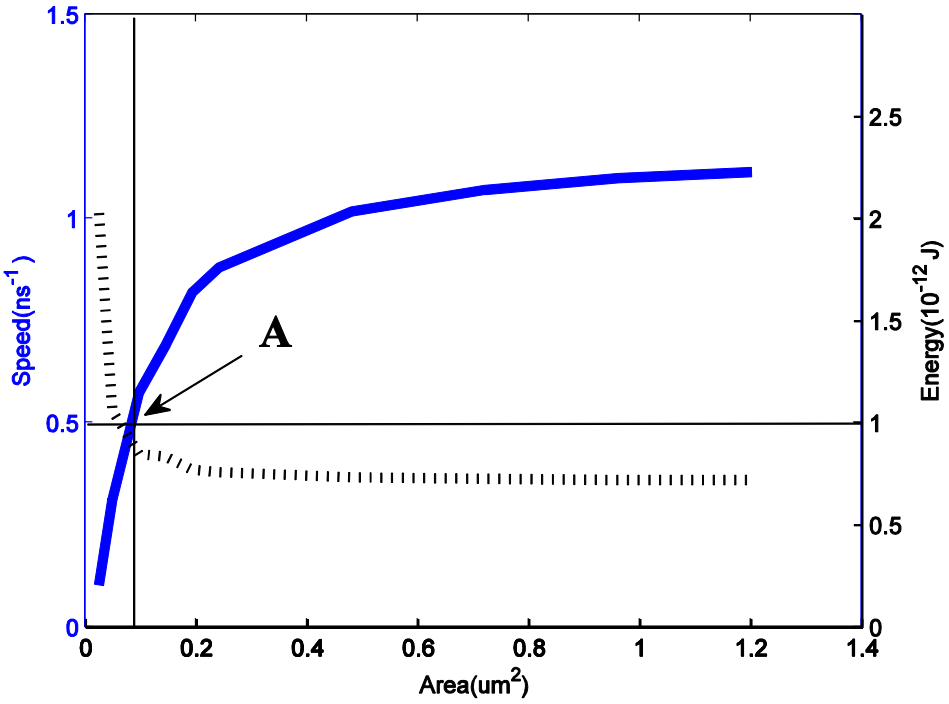


Figure 4.8 Dependence of circuit switching speed (solid line) and energy dissipation (dotted line) versus die area with four transistors (MN0-1, MP0-1).

4.3.3 Reliability of writing operation

For the advanced node below 90 nm, high reliability is becoming more and more crucial for the IC design. Thanks to the integration of STT stochastic behavior into this model, an overall reliability investigation becomes possible. Figure 4.9 shows the statistical Monte-Carlo simulations of 100 complete writing operations using the writing circuit shown in Figure 4.6 (however, instead of a couple of MTJs, there is only one MTJ connected in the circuit for this part of study). The complete writing operation includes the switching from parallel to anti-

parallel and from anti-parallel back to parallel. Similarly to the case of the sensing circuit, the writing current at each write event is different because of the mismatch and process variation of CMOS part. The switching delay times vary randomly due to the stochastic behavior of MTJ cell. Since writing current is normally larger than sensing current, the stochastic effect in writing operation is relatively weaker than that in sensing operation. This can be proven by Figure 4.9, the variation of every event is not so enormous.

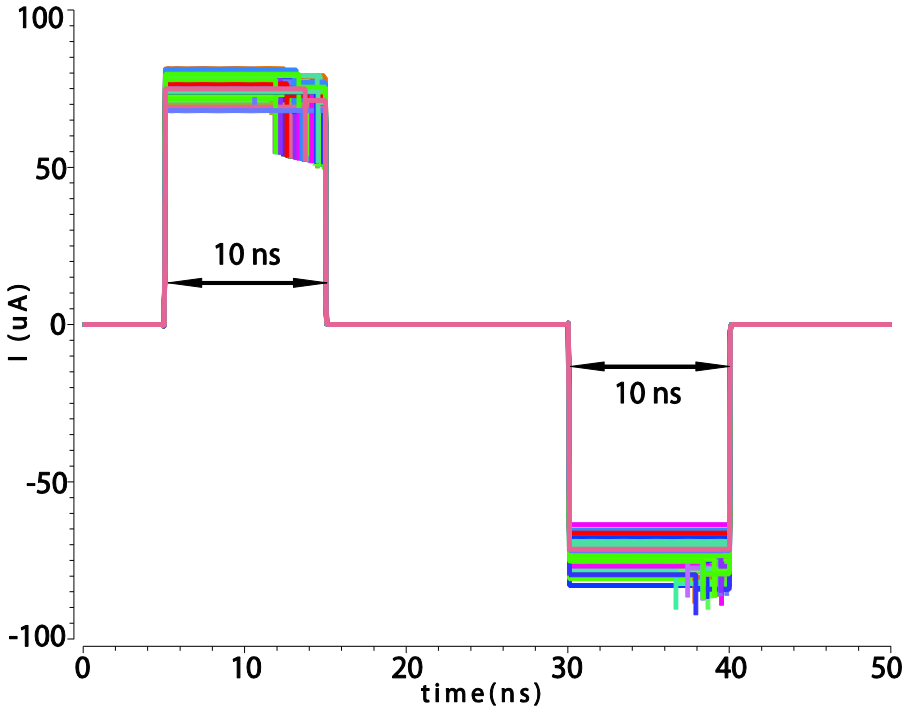


Figure 4.9 Monte-Carlo simulation of a whole writing operation implemented by the writing circuit.

Writing current magnitude and pulse duration are two key factors for the writing operation. As mentioned above, the writing current magnitude is dependent on the die area of writing circuit. We then perform the Monte-Carlo simulations for different writing pulse durations (5 ns, 10 ns and 20 ns) to observe the dependence of writing Bit Error Rate (BER_W) versus die area of writing circuit (four main transistors: MP0-1 and MN0-1). The simulation results shown by Figure 4.10 demonstrate their tradeoff relation: the increase of area can improve the BER performance. The reason is that a larger circuit allows larger write current, which in average reduces the time required to switch. For a given pulse duration, this increases the switching probability. Correlatively, it is observed that a longer pulse can also increase the reliability, which confirms the explanation mentioned above.

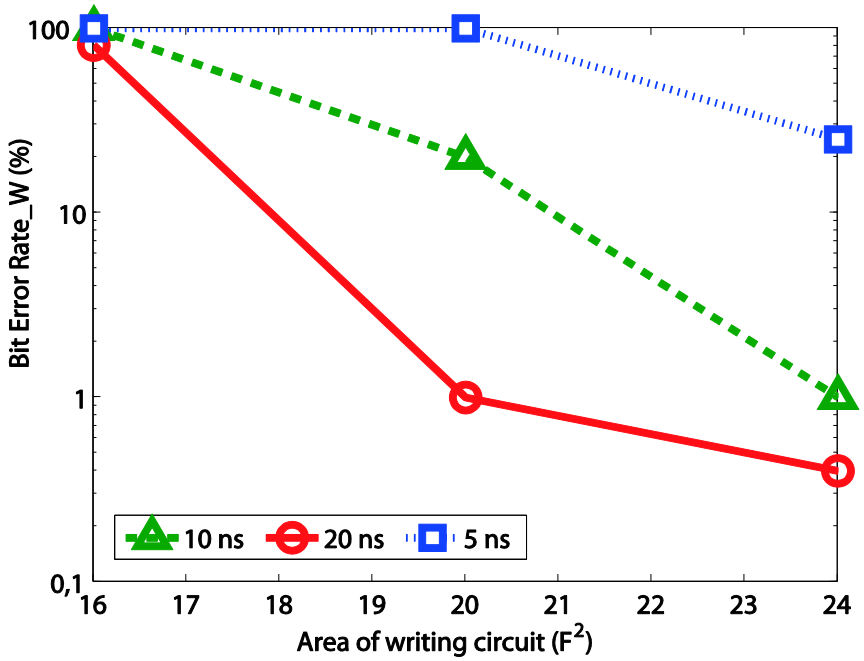


Figure 4.10 Dependence of writing Bit Error Rate (BER_W) versus die area of writing circuit.

4.4 Magnetic Full adder (MFA)

Aiming to overcome the issue of rising standby and dynamic power, magnetic processor based on spintronic devices is thus expected. Since addition is the basic operation of the arithmetic/logic unit of any processors, magnetic full adders (MFAs) attract a lot of attention and several designs based on diverse technologies are proposed in the last years [134, 136, 169]. Here, we present a 1-bit MFA based on PMA STT MTJ (STT-MFA) and a multi-bit MFA based on PMA racetrack memory. By comparing with CMOS based full adder, they show encouraged performances in terms of power consumption and area while keeping a relatively high speed.

4.4.1 1-bit MFA based on PMA STT MTJ (STT-MFA)

4.4.1.1 Structure of STT-MFA

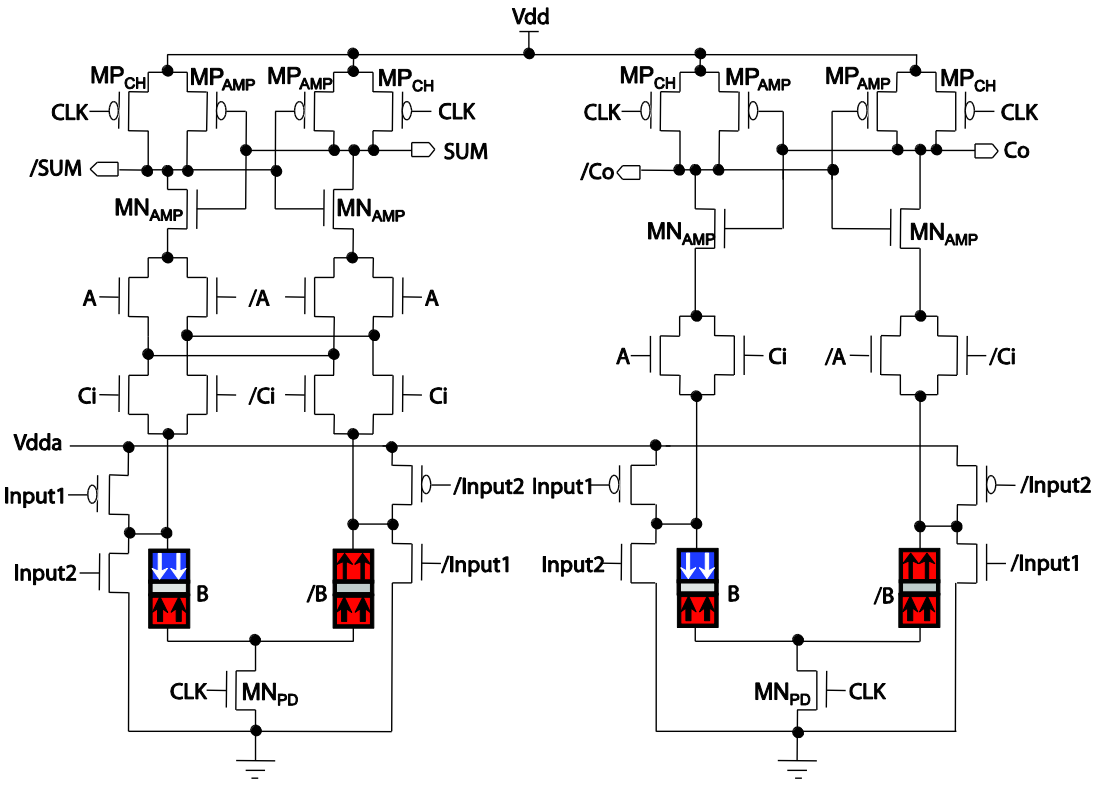


Figure 4.11 STT-MFA architecture with "*SUM* " (left) and output carry "C_o" (right) sub-circuits, "A" is volatile data for computing, "B" is non-volatile data using as quasi-constant.

Figure 4.11 shows a 1-bit STT-MFA circuit, which is based on the generic structure introduced in

the Chapter 2 (see Figure 2.16). To evaluate the logic function, PCSA circuit is used. The inputs are "A", "C_i" and "B", and the outputs are "SUM" and "C_o". Among them, the input "B" relates to non-volatile storage PMA STT MTJ. The MOS tree is designed according to Eqs. 4.6-4.9 and the truth table shown in Table 4.1.

$$SUM = A \oplus B \oplus C_i = ABC_i + A\overline{B}\,\overline{C_i} + \overline{A}B\,\overline{C_i} + \overline{A}\,\overline{B}C_i \tag{4.6}$$

$$\overline{SUM} = AB\overline{C_i} + \overline{A}\,\overline{B}\,\overline{C_i} + \overline{A}BC_i + A\overline{B}C_i \tag{4.7}$$

$$C_o = AB + AC_i + BC_i \tag{4.8}$$

$$\overline{C_o} = \overline{A}\,\overline{B} + \overline{A}\,\overline{C_i} + \overline{B}\,\overline{C_i} \tag{4.9}$$

Table 4.1 Truth table of "SUM" and "Co" logic gate for MFA

A	B	C	SUM	Co
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	0
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

For "SUM" logic, the MOS tree corresponds directly to the logic relationship among the inputs "A", "C_i" and "B", we can simply adapt it to the general structure with a couple of complementary PMA STT MTJ. However, it is a little difficult for "C_o" logic as there is the term AC_i in the logic function Eq. 4.8 and we cannot adapt the schematic to the general "logic-in-memory" structure. It can be inferred that the impact of the term AC_i on the resistance is equivalent to a sub-branch connecting PCSA and the discharging transistor (*MN2* in Figure 4.1). Table 4.2 exhibits the true table and the resistance configuration of "C_o" logic. R_{OFF} and R_{ON} are respectively the close and open resistances of MOS transistor. R_L and R_R are respectively the whole resistance of the left and right branch of PCSA. We can find that whatever the value of "A"

and "C_i", the sub-branches AC_i and $\overline{A}\overline{C_i}$ have no impact on the output. If "A" and "C_i" are different, the resistances of the two sub-branches are the same. If they are the same, their comparison corresponds to that of R_L and R_R in the condition of $R_{ON} > R_{AP}$, which is always true for PMA STT MTJ under present technology condition. This allows the term AC_i to be deleted from Eq. 4.8 and we can obtain the "C_o" logic circuit shown in Figure 4.11.

The PMA STT MTJs connect serially with a common central point. In order to program MTJs, we use a writing circuit composed of pass transistors, which are connected respectively to the bottom and top electrodes of the serial branch. In such a manner, as a control signal ("*Input1*" or "*Input2*") is activated, the first PMA STT MTJ noted "B" is put in high resistance state (R_{AP}) or low resistance state (R_P) while the second PMA STT MTJ noted "/*B*" is put in the complementary state R_P or R_{AP}.

It is noteworthy that there is neither capacitance for the data sensing and nor magnetic field for data programming in this new structure beyond the previous structures [134, 136, 169]. Therefore, this design allows efficient area minimization and is suitable for advanced fabrication nodes below 65 nm.

Table 4.2 Truth table and resistance configuration of "Co" for MFA

A	B	C	Resistance Comparison	Co	Sub-branch AC_i	Sub-branch $\overline{A}\overline{C_i}$
0	0	0	$R_L > R_R$	0	$2R_{OFF}$	$2R_{ON}$
0	0	1	$R_L > R_R$	0	$R_{OFF} + R_{ON}$	$R_{OFF} + R_{ON}$
0	1	0	$R_L > R_R$	0	$2R_{OFF}$	$2R_{ON}$
0	1	1	$R_L < R_R$	0	$R_{OFF} + R_{ON}$	$R_{OFF} + R_{ON}$
1	0	0	$R_L > R_R$	0	$R_{OFF} + R_{ON}$	$R_{OFF} + R_{ON}$
1	0	1	$R_L < R_R$	1	$2R_{ON}$	$2R_{OFF}$
1	1	0	$R_L < R_R$	1	$R_{OFF} + R_{ON}$	$R_{OFF} + R_{ON}$
1	1	1	$R_L < R_R$	1	$2R_{ON}$	$2R_{OFF}$

4.4.1.2 Simulation and performance analyses of STT-MFA

Figure 4.12 illustrates the transient simulation of 1-bit STT-MFA shown in Figure 4.11. It is performed by using PMA STT MTJ compact models introduced above and CMOS 40 nm design kit. The time-dependent behaviors of outputs ("*SUM*" and "C_o") confirm the logic functionality of full addition. For instance, for the operation "A" = '1', "B" = '0', "C_i" = '0', the result is '1' and no carry yields; for the operation "A" = '1', "B" = '0', "C_i" = '1', the result is '0' and the carry is '1'.

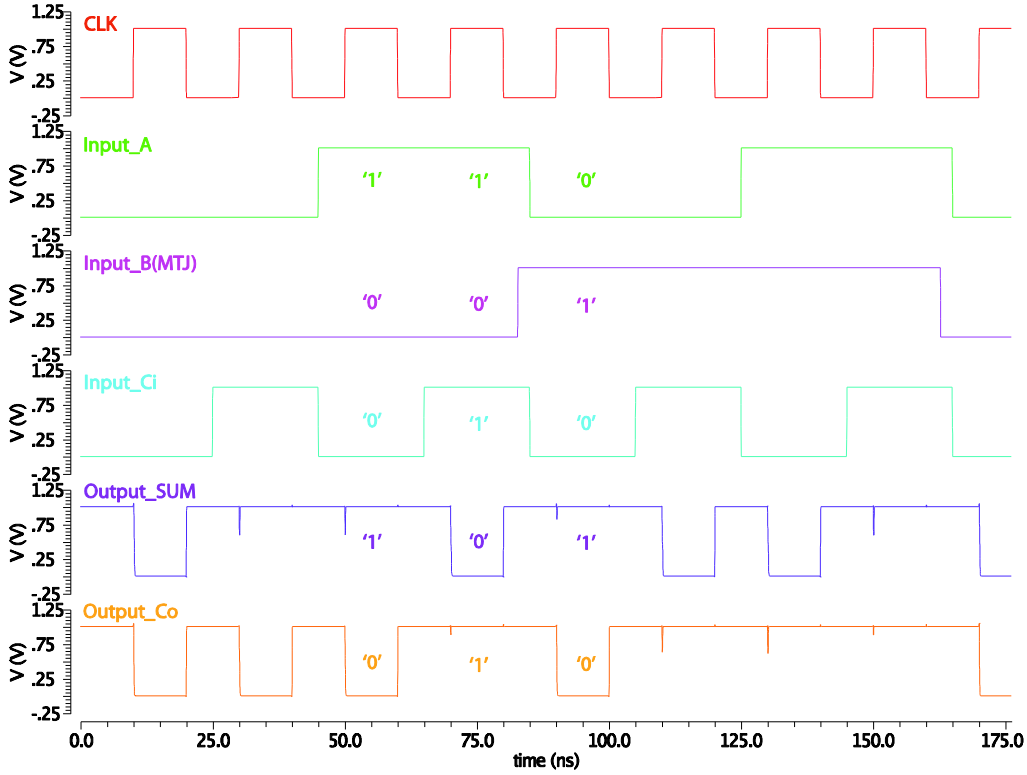


Figure 4.12 Transient simulations of 1-bit STT-MFA in 40 nm node.

Figure 4.13 emphasizes one sensing operation of this STT-MFA and shows the analog behaviors. It confirms the pre-charge, evaluation and amplification process described previously. Moreover, we find that the sensing delay of "*Output* _ C_o" (~127 ps) is shorter than that of "*Output* _ *SUM*" (~147 ps). This is due to the higher resistance of the branch associated with "*Output* _ *SUM*", leading to lower current and slower amplification.

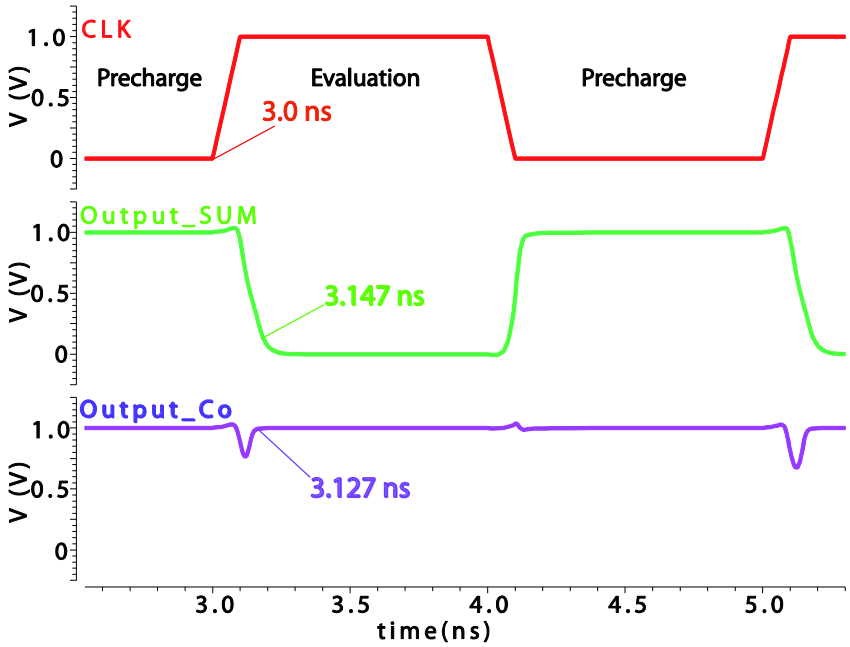


Figure 4.13 One sensing operation of the PCSA based STT-MFA: outputs are pre-charged as CLK is set to '0' and are evaluated as CLK is set to '1'.

The delay time and dynamic energy are generally two crucial parameters to evaluate the performance of computation system. We have studied the effects of three possible factors: the size of discharge transistor (MN_{PD} in Figure 4.11), PMA STT MTJ resistance-area product (*RA*) and TMR ratio. Figure 4.14 demonstrates the performance dependence of this STT-MFA in terms of delay time and dynamic power on the size of discharge transistor. We can find a tradeoff between the speed and power performance by varying the die area. A larger discharge transistor can drive a higher sensing current and faster amplification of PCSA circuit, but cost more energy.

Figure 4.15 shows the *RA* dependence for this STT-MFA. By decreasing *RA*, the delay time becomes shorter while keeping a relatively steady dynamic power performance. This confirms that the speed advantage of using low *RA*.

We also investigate the dependence between TMR ratio of PMA STT MTJ and STT-MFA performance. Figure 4.16 shows that faster speed is possible by increasing the TMR ratio while the dynamic energy changes slightly.

According to the above analyses, a PMA STT MTJ with lower *RA* and higher TMR ratio is expected to perform fast computation while keeping nearly the same dynamic energy. In the

recent experimental demonstration of the MTJ, a low *RA* (e.g. 5 $\Omega\mu m^2$) and high TMR ratio (e.g. 200%) can be achieved in PMA STT MTJ [40, 170-171].

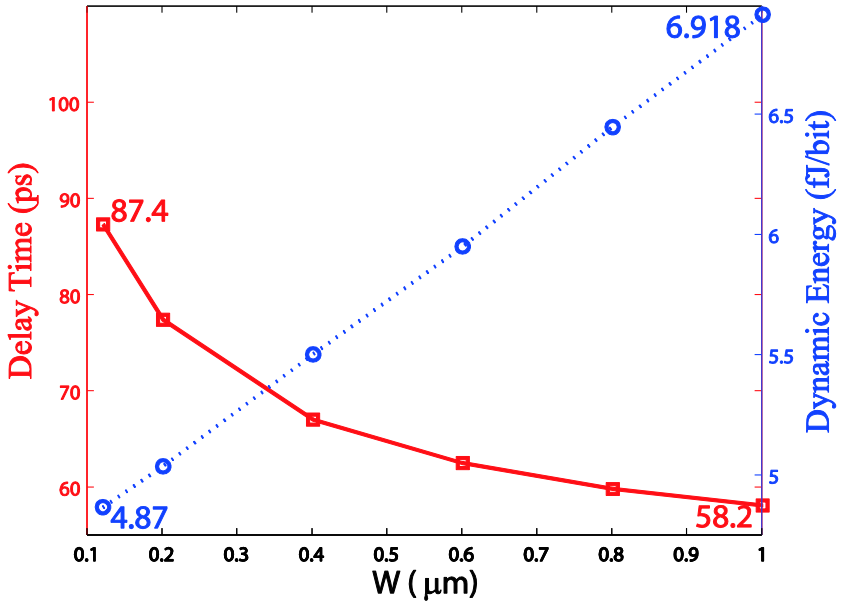


Figure 4.14 Dependence of delay time (red solid line) and dynamic energy (blue dotted line) on the width of discharge transistor for STT-MFA.

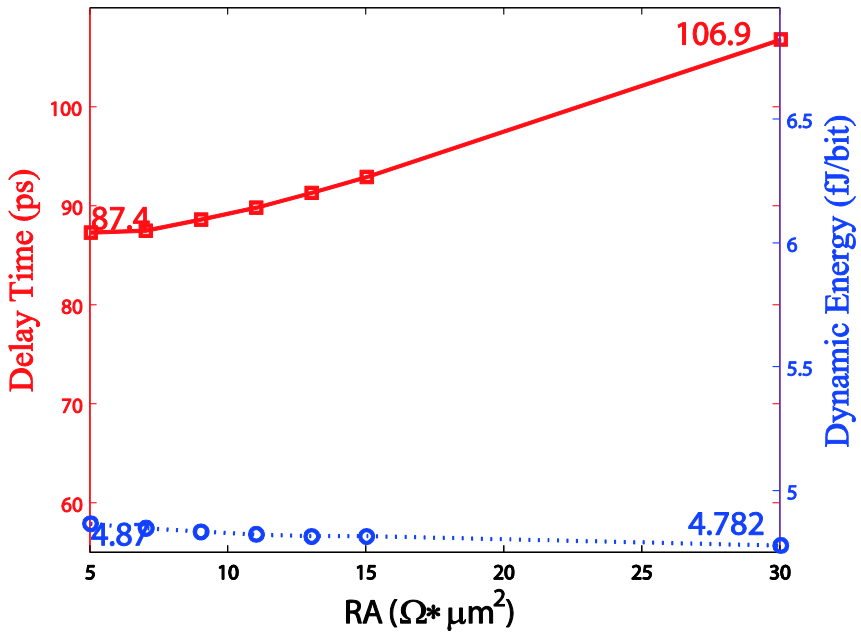


Figure 4.15 Dependence of delay time (red solid line) and dynamic energy (blue dotted line) on the resistance-area product (*RA*) of PMA STT MTJ.

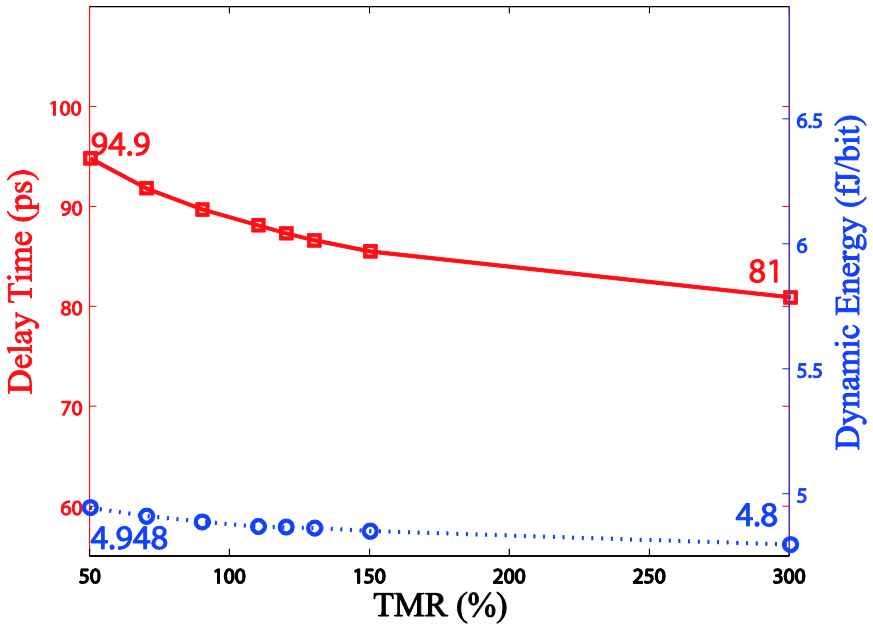


Figure 4.16 Dependence of delay time (red solid line) and dynamic energy (blue dotted line) on PMA STT MTJ TMR ratio for 1-bit STT-MFA.

4.4.1.3 Comparison with 1-bit CMOS-only full adder

We compare our proposed STT-MFA with conventional CMOS only full adder in terms of delay time, dynamic power, standby power, data transfer energy and die area (see Table 4.3). The CMOS-only full adder is taken from the standard cell library of STMicroelectronics 40 nm design kit. Two full latches are added to synchronize the outputs with clock signal.

In conventional computing architectures, logic and memory are completely separated [172]. In order to perform a logic operation, both the instruction and data need to be read from memory units (i.e. cache and main memories), and then moved to logic unit. The results are transferred back to the memory units after the computing. In the proposed MFA circuit based on "logic-in-memory" architecture, logic operations are processed directly with the magnetic data stored in MTJs and the addition result is written to other MTJs for the next operations. Long latency and high dynamic power due to data moving can be significantly economized. For example, the data transfer energy (~ 1 pJ/mm/bit @22 nm [122]) becomes much lower thanks to the shorter distance between memory and computing unit, which is about some μm or below in STT-MFA instead of some mm for CMOS only logic circuit.

Furthermore, thanks to the 3D integration of STT-MRAM, the die area of this design (38 MOS + 4 MTJs) is advantageous compared to those of the CMOS full-adder (46 MOS). However, its energy-delay product (EDP) exceeds that of a CMOS full-adder by approximately 10% since it takes more time for PCSA amplification process. Due to the non-volatility of PMA STT MTJ, the new chip can be powered off completely and this allows the standby power to be reduced significantly down to 0.75 nW [171]. Thereby, the STT-MFA can greatly reduce the consumption in a full computing system, especially for those normally in OFF state.

Another critical idea of this design is to use a programming frequency (e.g. 1 kHz) of STT-MRAM much lower than the computing frequency. Thereby, the switching power for non-volatile storage becomes insignificant to other power consumption in a full system. We can continue to reduce it by shortening the non-volatile data retention (e.g. 1 day). Moreover, the programming energy for the non-volatile data (bit "B" in Figure 4.11) can be reduced, following the area minimization [173] and new material development for MTJs (e.g. ~0.1 pJ/bit).

Table 4.3 Comparison of proposed 1-bit STT-MFA with CMOS only full adder

Performance	CMOS full adder (40 nm)	STT-MFA
Delay time	75 ps	87.4 ps
Dynamic power @500 MHz	2.17 µW	1.98 µW
Standby power	71 nW	<1 nW [31]
Data transfer energy	>1 pJ/bit	<1 fJ/bit
Die area	46 MOS	38 MOS + 4 MTJs

4.4.2 Multi-bit MFA based on PMA racetrack memory

PMA racetrack memory is distinguished as it can store and shift multiple bits of data through CIDW motion along a magnetic nanowire. This advantageous feature makes it possible to design a high speed and compact multi-bit serial MFA. Besides the information storage means, the logic evaluation circuits (SUM circuit and CARRY circuit) are similar to them of the 1-bit MFA circuit previously proposed in the section 4.4.1.

4.4.2.1 Structure of proposed multi-bit MFA

Figure 4.17 shows the detailed schematic of CARRY circuit of our proposed multi-bit MFA

based on PMA racetrack memory including MTJ writing circuit. "A" and "B" are multi-bit input data stored in different nanowires. Each data is designed to be stored in dual magnetic nanowires with exactly opposite configuration to minimize the variation between two complementary data (e.g. "A" and "$\overline{A}$") as the same I_{shift} is used in the dual nanowires to move the DWs [24]. At each rising edge of CLK, "C_o" and "$\overline{C_o}$" are evaluated through the PCSA circuit and become inputs of a writing circuit, which generates writing current I_{write} to reverse or just conserve the state of nucleation MTJs ("$C_{nucleation}$" and "$\overline{C_{nucleation}}$"). At each falling edge of CLK, propagating current I_{shift} induces the DW motion of all magnetic nanowires ("A", "$\overline{A}$", "B", "$\overline{B}$", "C", "$\overline{C}$") simultaneously, moves next magnetic domains under the read MTJ for next adding operation.

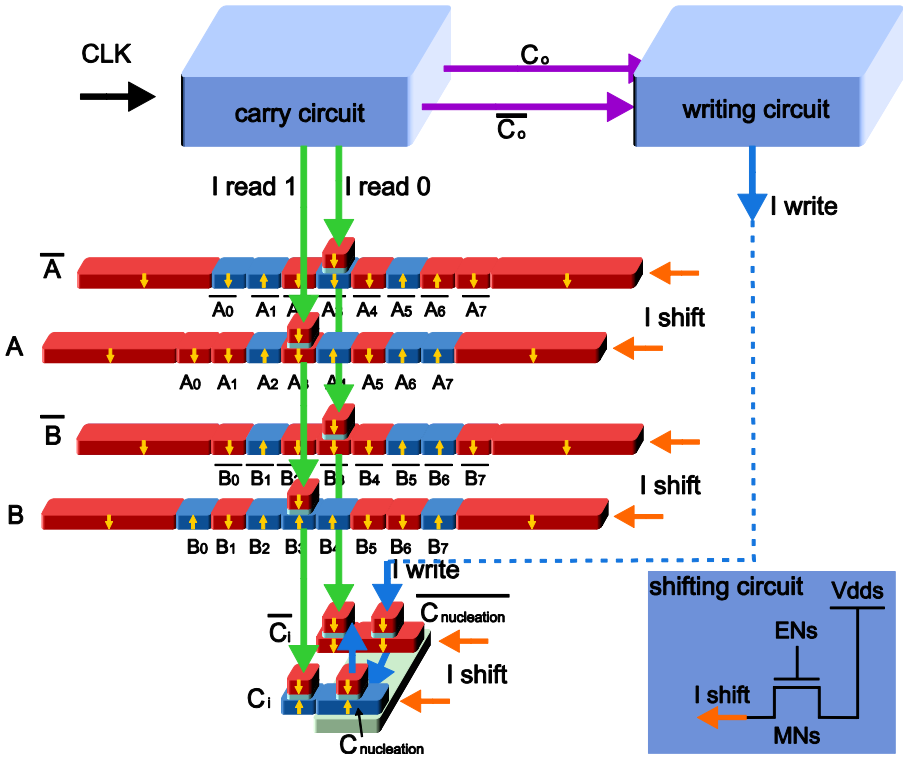


Figure 4.17 CARRY circuit of multi-bit MFA based on PMA racetrack memory including MTJ writing circuit.

The operation of SUM circuit is similar to that of CARRY: the SUM output and its complement are evaluated through the SUM PCSA circuit and become inputs of SUM writing circuit, which generates the writing current to write these values into the most significant bits of SUM nanowire.

These bits are then shifted in the way that at the end of all additions, the sooner the bit is calculated, the greater its weight is.

4.4.2.2 Simulation and performance analyses of proposed multi-bit MFA

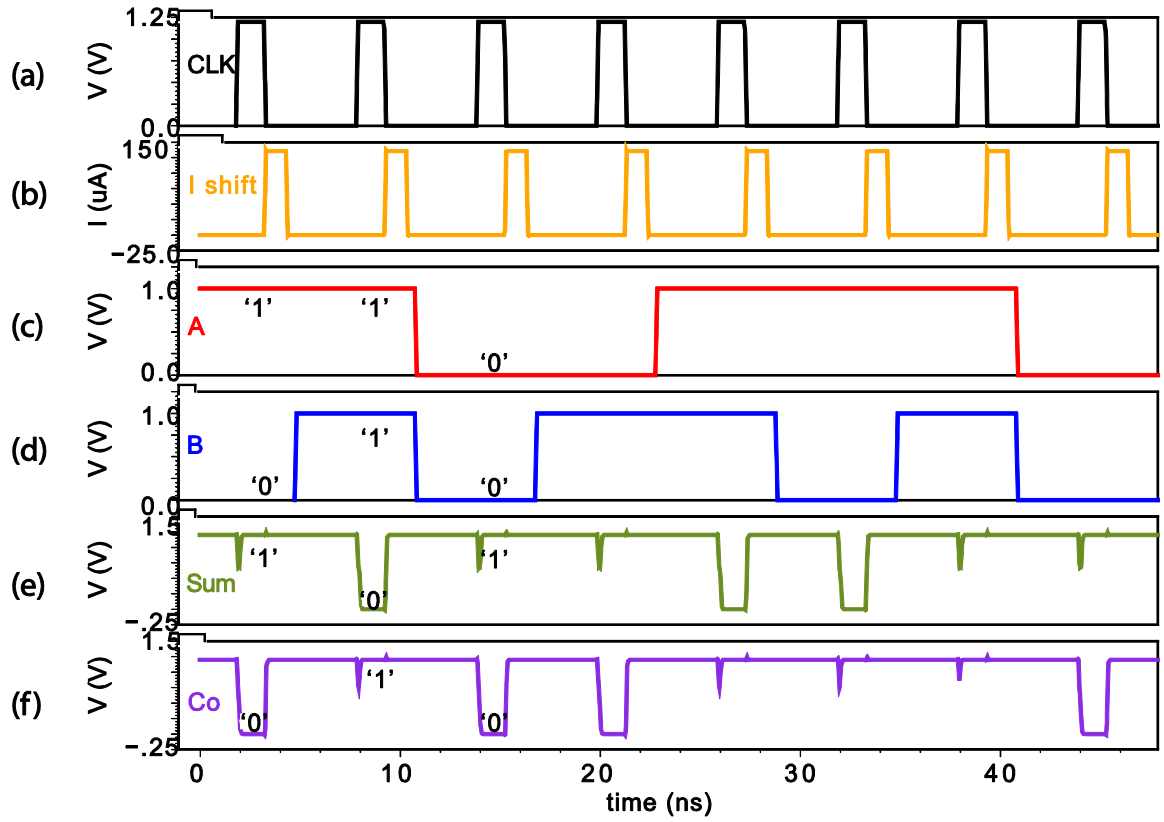


Figure 4.18 Transient simulation of proposed multi-bit MFA. (a) CLK signal (b) Data shifting current pulse I_{shift} (c) Input data "*A*" (d) Input data "*B*" (e) "*SUM*" (f) "C_0".

To validate the correct functionality, STMicroelectronics CMOS 65 nm design kit and compact model of PMA racetrack memory have been used to perform SPICE simulations for this multi-bit MFA. The transient simulation shows the addition operation of two random 8-bit words: "A"= "01110011" (Figure 4.18(c)) and "B"= "01011010" (Figure 4.18(d)). "CLK" (Figure 4.18(a)) drives PCSA circuit and "I_{shift}" (Figure 4.18(b)) induces DW motion in the magnetic nanowire. The outputs "*SUM*" (Figure 4.18(e)) and "C_o" (Figure 4.18(f)) are firstly pre-charged to logic '1' when "CLK"= '0' and are evaluated when "CLK" is set to '1'. The DW motions are implemented in the pre-charge phase in order to avoid the disturbance to the output evaluation. The serial addition is performed from the least significant bit and the simulation result "*SUM*" =

"11001101" and "C_o" = "01110010" confirms the correct operation of MFA.

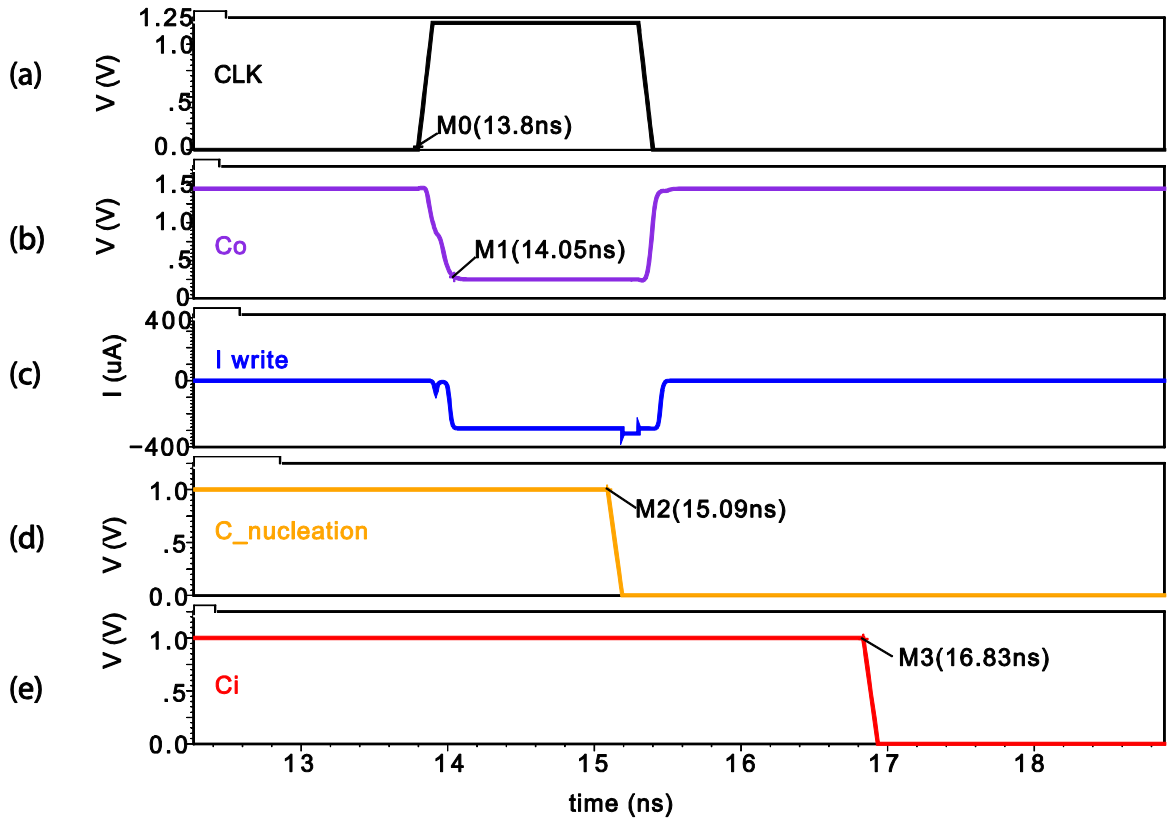


Figure 4.19 Transient simulation of CARRY storage (a) CLK signal (b) "C_0" (c) DW nucleation current I_{write} (d) State of DW nucleation MTJ (e) Carry in for next adding operation

Between two addition evaluations, there is a data transition process to achieve multiple bits operation. Figure 4.19 demonstrates the CARRY transition including DW nucleation and motion. Carry-out "C_o" (Figure 4.19(b)) is firstly pre-charged to "Vdd" before the time "M0", after the rising edge of "CLK" (Figure 4.19(a)), "C_o" is evaluated by the PCSA and becomes the input signal of writing circuit (Figure 4.19(b)). I_{write} is generated to nucleate DW in the magnetic nanowire (Figure 4.19(c-d)). I_{shift} is in the following activated to propagate the DW and replace the value of carry-in "C_i" with "C_o" at the time "M3" for next cycle of addition (Figure 4.19(e)).

We then analyze the performance of this multi-bit MFA in terms of delay and power dissipation. In our simulation, several parameters, such as the size of transistors and "Vdd", can affect greatly on them. A first look at the repartition of consumed energy in this MFA shows that the energy for

nucleation and propagation is of the same order of magnitude, and higher than that of data sensing or logic computing.

In order to propagate the magnetic domains to their next positions (shift 1 bit) using our compact model of PMA racetrack memory, we must supply a current I_{shift} in a period t_{pulse}. The period t_{pulse} is the necessary time for all magnetic domains to move from their current positions to their next positions. It corresponds to the propagation delay. In our model, this delay is inversely proportional to I_{shift}. Consequently, the propagation energy does not vary much (seeing that this energy is the integral of the product $V_{pulse} \times I_{shift} \times t_{pulse}$ and the power supply voltage V_{pulse} is kept invariable). Simulations show that energy needed for shifting all racetrack memories 8 bits is about 29 pJ.

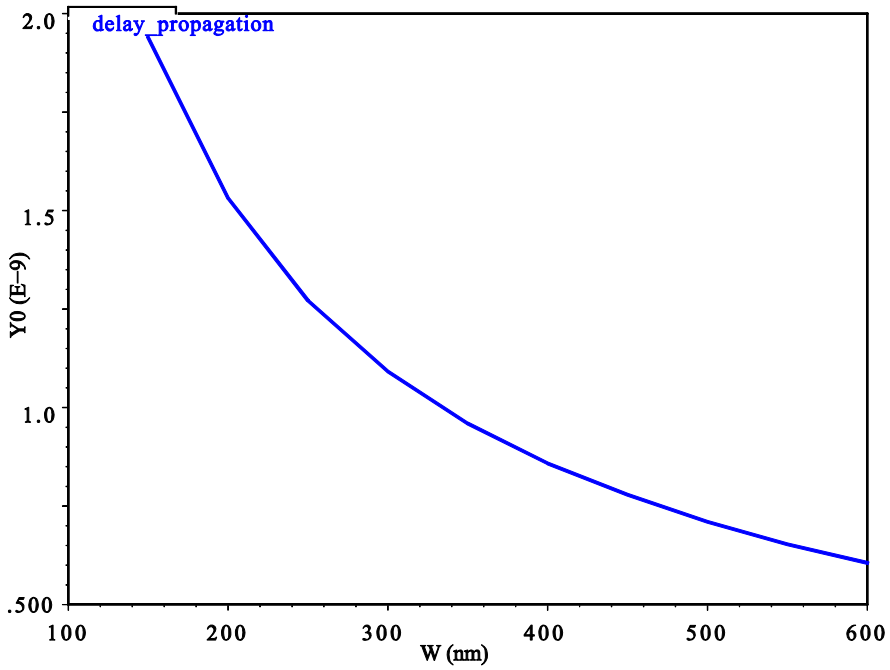


Figure 4.20 Dependence of domain wall propagation delay on the transistor width of propagation circuit.

Since the energy needed for propagation is almost invariable, we can reduce the propagation delay by increasing I_{shift}. Normally, one transistor based current source is used to generate the DW propagation current, thus the size of transistor determines the generated current in propagation circuit. Figure 4.20 shows the tradeoff dependence of propagation delay on the width of transistor MN2: the reduction of propagation delay at the cost of satisfying the area.

The writing circuit nucleates domain walls under the MTJ write head (e.g. MTJ0 in Figure 2.14) by passing through a bi-directional current I_{write}. I_{write} is proportional to both supply voltage V_{write} and transistors' size. V_{write} will be set as high as possible in order to minimize the size of transistors (MN0-1 and MP0-1) while keeping the switching current at fixed value. In this simulation, V_{write} is set to 2 V to avoid the breakdown of oxide barrier at 65 nm technology node. A higher I_{write} can reduce the switching delay, but increase the power consumption.

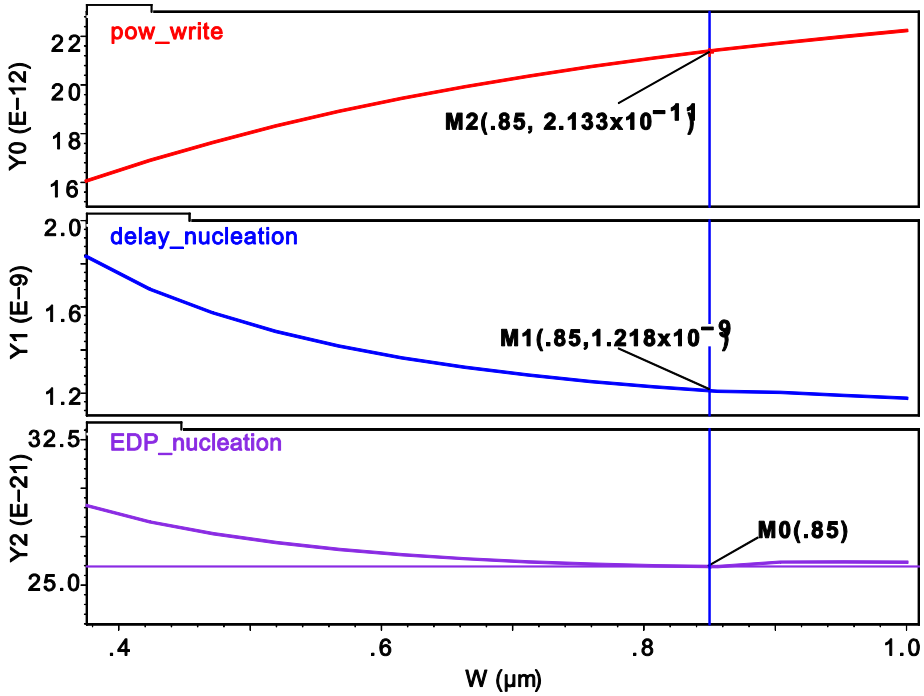


Figure 4.21 Dependence of energy consumption (pow_write), switching speed (delay_nucleation) and the energy-delay product (EDP_nucleation) versus the width of 4 transistors using in writing circuit.

A study of the tradeoff among the width of transistors, switching speed and power dissipation have been made to find out optimal operation point (see Figure 4.21). In this analysis, the width of transistors W is started at 0.35 µm because I_{write} is not high enough to switch the state of MTJ below this value. This curve shows that one can increase W to reduce the factor of merit EDP until the point "M0" (W=0.85 µm) and then it slightly goes up. Thereby the operating points should be chosen around the optimum (e.g. W=0.75-1.0 µm) to address different applications. For instance, the two markers "M1" and "M2" show that when W=0.85 µm, the switching power and latency are 21.33 pJ and 1.22 ns respectively. The switching current I_{write} equals to 291 µA

in this case.

4.4.2.3 Comparison with CMOS-only multi-bit full adder

In order to understand the advantages and disadvantages of this multi-bit MFA based on PMA racetrack memory, we compare its performance with that of a CMOS-only series adder (see Table 4.4), which uses a full adder taken from the library of STMicroelectronics 65 nm design kit.

For the comparison with CMOS only multiple bits full adder regarding writing and transferring data, we see that the chip area of the MFA based on PMA racetrack memory is significantly reduced. The proposed 8-bit MFA uses only 23 MOS transistors, 18 MTJs and 8 magnetic nanowires instead of 22 MOS plus 8×3 Flip-Flops (310 MOS transistors totally) for an 8-bit series CMOS full adder. Although the number of transistors decrease 13 times, the area reduction is about 4.5 times since the writing circuit and propagating circuit requires the transistors with 6.3 times and 3.3 times minimum width (0.135 µm@65 nm technology node), respectively. The total delay of one operation of the new MFA is ~2.1 ns, composed of DW nucleation (~1.2 ns), motion (~0.7 ns) and detection (~180 ps). It can be thus driven by a CLK frequency up to 470 MHz, which can be further increased with the feature size shrinking. This latency is of the same order with that of CMOS circuit (read time + transfer time + operate time + transfer time + write time). On the contrary, the proposed MFA consume 6 times dynamic energy more than the CMOS only full adder since energy needed for nucleation and propagation is still too large with current technology. However, we have not yet addressed the static energy in this comparison. Regarding that power must be supplied in order to maintain stored data in CMOS-only storage circuit, the proposed MFA does not require energy to conserve information thanks to its total non-volatility. This allows the circuit to be turned off safely in “idle” mode without data backup. All the operations can be retrieved instantly after power-on. This instant on/off capability promises to overcome completely the rising standby power issue due to leakage currents and could be very useful for normally-off systems [171].

It is important to note that for this non-volatile MFA, operations are performed directly with the data (“A”, “$\overline{A}$”, “B”, “$\overline{B}$”) stored in magnetic nanowires, which plays the role of shift registers. We do not take into account the writing circuits of “A” and “B” to keep the same comparison condition as the writing circuits of data are considered in the CMOS shift register

part, not in the adder. The number of writing circuit is then reduced to 2 for respectively SUM and CARRY circuits, which are shared by the eight bits.

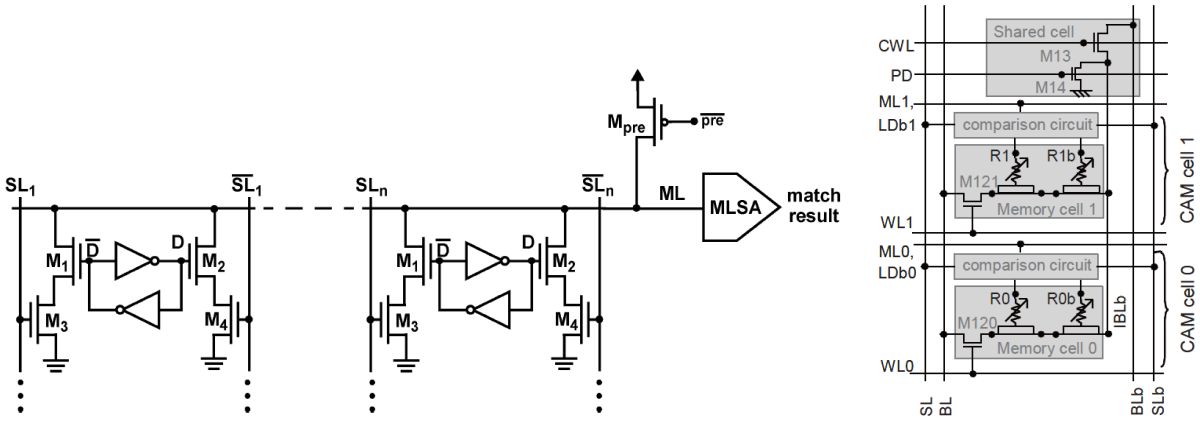
Table 4.4 Comparison of proposed 8-bit MFA based on racetrack memory with CMOS only full adder included transferring and writing data

Performance	CMOS full adder (65 nm)	Proposed MFA
Write time	200 ps	2 ns
Write energy	16 fJ/8 bits	(21.39+29) pJ/8 bits
Transfer time	~ ns	0
Transfer energy	8 pJ/mm (for 8 bits)	~0
Die area	310 MOS	23 MOS + 18 MTJs

4.5 Content addressable memory (CAM)

4.5.1 Introduction

Content addressable memory (CAM) is a computer memory that can output the address of serach data. It compares search data with stored data and returns the match location with its high-speed fully-parallel manner. Therefore it is widely used in mobile, internet routers and processors to provide fast data access and ultra-high density [174]. The mainstream CAMs are composed of large-capacity volatile SRAM blocks (see Figure 4.22(a)), which lead to high static power and large die area [128]. These become the key challenges for the future R&D of CAM. Replacing volatile memories by non-volatile memories or applying hybrid non-volatile logic-in-memory circuits is a promising solution to build non-volatile CAM and overcome both these drawbacks. This topic is currently under intense investigation. For instance, a DW motion MRAM based CAM (DW-CAM) was prototyped recently (see Figure 4.22(b)), which demonstrated important progress in terms of power and density [175]. However, this DW-CAM used a three-terminal MTJ as storage element and every memory cell had one comparison circuit and one selected transistor, which lead to a high bit-cell cost and cannot allow the expected ultra-high density.



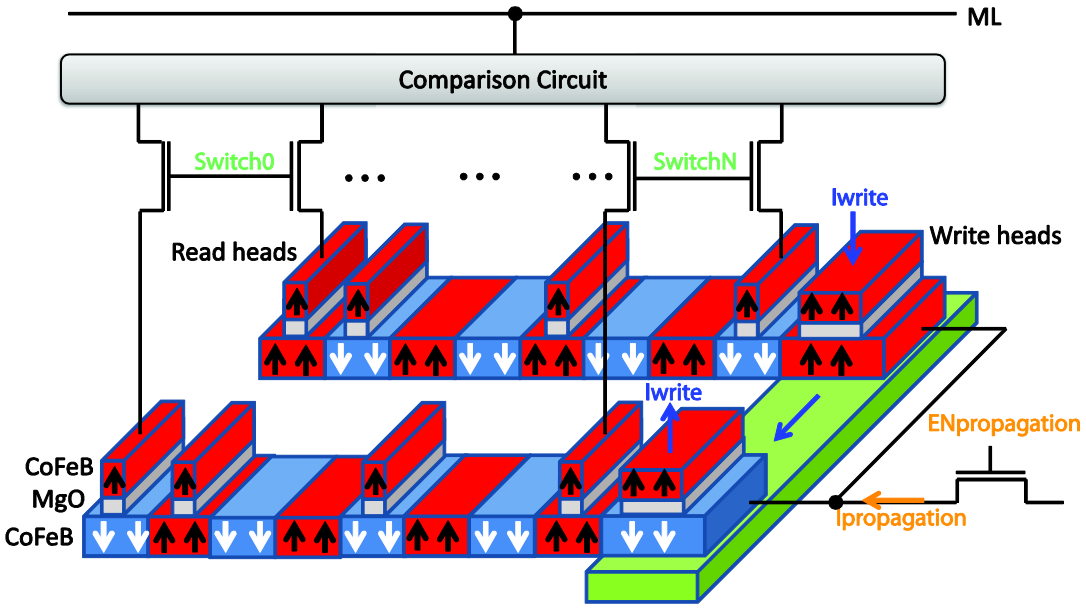
(a) (b)

Figure 4.22 Conventional CAM. (a) SRAM based CAM. (b) DW-CAM. [174-175]

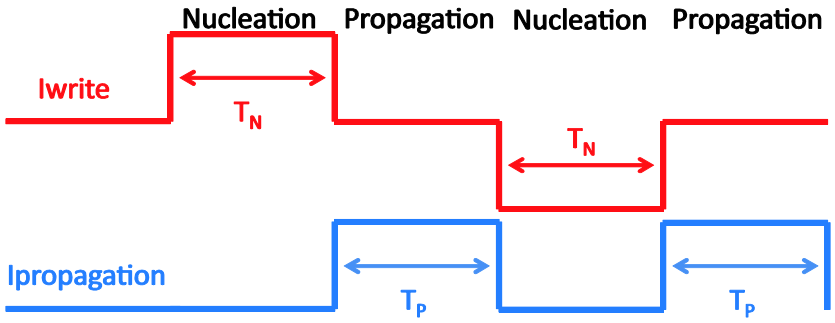
In this situation, we propose a design of CAM based on complementary dual PMA racetrack memories (RM-CAM). Its non-volatile feature can reduce the static power due to leakage currents. The CMOS based DW nucleation and sensing circuits are globally shared to scale down the cell area. The complementary dual nanowires structure allows the local sensing and fast data search operation.

4.5.2 CAM based on PMA racetrack memory (RM-CAM)

4.5.2.1 Structure of RM-CAM



(a)



(b)

Figure 4.23 (a) Structure of dual nanowires based RM-CAM. One writing current pulse nucleates a couple of MTJs with complementary configurations. A propagation current pulse drives the dual nanowires synchronously. Every dual-wire shares a comparison circuit. (b) One example of current pulse configuration for I_{write} and $I_{propagation}$. T_N and T_P are respectively their pulse durations.

The RM-CAM is composed of comparison circuits, PMA racetrack memories and DW nucleation/propagation circuits. A couple of complementary magnetic nanowires are used to present one word (see Figure 4.23(a)) in order to obtain the most reliable and fast access operation for CAM applications as this solution benefits the maximum TMR value instead of

TMR/2 for conventional single nanowire structures. We design the comparison circuit based on PCSA, which allows minimum power and sensing errors. This RM-CAM includes a couple of PMA MTJs connected together as the write heads. Due to the different directions of the writing current pulse I_{write} through these two MTJs, they can nucleate the complementary configurations through STT switching mechanism under the same I_{write} pulse. One of the critical challenges for complementary magnetic nanowires is to synchronize precisely the domain wall positions. Here, the same current pulse $I_{propagation}$ propagates domains in the dual nanowires and we implement the DW pinning constrictions with the same distance in the magnetic nanowires [176]. To avoid the interference between the DW nucleation and the previous data, write heads do not hold the data storage and there is always a $I_{propagation}$ pulse following each DW nucleation (see Figure 4.23(b)). There are also a couple of PMA MTJs at each bit of storage elements as read heads. Since lower resistance can reduce the rate of breakdown and higher resistance can improve the sensing performance, the size of the read heads should therefore be smaller than that of the write heads to obtain the best switching and sensing reliability.

The comparison circuit (see Figure 4.24) consists of two parts: a PCSA detects the complementary magnetizations of the read heads by two reading current pluses (I_{read} and I_{readb}) and outputs a logic value; the transistors MN3-MN6 build a classical NOR-type CAM. The signal "MLpre" is used to pre-charge the match line (ML). In case that the search line "SL" ("SLb" is its complementary signal) matches the stored data, there is no path to discharge and ML will thus be asserted. In contrast, ML will be discharged.

The fast search operation as shown in [175] can be also expected in the RM-CAM. At first, we program the magnetic nanowires, and the switch signals then select each bit of magnetic nanowires to be loaded in the comparison circuit. By sequentially triggering the switch signals, all the words can be explored. If there is no match case, DW nucleation and propagation will be carried out to enter new words for the next search. The programming speed of magnetic nanowires depends on T_N and T_P, which are respectively the pulse durations of I_{write} and $I_{propagation}$. They can be both sped up to ~1 ns. According to the current pulse configuration shown in Figure 4.23(b), the worst case of programming duration is $N \times (T_N+T_P)$, where N is the number of pinning potentials in the magnetic nanowire. We can benefit a higher speed for the

repeated bits such as “111” and “000” when only one DW nucleation is required for three bits.

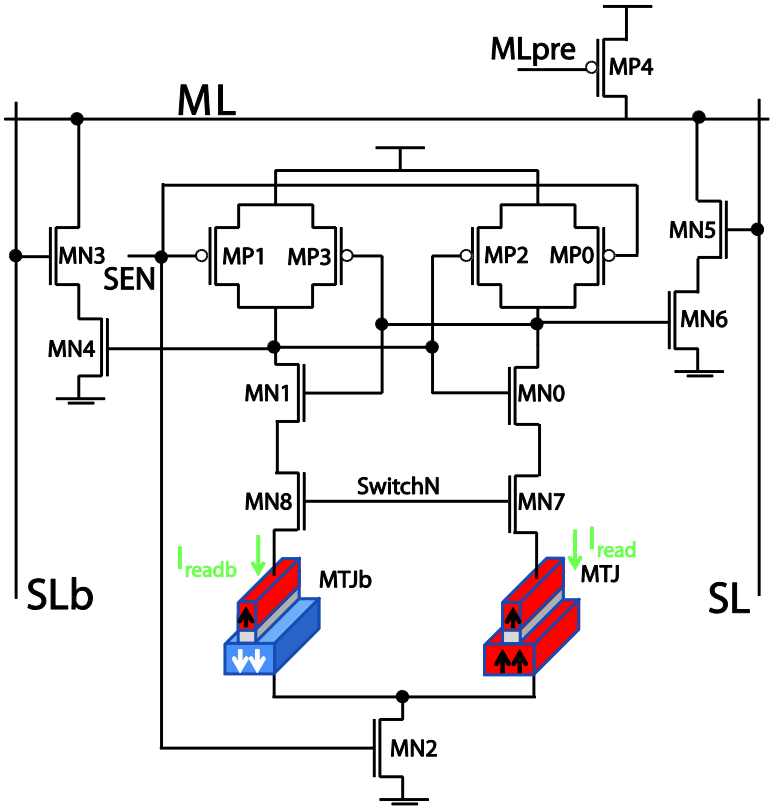


Figure 4.24 Schematic of the comparison circuit. It outputs the logic value ‘1’ or ‘0’ according to the configuration of complementary MTJs. MN3-6 transistors build up a NOR-type CAM.

In order to improve the area efficiency, every couple of dual nanowires shares the comparison circuit in this RM-CAM (see Figure 4.23(a)). Unlike the DW-CAM where there is a large transistor for nucleation for every storage cell, the same write head is shared for one magnetic nanowire in RM-CAM, and the CMOS area dedicated for each storage cell becomes ignored for a long track with numerous pinning constrictions. This structure thus allows an ultra-high density.

4.5.2.2 Simulation and performance analyses of RM-CAM

By using the compact model of PMA racetrack memory and STMicroelectronics CMOS 65 nm design kit, an 8-bits-width-8-words-depth PMA RM-CAM (see Figure 4.25) has been simulated.

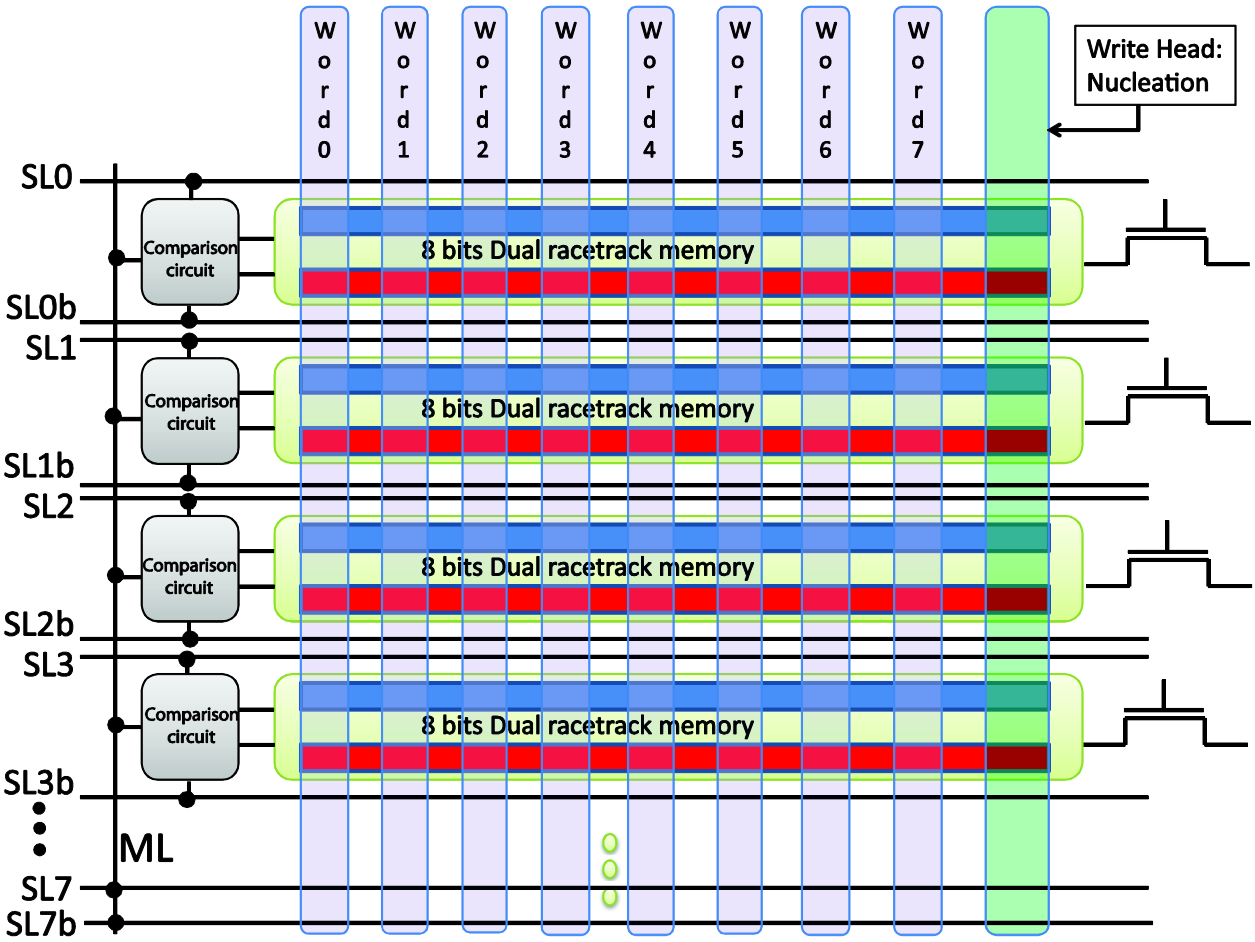
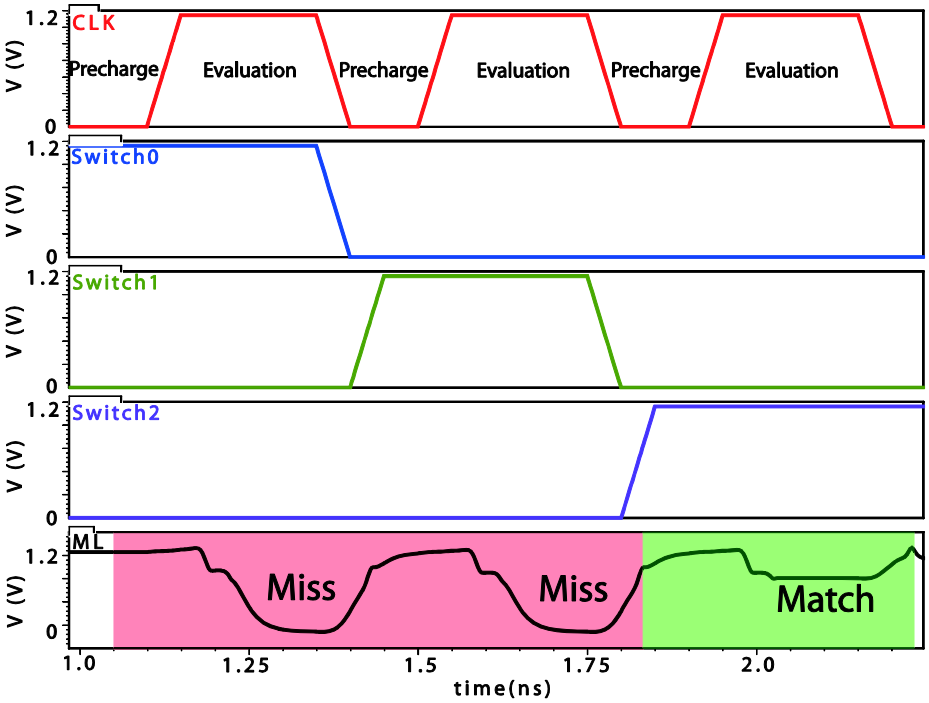
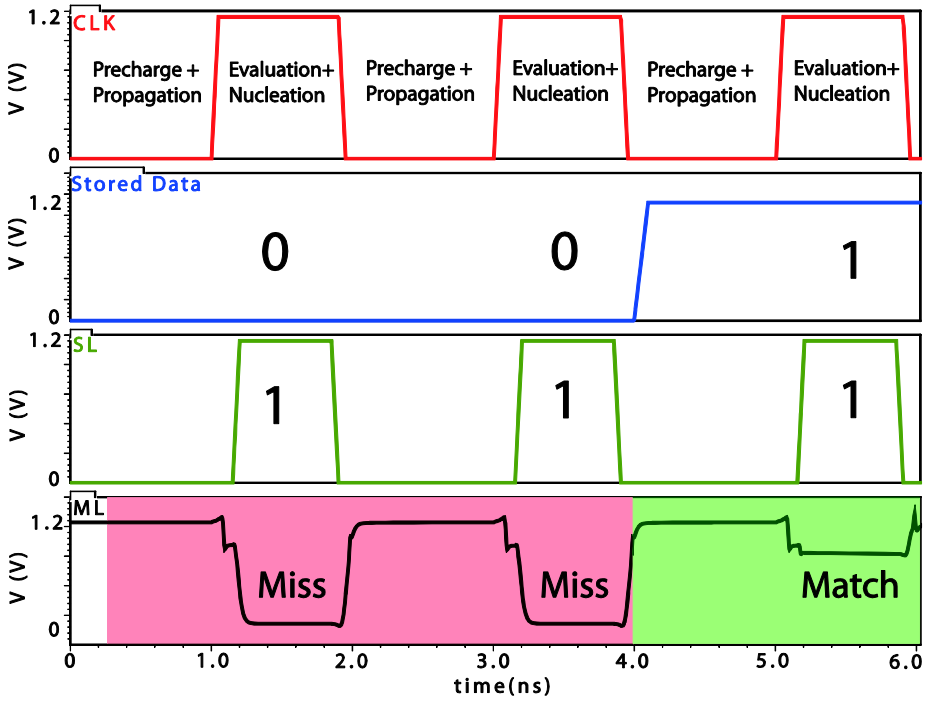


Figure 4.25 Schematic of an 8 x 8 bits RM-CAM. Each word is composed of the bits at the same positions in 8 different dual nanowires; they can be driven to move simultaneously by the propagation currents.

Firstly, we implement the transient simulation for the search operation without DW propagation (see Figure 4.26(a)). The clock signal "CLK" involves the "Pre-charge" phase and the "Evaluation" phase. During the "Pre-charge" phase, both of the signals "SEN" and "MLpre" (see Figure 4.24) are set low to pre-charge the PCSA circuit and the match line "ML". The first word "Word0" has been loaded by enabling the signal "Switch0". With the response of the signal "Miss", "Switch1" will be then activated and so on. This process doesn't stop until the appearance of the match case. We find that this search operation needs only ~0.45 ns, which is faster than that of conventional SRAM-based CAM and DW-CAM. In addition, the energy consumption of searching is as low as ~12 fJ/bit/search, which can be further reduced by the decrease of activity rate thanks to the segmentation of the match line [177].



(a)



(b)

Figure 4.26 Transient simulations of the RM-CAM: (a) Without DW nucleation and propagation. (b) With DW nucleation and propagation.

In case that no storage data can match the search word, a new word will be nucleated and

propagated into the magnetic nanowire for the next round of search. Figure 4.26(b) shows the transient simulation result of the worst case: 1-bit miss process. It means that the rest 7 bits of the search word match the stored data, only one bit is different from the stored data. As shown in Figure 4.26(b), the search bit is '1', if no match is found, the propagation current pulse will start to drive the DW propagation, until "SL" and "Stored data" match each other. We can find the whole operation, consisting of "Pre-charge", "Propagation" and "Evaluation" phases, only requires ~2 ns. This suggests a high operating frequency up to 500 MHz, comparable to that of traditional CAM [178].

We estimate the cell area for RM-CAM with Eq. 4.10:

$$A_C = \frac{A_{CO} + A_{NU} + A_{PR} + N \times MAX(A_{BT} + A_{LS})}{N} \tag{4.10}$$

where A_{CO} denotes the area of a comparison circuit, which is ~50 F^2, A_{NU} denotes the area of a DW nucleation circuit, which is ~48 F^2, A_{PR} denotes the area of a propagation current generating circuit, which is ~7 F^2, A_{BT} is the area of every bit in racetrack memory, A_{LS} is the area of two load selecting transistors for every bit and N is the number of bits per word.

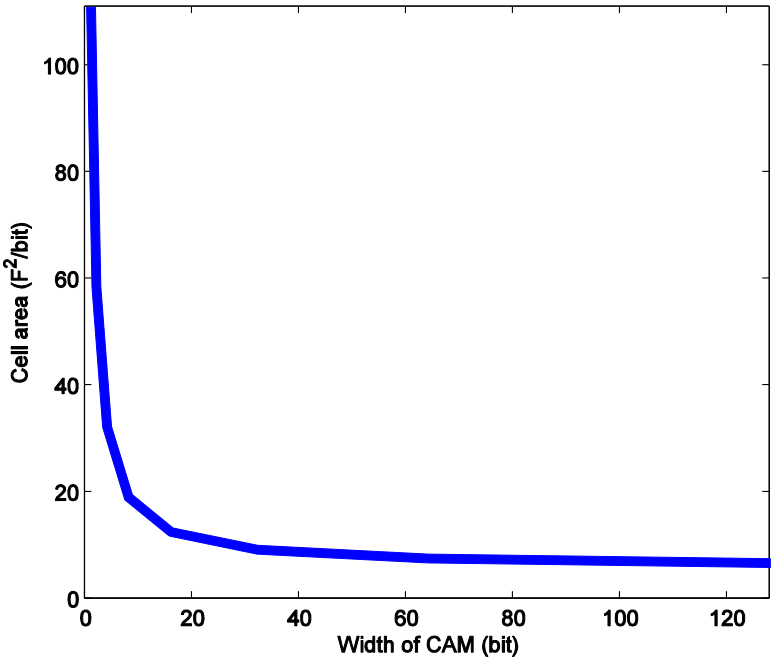


Figure 4.27 Dependence of full area versus number of bits per word.

Due to the 3D integration of MTJs above CMOS circuit, only the larger one between the MTJs' area and the selecting transistors' area will be involved for calculating the full area. For our design, A_{BT} is ~6 F^2 considering 2 F between two adjacent constrictions. Coincidentally, A_{LS} is also ~6 F^2 with the minimum size. If the distance between two adjacent constrictions exceeds 2 F, only A_{BT} would be taken into account in Eq. 4.10. As N = 8 for our simulation, the cell area per bit is therefore ~19 F^2, which is much lower than that of SRAM-based CAM or DW-CAM [174-175]. Meanwhile, with the increase of the bit number per word, the area of shared CMOS circuits for data comparison, DW nucleation and motion would become negligible (see Figure 4.27). The cell area per bit will approach to $MAX(A_{BT} + A_{LS})$ (e.g. ~6 F^2 for our design).

4.5.2.3 Comparison of CAMs based on different technologies

With the performance analyses above, we summarize the comparison of CAMs based on different technologies. From Table 4.5, we can find that non-volatility of racetrack memory allows RM-CAM to eliminate the static power. DW propagation in the racetrack memory benefits for improving the search speed. Most importantly, RM-CAM shows a great advantage in terms of density thanks to 3D integration and sharing of CMOS circuits (comparison circuit, DW propagation circuit and DW nucleation circuit). However, we have to mention that the cell area of DW-CAM shown in Table 4.5 does not consider the 3D integration. If it applies 3D integration, it could reach N times of our proposed RM-CAM. For example, if N=8 for RM-CAM, the cell area of DW-CAM would be about 160 F^2/bit.

Table 4.5 Comparison of CAMs based on different technologies

Type	SRAM based CAM	DW-CAM	RM-CAM
Cell area (F^2/bit)	540	~815	~19
Cycle time (ns)	2	5	~0.45
Energy (fJ/bit/search)	9.5	~30	~12
Static power	Yes	No	No

4.6 Conclusion

In this chapter, by using the compact models of PMA STT MTJ and PMA racetrack memory developped in the Chapter 3, we designed and investigated a number of hybrid spintronics/CMOS logic and memory circuits based on 65 nm or 40 nm CMOS technology node, from the basic writing and sensing circuits for MTJ to more complex MFA and CAM based on PMA STT MTJ or PMA racetrack memory. Thanks to non-volatility, high density, high speed and nearly infinite endurance of PMA STT MTJ and PMA racetrack memory, the circuits based on them all demonstrate the performance advantages in terms of area and/or speed and/or energy.

In details, non-volatility allows the hybrid systems to be powered off while saving the data, and then to eliminate the static power consumption. This feature can reduce greatly the overall power consumption, especially for normally-off systems. 3D integration technology can improve the system's density efficiency. Moreover, it can shorten the distance between logic and memory, which helps to save considerably the transfer energy and time. Although the switching speed of MTJ doesn't show an evident advantage compared to conventional CMOS, it is still sufficient for logic and memory application. To overcome this challenge, using of the DW motion is an alternative solution. Considering that the distance between two adjacent DW is 40 nm and the propagation speed can be as high as 100 m/s, switching a state by propagating DW can be as fast as 400 ps. That is why we believe the DW motion based racetrack memory design has a great potential for the future high-speed low-power systems.

Nevertheless, there are still material and fabrication challenges as well. For instance, the relatively high critical current for DW propagation is always an obstacle for the progress of racetrack memory. In the following chapter, we will carry out the investigation of this issue and propose the design optimizations to resolve it.

Chapter 5 Design Optimization for STT-MRAM and PMA Racetrack memory

5.1 Introduction

As mentioned above, STT switching mechanism allows a simple switching process for MRAM and provides a prospective for the scalability of MRAM. However, there are still a lot of challenges to achieve high density. In order to address these density limitations, several schemes have been proposed, such as cross-point architecture, 3D configuration. In this chapter, we present another concept, multi-level cell (MLC), to improve the density of STT-MRAM [179]. This concept is a promising way to multiply the memory density with little area overhead so that it has been applied in other non-volatile memories, for example, flash memory. Our proposal is so specific that it benefits from the STT stochastic behavior which is normally considered as a disadvantage for STT-MRAM. It not only allows higher storage density, but also reduces programming power and delay. This new cell can be also used as electrical synapses to build neuromorphic computing systems or other biological networks [180-181].

On the other hand, we also endeavor to optimize the racetrack memory design. Among the complex features of racetrack memory, DW motion is elemental. In particular, the velocity of DW motion is the major issue that plays an important role for the implementation of racetrack memory. There is a critical current density that a current pulse must exceed to trigger the propagation of DWs. However, a difficulty for wide application of racetrack memory is that the critical current is considerably high, which can derive a series of problems: 1) current can hardly reach or exceed the critical current for a long-length magnetic nanowire, which will lead to a limitation of storage capacity; 2) the increase of joule heat can lead to a risk of structural destruction. With the scaling-down of nanowire in racetrack memory, its resistance increases continuously, the drawback of high current threshold become more and more evident.

The peripheral circuit has the difficulty to provide a current both high and stable to propagate DWs, especially for a long nanowire to achieve high capability. Optimization of material resistivity could be a solution to overcome this issue; however the resistivity of ferromagnetic materials is difficult to be reduced below 10^{-6} Ωm at present, the further reduction depends mainly on the material technology advancements. Finding an alternative to decrease the current for propagating DWs is indispensable for future development of DW-based device, not merely racetrack memory.

As a consequence of Walker breakdown effect, DW motion induced and propagated below critical current in magnetic field environment has recently been discovered [182]. This phenomenon encourages the magnetic field assistance to be employed in racetrack memory in order to improve the storage capacity and feasibility.

In this chapter, we present a MLC design for STT-MRAM and a nanowire system of PMA Co/Ni racetrack memory with magnetic field assistance. Analyses are successively carried out to show their advantageous performances.

5.2 Multi-level cell (MLC)

Multi-level cell (MLC) is a concept with respect to Single-level cell (SLC), which is able to store multiple bits of information in one cell. It has been maturely applied in some non-volatile memories, such as flash memory [183] and resistance random access memory (RRAM) [131, 184-185]. It is also a promising technology to improve the density of MRAM and implement the neuromorphic computing. In this section, we take advantages of stochastic effect in PMA STT MTJ to realize MLC, even though the stochastic effect is normally considered as a disadvantageous factor. The simulations and analyses display that this MLC based on stochastic effect can provide not only high density but also low power and high speed.

5.2.1 Structure of MLC

Two traditional structures of MLC are the parallel and the serial structures (see Figures 5.1-5.2) [186]. MLC usually uses a selected transistor (Tp for parallel and Ts for serial) to enable the switching process, which is similar to the 1T/1MTJ structure used in MRAM. In addition, the number of levels is determined by the number of MTJs. *N* MTJs in the MLC can yield *N*+1 levels meaning $\log_2(N+1)$ bits of data correspondingly.

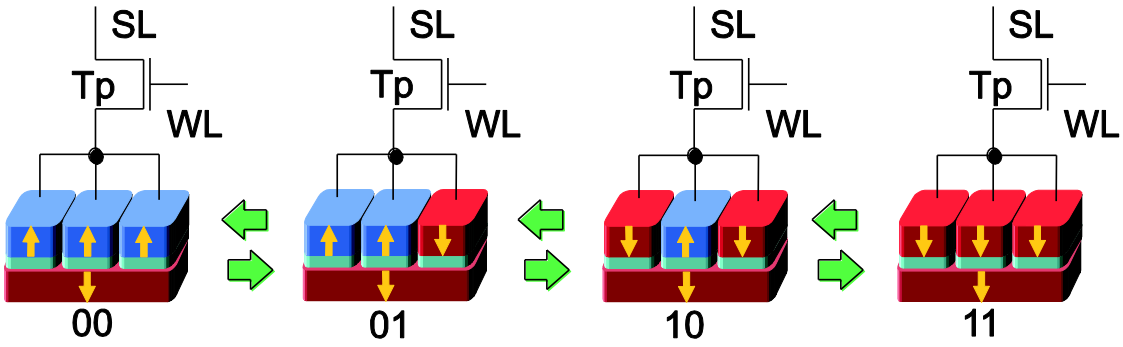


Figure 5.1 2-bit parallel MLC

Figure 5.1 shows an example of 2-bit parallel MLC structure where 3 separated free layers are grown on an entire pinned layer, although it is not yet achievable with the existing technology capability. In the following simulations, we used the clusters of parallel-connecting MTJs as the parallel MLCs. The serial MLC is a stack of MTJs connected vertically as shown in Figure 5.2. Ordinarily, the serial MLCs requires a lower current but a higher voltage than the parallel ones,

thereby it is preferable to use current source for programming the serial MLC and voltage source for parallel MLC. Furthermore, in order to generate a higher current for the parallel ones, the selected transistor used in the circuit will be larger than those for serial MLC. Thus, the serial structure seems more achievable between these two types of structures.

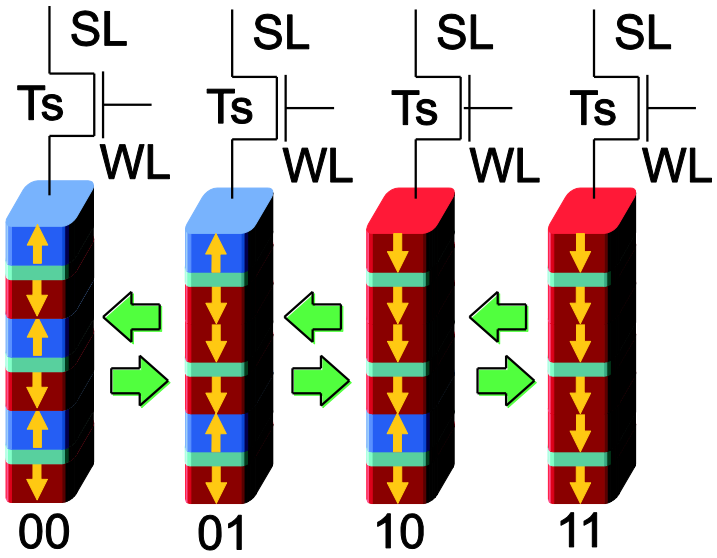


Figure 5.2 2-bit serial MLC

Moreover, from the point of view of real TMR ratio, the serial structure is more favorable than the parallel structure. According to the definition of TMR ratio, by taking the resistance of selected transistor into account, the real TMR ratios can be described as Eqs. 5.1-5.2.

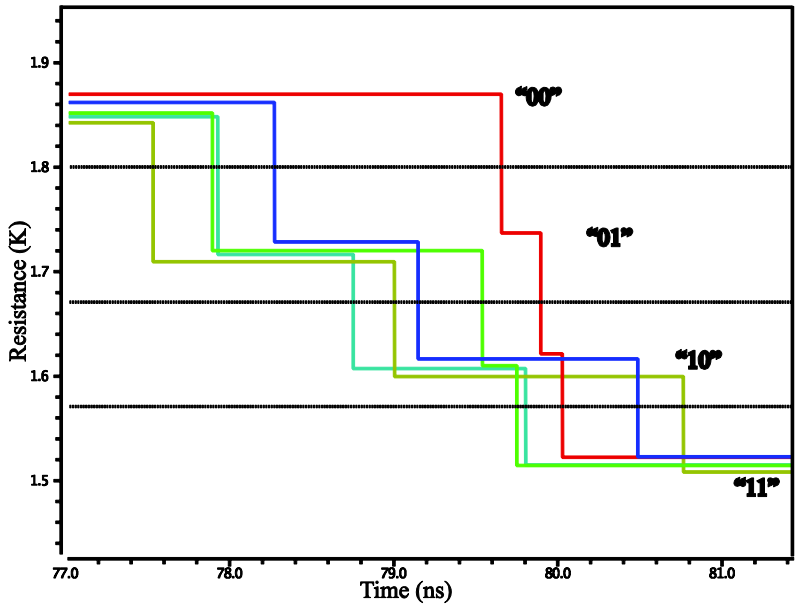
$$TMR_{rs} = \frac{R_{AP} - R_P}{R_P + \frac{1}{N} R_{ST}} \tag{5.1}$$

$$TMR_{rp} = \frac{R_{AP} - R_P}{R_P + N R_{ST}} \tag{5.2}$$

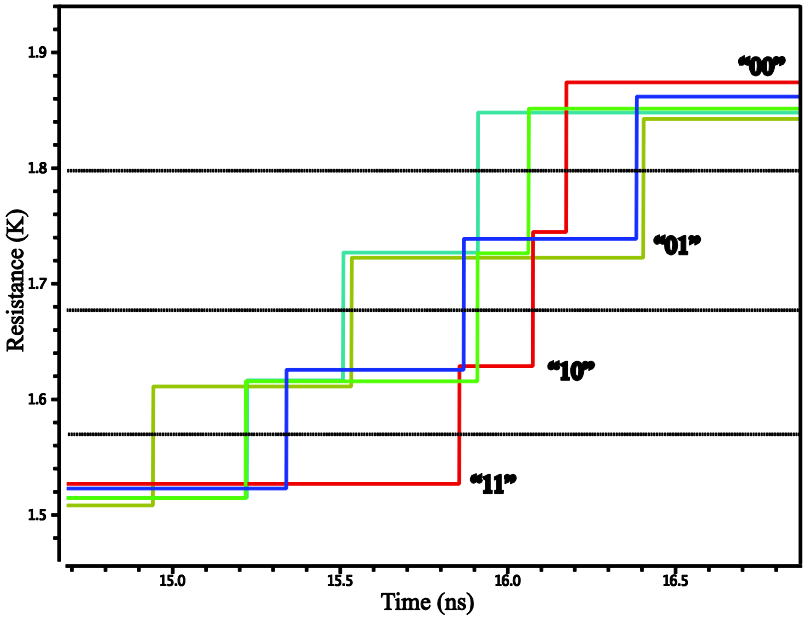
where TMR_{rs} and TMR_{rp} are the real TMR ratios for serial and parallel structures, R_{ST} is the resistance of the selected transistor, N is the number of MTJs in the MLC. Note that R_{ST} is normally larger than the resistance of MTJ, the real TMR ratio of the serial structure is thus larger than that of the parallel one.

For these structures, due to the stochastic behavior, the 3 MTJs are switched randomly and successively. If we consider antiparallel state of MTJ as the logic '0' and parallel state as '1', the MLCs can be programmed from '00' to '11' or from '11' back to '00'.

5.2.2 Simulation and performance analyses of MLC

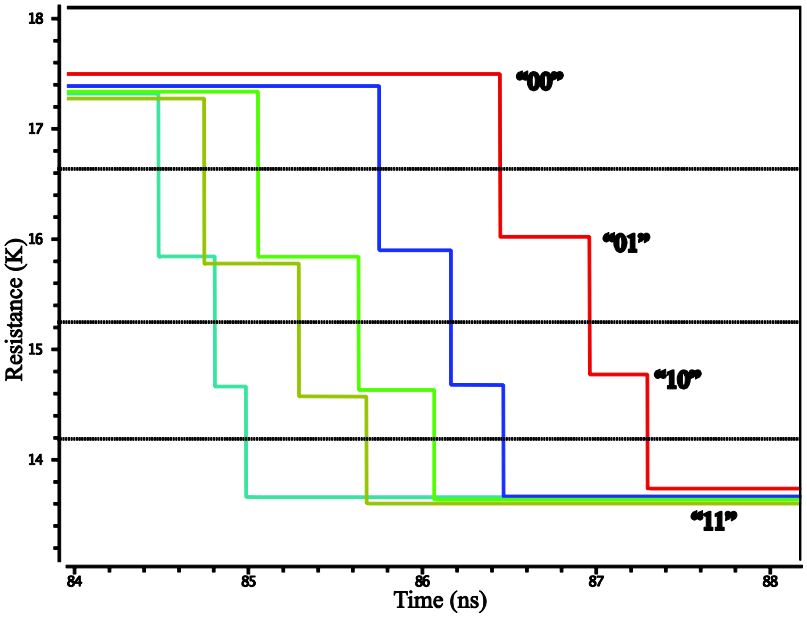


(a)

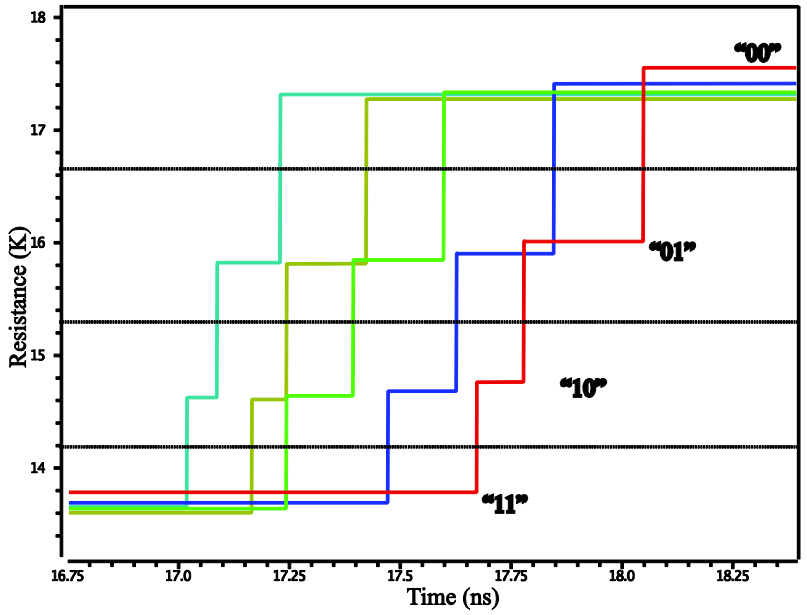


(b)

Figure 5.3 Monte-Carlo simulations for 2-bit parallel MLC. (a) Data switches from '00' to '11'. (b) Data switches from '11' to '00'.



(a)



(b)

Figure 5.4 Monte-Carlo simulations for 2-bit serial MLC. (a) Data switches from '00' to '11'. (b) Data switches from '11' to '00'.

The functionalities of these two structures are respectively validated by the Monte-Carlo simulations (see Figures 5.3-5.4). The MTJ's key parameter variations, such as TMR ratio, oxide barrier and free layer thickness, are also considered. Figure 5.3 shows the case for parallel MLC programming from '00' to '11' and back to '00'. Figure 5.4 illustrates the other case for serial

structure. Each process has actually been executed 100 times under a 10 ms pulse. However, we only show 5 times among them for the purpose of clarity. We can find that the states of MLCs vary randomly and step by step, which is the functionality that we anticipated. Note that these two simulations are executed in the aforementioned thermally assisted regime (Néel-Brown model). When expecting to accelerate the programming procedure, higher current can be applied to drive the MLC to work in the precessional switching regime (Sun model). In this condition, the switching duration can reach nanosecond or sub-nanosecond, however the multi-level performance degrades accordingly. By comparing these two simulation results, the difference between two adjacent resistance levels of serial MLC is larger than that of parallel one, which proves the advantage of serial MLC for sensing operation.

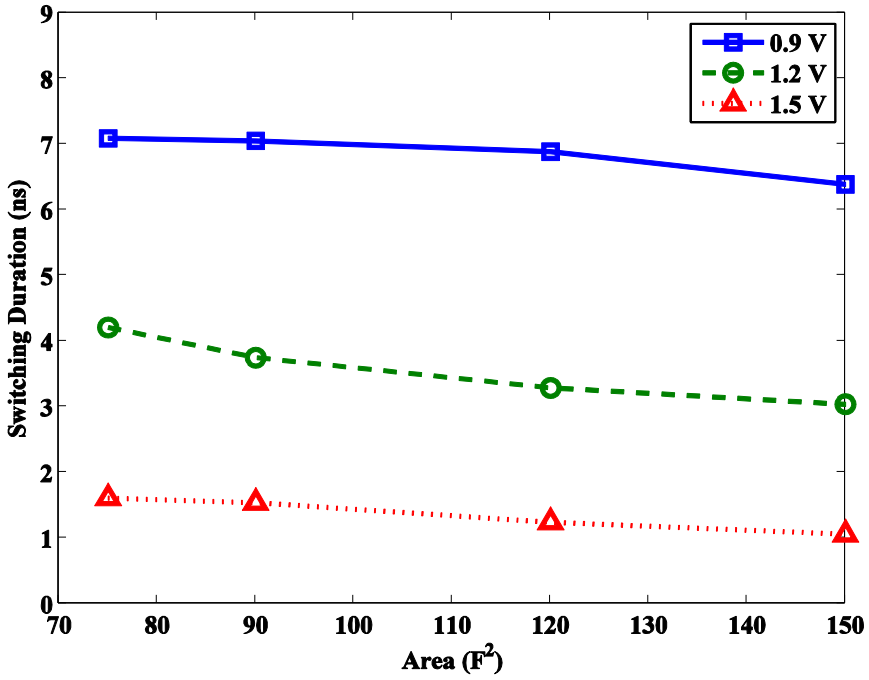


Figure 5.5 Tradeoff dependence of switching duration versus area and pulse magnitude for serial MLC.

Thanks to the 3D integration technology, the area efficiency of hybrid MTJ/CMOS circuits can be improved without nearly no area overhead. However this also makes the CMOS part dominate the overall area. On the other hand, the currents through the MLCs govern the programming speed, which is also determined by the area overhead of transistors. Figure 5.5 demonstrates a slight tradeoff relationship between the average switching duration and the area of CMOS part in the case of serial MLC: the increase of area can improve the speed very slightly. This is because

that the overall resistance of serial MLC is comparable or larger than that of selected transistor, the resistance variation of transistor cannot influence the current obviously by fixing the voltage. On the other hand, we also take the effect of pulse magnitude into account by applying 3 different magnitudes of pulses (i.e. 0.9 V, 1.2 V and 1.5 V). It shows that the higher voltage improves the speed. Thereby, in order to obtain high capacity and high performance, we suggest using serial MLC with a voltage as high as possible in practice.

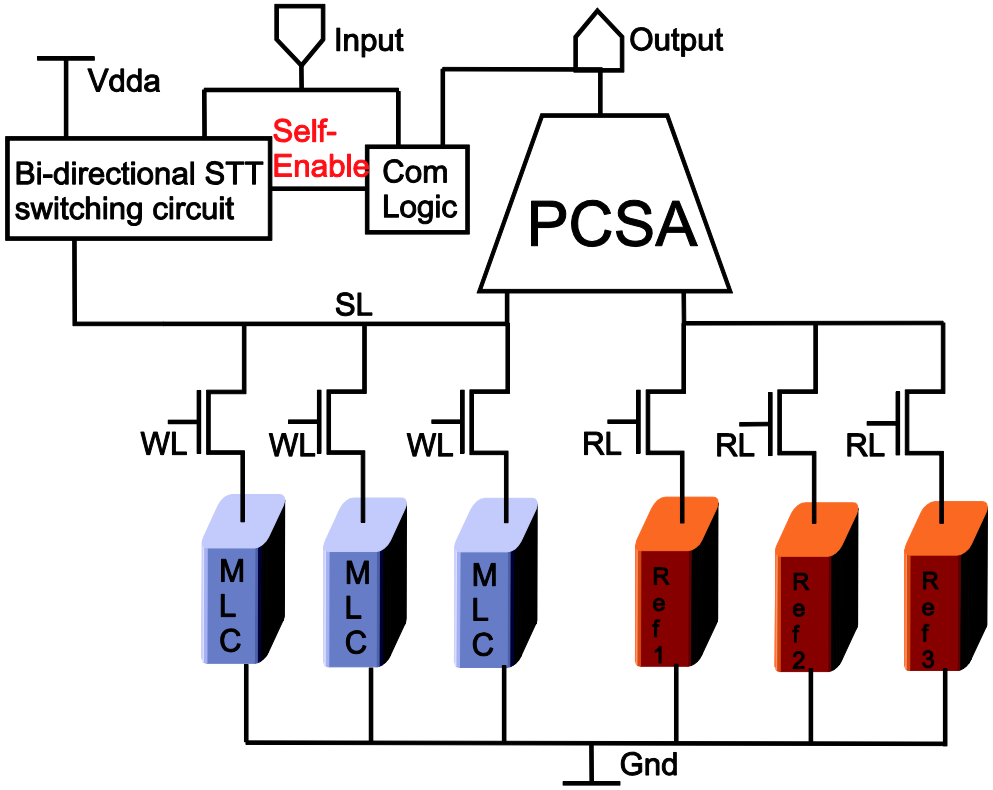


Figure 5.6 MLCs served by "Self-Enable" switching circuit.

Moreover, due to the stochastic feature, the switch for adjacent states could be ultra-fast. If it can be controlled strictly and properly, for instance, by integrating with "Self-Enable" switching mechanism (see Figure 5.6), this device can be advantageous in terms of fast speed and low power consumption. In the "Self–Enable" switching circuit, besides the synchronized PCSA and the bi-directional STT switching circuit, the comparison logic (Com Logic) circuit is able to control the input signal by monitoring the state change of MLCs. As shown in Figure 5.6, the PCSA detects the state of MLC and outputs a logic value to the "Com Logic" circuit. By comparing this logic value with the input data, it can automatically block the input pulse once the state of MLC changes. This makes the switching operation deterministic instead of stochastic,

which saves a great deal of energy consumed by the conventional iterations for probabilistic switching [179].

5.3 Racetrack memory with magnetic field assistance

5.3.1 Peripheral circuit and material resistivity optimizations and their drawbacks

Similar to STT-MRAM, peripheral CMOS circuits of racetrack memory are also the big challenges as they define essentially the final area, speed and reliability performance. In the last years, the circuits for STT-MRAM writing and sensing have been well studied, which have also been applied compatibly for racetrack memory, for example, the sensing and writing circuits shown in Figure 5.1 and Figure 5.6. Beyond them, the circuit for propagating DWs in the magnetic nanowire is a critical part of racetrack memory. As data transferring operation and its speed depend deeply on the DW propagation current [121, 123], an efficient current generator source is thus required for practical application of racetrack memory in logic and memory. It should provide enough high current to ensure the DW motion and keep it with a steady speed. However, due to the high material resistivity and the limitation of CMOS circuits, this requirement of current becomes the main bottleneck delying the progress of high-capacity racetrack memory. As shown in the first prototype of racetrack memory, an extra magnetic field is required for the DW motion [124]. This is probably caused by the insufficient current generation of CMOS peripheral circuits to drive the DW motion. We compare several designs for the current generation and address the material resistivity issues. The discussions can help to propose the optimized solutions for racetrack memory.

5.3.1.1 Voltage source for current generation

A NMOS or PMOS transistor constitutes the simplest current generator (see Figure 2.14). By applying an external voltage supply, it delivers sufficient current to move DWs in the magnetic nanowire. Ideally, a control gate signal switches the transistor between its saturated and blocked operation region. Nevertheless, storage capacity is limited by the resistance of nanowire. For a given voltage supply, if we increase the nanowire length, its resistance will become higher. This will slow down the DW motion speed until the current is lower than the propagation threshold value, leading to the motion failure. We denote L_{Max}, maximum length of DW propagated by the shifting current.

The curves on Figure 5.7 demonstrate the current generation of voltage source composed of NMOS transistor with fixed geometric factors (width and length) corresponding to different length of nanowire. We can find that the increase of nanowire length will decrease greatly the DW shifting current value I_{shift} and then influence the DW motion speed. In this case, we can get L_{Max}, which equals to 10 µm for the thin nanowire with 2.5 nm thickness. This represents 250 bits data storage in 40 nm technology nodes.

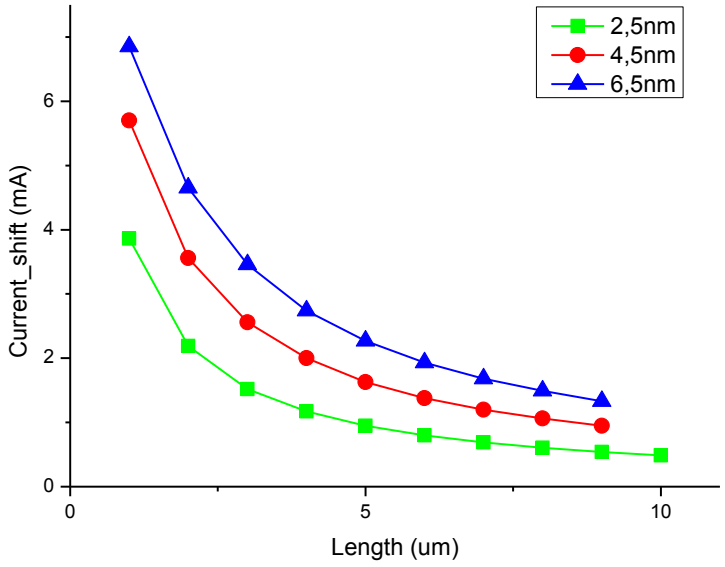


Figure 5.7 Static simulations to evaluate voltage sources performances using a varying magnetic nanowire length powered by a fixed voltage 3V. Three curves describe the thickness impact on the voltage source performance.

For different thickness, a wide range is observed between the maximum level of shifting current and the minimum one. The variation rate of current intensity is estimated to ~87% for thickness of 6.5 nm and ~80.5% for thickness of 2.5 nm. This large variation highlights the instability degree of voltage source for the current generation, which will limit the data access speed control and DW motion tolerance to the process variation as shown in [124]. The reason of this weakness is essentially due to MOS functional characteristics. As magnetic nanowire resistance increases, I_{shift} corresponding to the drain to source current of transistor I_{DS} decreases at a fixed voltage

gate level V_{GS} . This will switch rapidly the NMOS transistor functionality region from saturation configuration to linear one.

5.3.1.2 Current source for current generation

In order to optimize the stability of current generation against the shape variation while keeping high current amplitude, we design several current sources based on current mirror [187] as alternative solutions for driving the DW motion (see Figure 5.8). By choosing different integration method of shifting current pulse, three schematics have been designed. For instance, in the schematic Figure 5.8(a) a current mirror is added directly to the NMOS transistor driven by shifting control pulse V_{pulse}. In the schematic Figure 5.8(c), the NMOS transistor is placed in the reference current branch to allow the PMOS transistor M3 to connect the nanowire and provide high current amplitude.

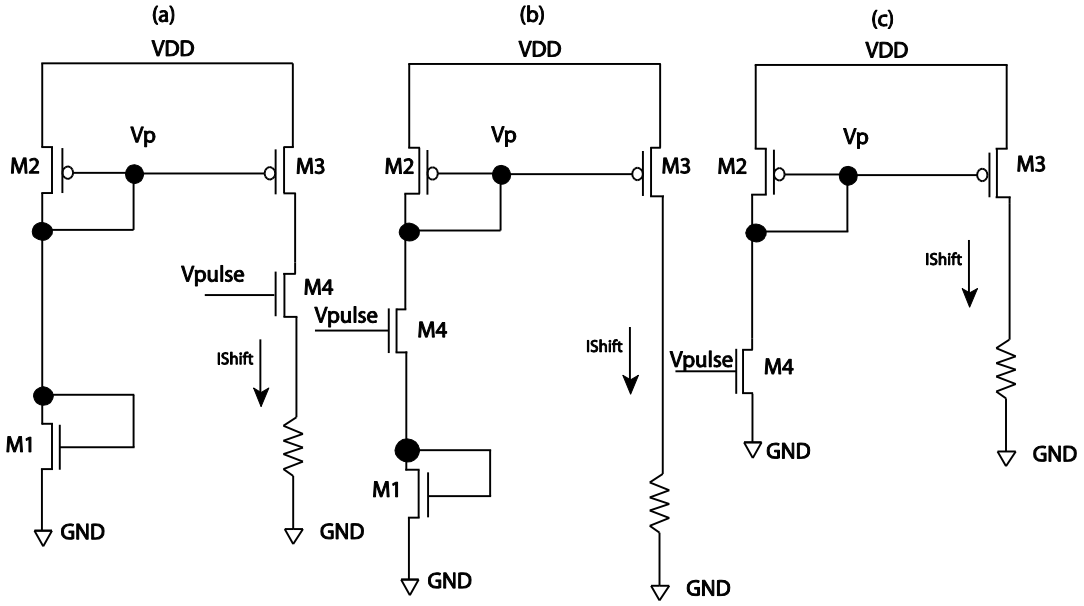


Figure 5.8 Current source based on current mirror to optimize current amplitude and stability: (a) pulse generator transistor "M4" is designed on the same branch of magnetic nanowire. (b) Pulse generator transistor "M4" is set on the reference circuit branch. (c) The total charge on the reference branch is lessened by deleting the active charge.

Among them, the schematic shown in Figure 5.8(b) presents a best shift current intensity variation (see Figure 5.9), which is ~23.7% compared with the ~80% of voltage source. This

relatively high stability is due to the using of a lower gate voltage at "M3" PMOS transistor rather than "Vdd" used in voltage source. And this lower voltage is locked by the transistors "M1" and "M4". However, we can clearly find that the current amplitude is reduced; as a result, L_{Max} is reduced from 9 µm to 7 µm.

In summary, it is difficult to achieve both high and stable current generation caused by the linear resistance rising of long magnetic nanowire. Current source could provide a steady current despite the variation and length of nanowire; voltage source allows the best L_{Max} to be obtained.

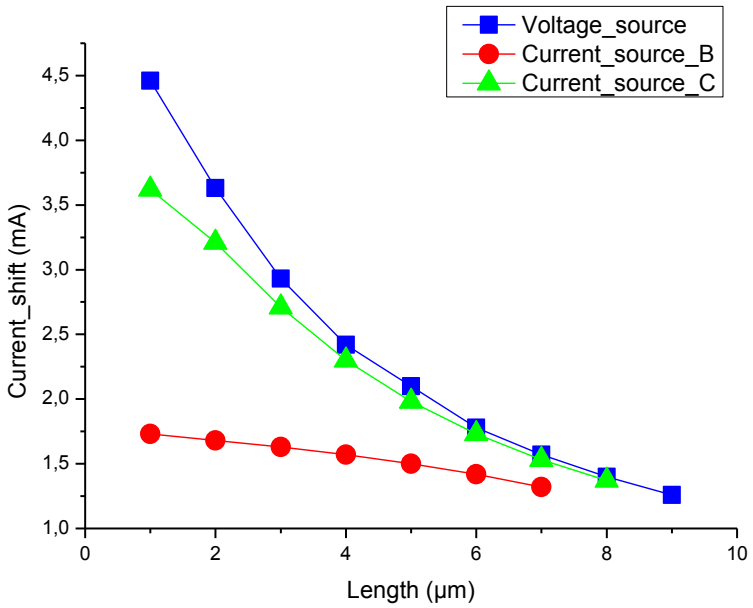


Figure 5.9 Dependence of nanowire length on shifting current for different current and voltage sources by fixing the thickness at 6.5 nm.

5.3.1.3 Material resistivity optimization

In the sections above, we show the challenge of designing an efficient current generation circuit against the linear resistance increase for long magnetic nanowire. This is mainly due to the relatively high material resistivity of magnetic alloy, which is often higher than 10^{-7} Ωm. In the Ta/CoFeB/MgO structure, the resistivity is as high as 10^{-6} Ωm. Figure 5.10 shows that the maximum nanowire length for material CoFe with resistivity of 10^{-7} Ωm can be achieved up to 24 µm; however the maximum length for Ta/CoFeB/MgO structure is only 4 µm. For NiFe

nanowire used in the prototype of IBM [124], L_{Max}≈10 µm, which confirms that the nanowire length varies between 6 and 12 µm. Moreover, the thickness of magnetic nanowire has limited impact on L_{Max}.

It is important to note that only few structures allow both PMA and high TMR ratio (e.g. 100%). We expect to reduce its resistivity by optimizing the thickness of thin films in the nanopillar configuration. Recently, Toshiba presents a new magnetic material [114], which demonstrates a high TMR ratio up to 200% while keeping low nucleation current. We believe that this new structure can be very promising to build up future PMA racetrack memory performing ultra-high density and fast speed.

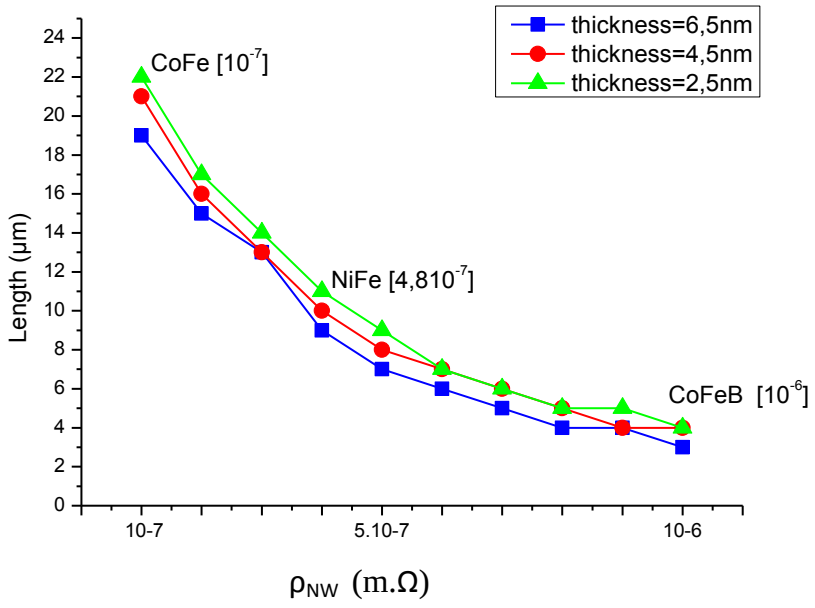


Figure 5.10 Maximum nanowire lengths for magnetic DW motion L_{Max} versus different magnetic material resistivity.

5.3.1.4 Discussion

As L_{Max} is limited to ~2 µm due to the high resistivity (e.g. 10^{-6} Ωm), we cannot use one nanowire to store a large number of data. The conventional integration architecture composed of one bit address per nanowire is thus not suitable for the material like Ta/CoFeB/MgO which has a

high resistivity. However, we can use a short magnetic nanowire to represent one word. For instance, a word composed of 64 bits needs 2.6 µm, which is able to be easily driven by both voltage source and current sources according to Figure 5.9. Figure 5.11 shows an example of the architecture where a racetrack memory representing a word.

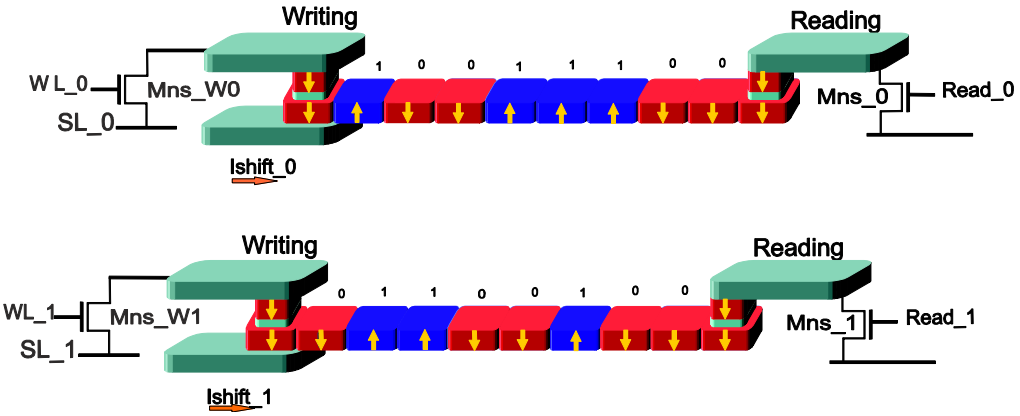


Figure 5.11 Short magnetic nanowire based racetrack memory is more suitable for Ta/CoFeB/MgO structure and there is a word per wire.

5.3.2 DW motion below critical current triggered by Walker breakdown

Recent progress shows that DWs could be shifted below the critical current in an external magnetic field of the order of DW pining field in PMA Co/Ni material [182]. This counterintuitive phenomenon is ascribed to Walker breakdown. Figure 5.12 shows the experimental results measured for a multilayer 3Ta/1.5Pt/0.2Co/(0.6Ni/0.2Co)4/1.5Pt/3Ta (thickness are in nanometer) reproduced from the Ono group [182]. As shown in Figure 5.12(a), it can even occur in the case that the direction of DW driven by current is opposite to that of magnetic field. Figure 5.12(c) shows the relation between the applied magnetic field and the critical current density required to propagate DW in this case. The red open circles represent DW depinning solely by magnetic field and the intrinsic current threshold is marked by the grey shadow. We can find that the critical current density can be reduced 40% compared to the intrinsic current threshold. With relation between the current density and the DW velocity, the relation between magnetic field and DW velocity can be obtained as Figure 5.12(d). It confirms experimentally the additivity of magnetic field induced velocity and current induced velocity above 500 Oe (below 500 Oe, it should consider 2D DW dynamics), which is also a good

agreement between the experimental measurements and 1D model theory. This discovery opens a route to achieve the goal of reducing the critical current with a different mechanism and encourages the magnetic field assistance to be employed in racetrack memory design in order to improve the storage capacity and feasibility.

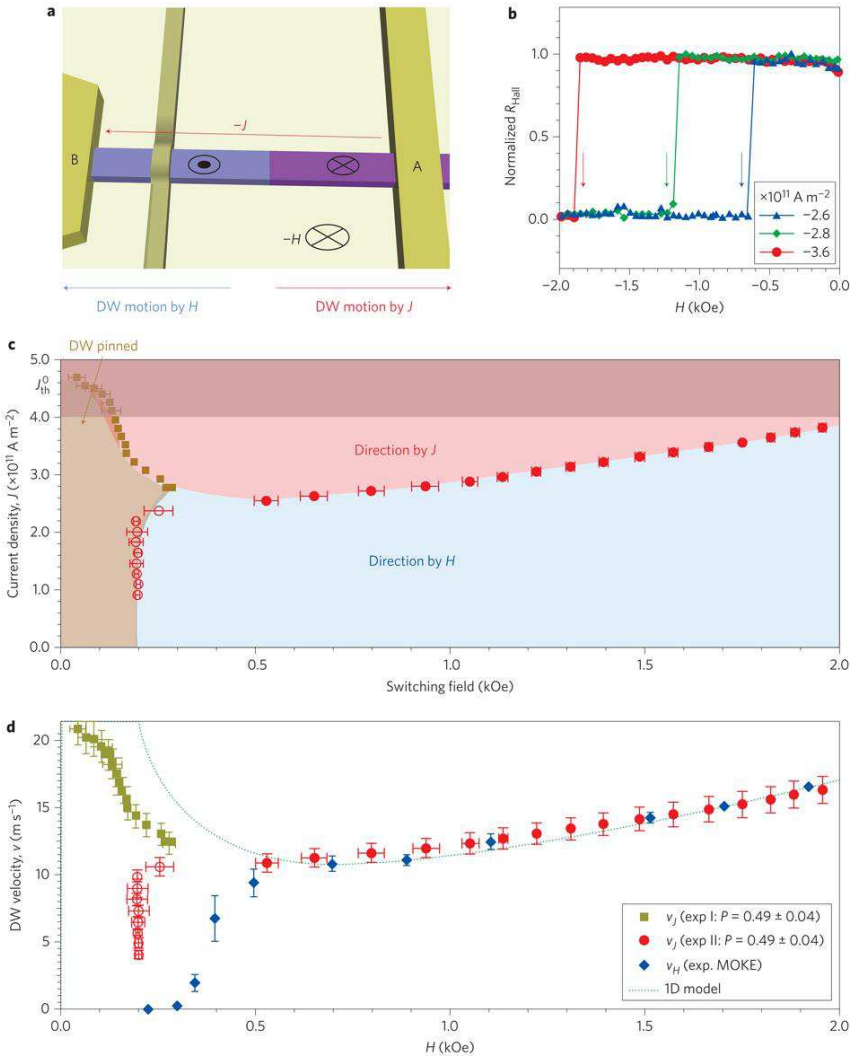


Figure 5.12 Direction of DW motion with the effects of current and field. (a) Experiment schematic. (b) Hall resistance as a function of magnetic field. (c) Relation between the switching field and current density. (d) Relation between magnetic field and domain wall velocity. [182]

5.3.3 Racetrack memory with magnetic field assistance

Taking the above phenomenon and theory into account, we propose a PMA Co/Ni racetrack memory with global magnetic field assistance. It is a promising way to solve the capacity

bottleneck of racetrack memory caused by the high critical current for DW motion. Reduction of DW shifting current benefiting from magnetic field assistance can make racetrack memory more feasible to achieve ultra-denser storage.

5.3.3.1 Structure of racetrack memory with magnetic field assistance

Figure 5.13 shows the structure of our proposed racetrack memory with magnetic field assistance. We apply a current flowing through high-level metal wire to generate magnetic field. A couple of racetrack memories share one metal wire to economize the area cost. As the distance between metal wire and magnetic nanowire is constant and the size of nanowire is relatively small, the local magnetic field can be considered as a global field for every entire racetrack.

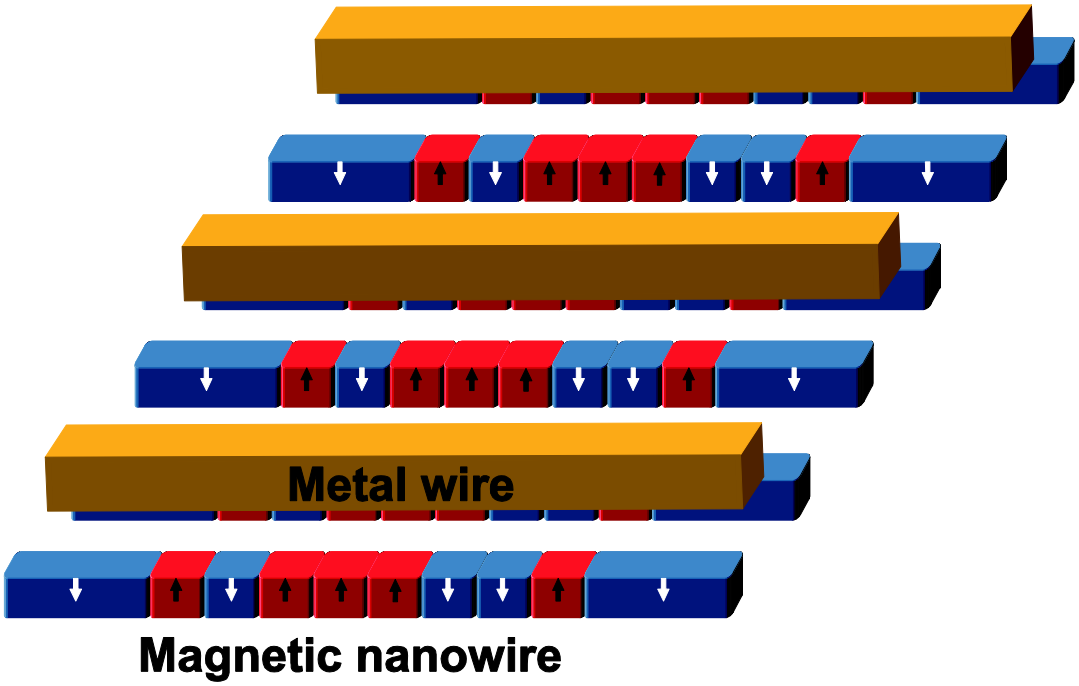


Figure 5.13 Structure of PMA Co/Ni racetrack memory with magnetic field assistance: a high-level metal wire is shared by a couple of racetrack memories.

Generation of magnetic field is a crucial part for this design. There are many ways to generate the magnetic field, for example coils, permanent magnet and current passing through the metal line. However, coils cost more power consumption and area; permanent magnet is difficult for advanced technology integration and nano-level implementation. As the racetrack memory is a linear system, placing metal lines is easier for realization and more beneficial for the

miniaturization. Nevertheless, it should be under some constraints: firstly, as the relatively high current (e.g. 15-20 mA) is required, we should use thick metal in the back-end integration process to pass through these currents so as to prevent the electromigration issues; secondly, the generation of magnetic field coherent to magnetization orientation of domains requires an interval between magnetic nanowire and metal wire.

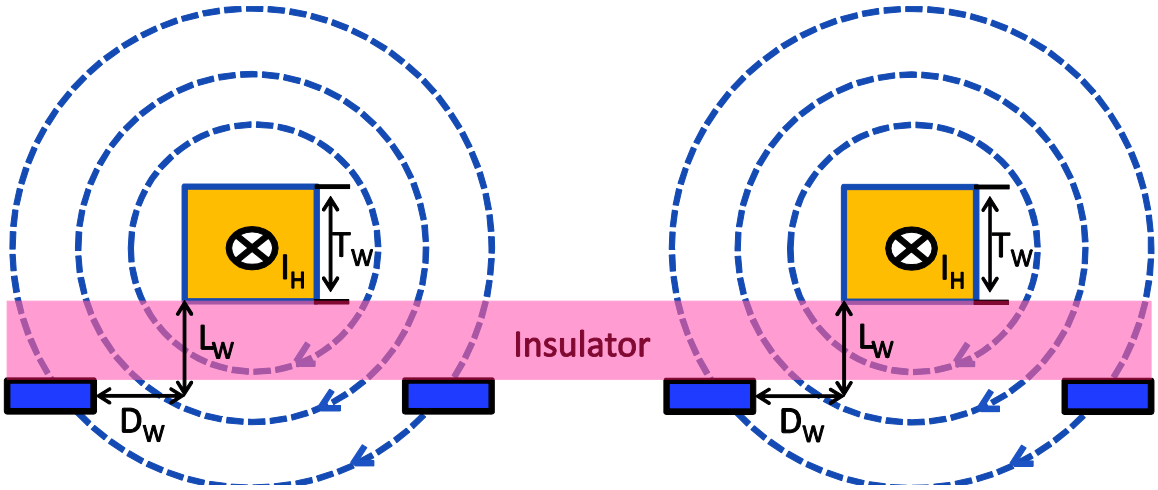


Figure 5.14 Sectional view of the structure: T_w is the thickness of the metal wire, L_w is the distance between the metal wire and racetrack memory, D_w is the interval between the metal wire and racetrack memory.

Figure 5.14 illustrates sectional view of the structure of racetrack memory with magnetic field assistance, we can deposit an insulator layer (e.g. MgO) to separate magnetic and metal wires. T_w is the thickness of thick metal wire (160 nm in our design). L_w and D_w represent the distance and interval length between magnetic nanowire and metal wire (L_w is supposed to be equal to D_w in the following analysis). According to *Biot-Savart-Laplace* law, the magnetic field will scale with the current if heating is not considered; meanwhile the distance has a great impact to the generation of magnetic field (see Figure 5.15). For example, in the case of $L_w = 0.5 T_w$, 10-20 mA current can create 10-20 mT magnetic field. However, if $L_w = T_w$, the same current only yield 6-12 mT field. In order to generate a magnetic field as high as possible with a current as low as possible, the thickness limitation of insulator could be defined as the thickness of metal line ($L_w = T_w$).

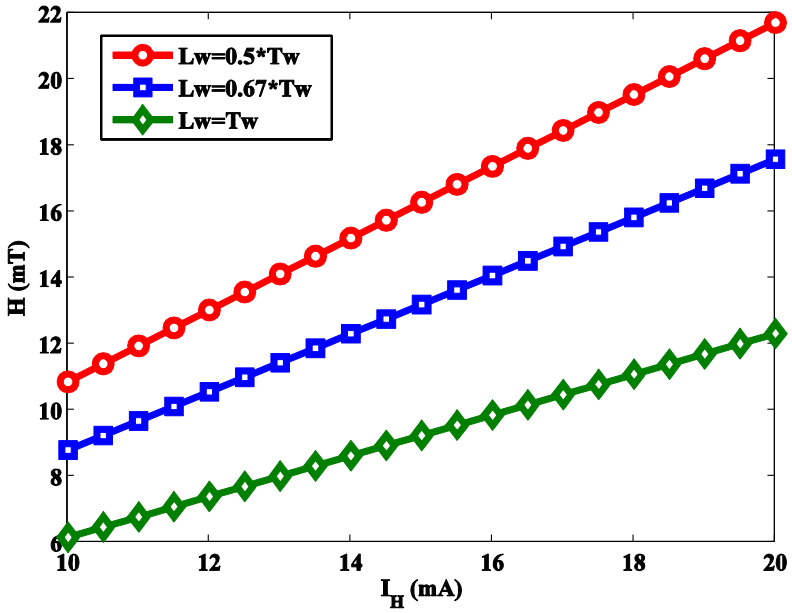


Figure 5.15 Generation of magnetic field by current for different distance between metal wire and racetrack memory.

5.3.3.2 Integration of DW motion 1D model

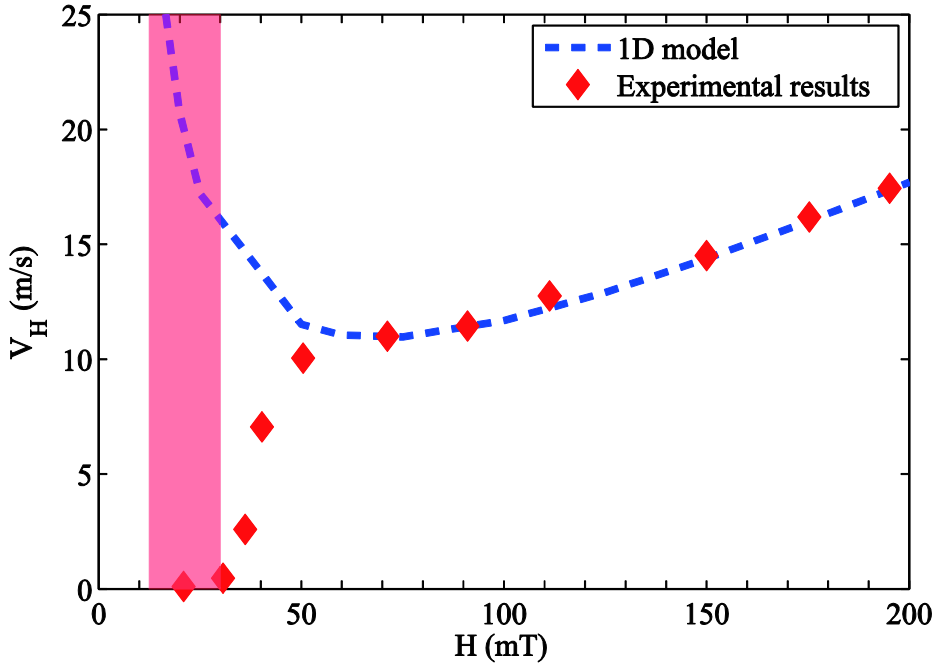


Figure 5.16 Relation between DW motion velocity and magnetic field. A good agreement of 1D model and experimental results appears above 500 Oe. These experimental results are based on Co/Ni material reproduced in [182].

Beyond the model solely concerning the current induced effect, here we integrate 1D model including the field-induced part (see Eqs. 3.22-3.23) in our DW compact model. We then simulate this DW compact model including the contribution of field and compare with the experimental results (see Figure 5.16) [182]. Same as the elucidations of [182], there is a discrepancy between the experimental results and the 1D model below 500 Oe. In order to match the reality, we directly interpolated from the experimental data for fields below 500 Oe, and above this threshold we used the 1D model results. The constants and the default values of parameters used in this model are shown in Table 5.1.

Table 5.1 Parameters in the compact models of PMA racetrack memory with magnetic field assistance

Parameter	Description	Unit	Default value
α	Gilbert damping constant		0.045
P	Spin polarization rate		0.49
M_s	Saturation magnetization	MA/m	0.66
γ	Gyromagnetic ratio	THz/T	0.176
λ	DW width	nm	10
$K_\perp$	Hard axis magnetic anisotropy	MJ/m^3	0.41
TMR	TMR of write/read head MTJ		120%
$J_{c\_nucleation}$	DW nucleation critical current density	GA/m^2	57
J_{c_motion}	DW motion critical current density	TA/m^2	0.31 (assisted) 0.47 (non-assisted)

Importantly, for the reason that the magnetic nanowire simultaneously stores multiple alternating domains, a strong magnetic field may drive alternating DWs in opposite directions, which leads to the annihilation of bits. To avoid doing so, we should apply a relatively low magnetic field that only plays a role in depinning the DWs. Especially, applying a field around the red marking region in Figure 5.16 (e.g. 20 mT) has a negligible contribution to the overall velocity of DW. In this situation, the current always acts as the dominant role to shift the DWs and governs the direction of shifting. On the other hand, despite the polarity of the field, the field will always assist to decrease the critical threshold current, which is relevant to our requirement for lowering current for DW shifting (3.1×10^{11} A/m^2 at 20 mT) and favorable for the implementation of field assistance.

5.3.3.3 Validation of proposed PMA racetrack memory

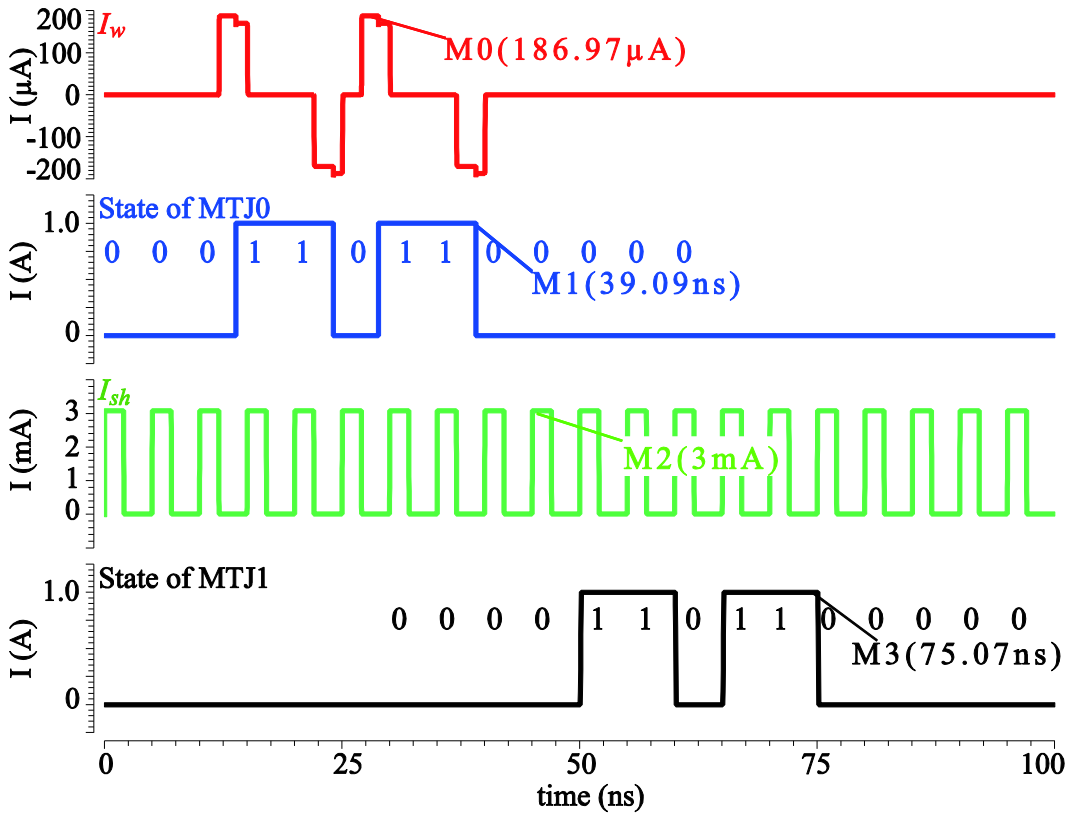


Figure 5.17 Transient simulation of 8-bit PMA racetrack memory with magnetic field assistance. Current pulse I_w is for switching the state of MTJ0, the data are correctly read by MTJ1 after 8 pulses of DW shifting current.

Based on aforementioned parameters and peripheral circuit (writing, sensing and shifting circuits), Figure 5.17 and Figure 5.18 demonstrate respectively the transient simulation results of an 8 and 16 bits PMA Co/Ni racetrack memory with magnetic field assistance. In these simulations, we apply a current of ~20 mA to generate a magnetic field of 19 mT. Current threshold for shifting under this condition is 3.1 x 10^{11} A/m^2 (compared to the intrinsic critical current 4.7 x 10^{11} A/m^2). As shown by the simulation results, the states of storage data along the magnetic nanowire are initialized (logic '0' for 8-bit case and logic '1' for 16-bit case). A writing pulse passing through the write head (MTJ0) generates a logic pattern ("…0110110…" for 8-bit case and "…011010…" for 16-bit case). I_w in these figures demonstrates the passing current during the writing operation, from which we can find the switches of state between two resistance levels. When we inject a current higher than the threshold, the read head (MTJ1) probes the same

data emitted by write head (MTJ0) after 8 or 16 pulses of shifting current. These simulations validate the functionality of magnetic field assisted racetrack memory. Moreover, from the consumption's point of view, the magnetic field should not always be on. Except being on during the DW shifting operation, it should be activated as rarely as possible (see inset of Figure 5.18).

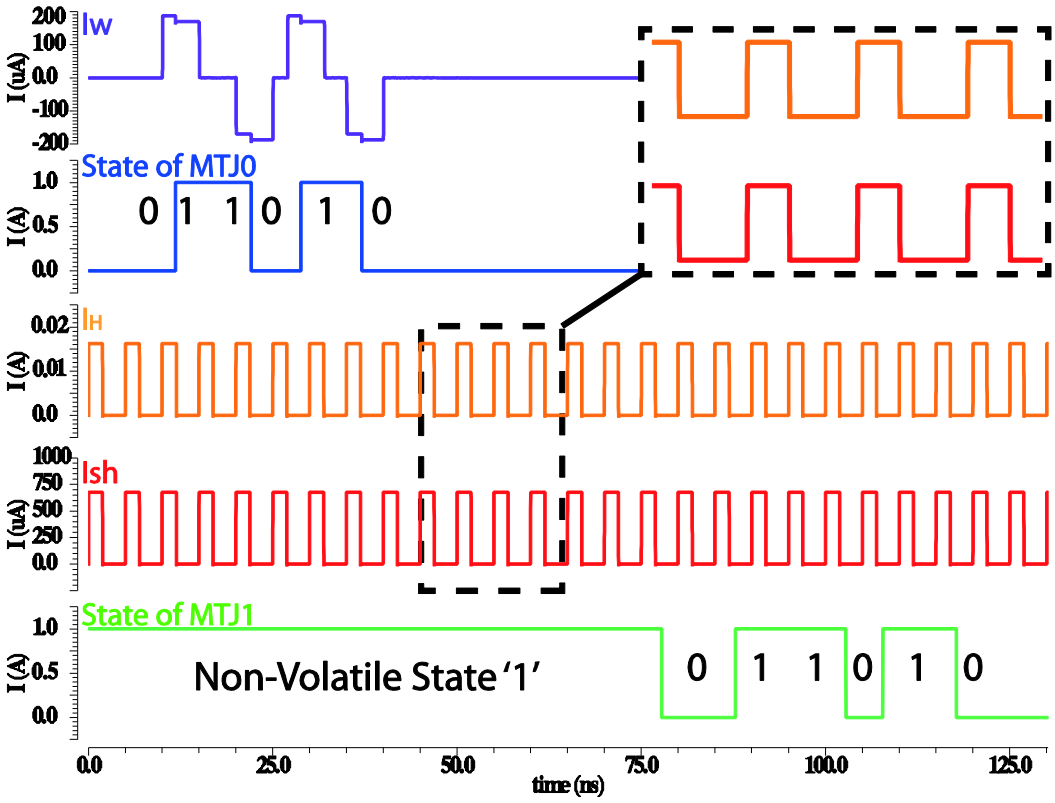


Figure 5.18 Transient simulation of 16-bit PMA racetrack memory with magnetic field assistance. Current pulse I_w is for nucleation in the write head (MTJ0), the data are correctly read by MTJ1 after 16 pulses of DW shifting current.

5.3.3.4 Capacity, area and energy analyses of proposed PMA racetrack memory

As the limitation of CMOS peripheral circuit serving for racetrack memory, the voltage supply cann't arbitrarily increase. It imposes a serious restriction for the application of racetrack memory, particularly on the high-density potential. Lowering the shifting current threshold by magnetic-field-assisted alternative can alleviate this restriction. Supposing that the voltage supply is fixed at 3 V, Figure 5.19(a) shows the maximal length of magnetic nanowire (L_{Max}) depending

on different shifting critical current densities. In addition, the magnetic nanowire thickness has also been studied.

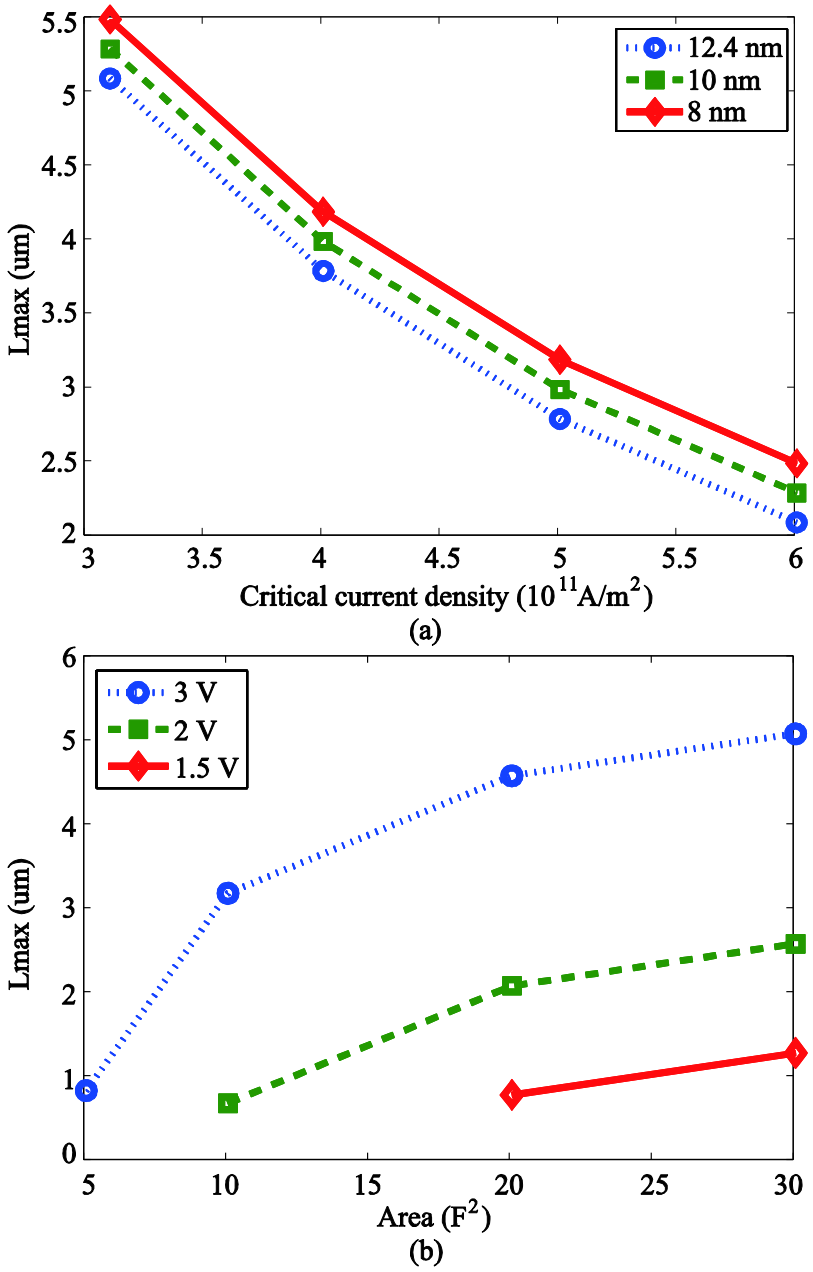


Figure 5.19 Dependence of maximal length of magnetic nanowire versus (a) critical current density of DW shifting at different nanowire thickness. (b) shifting NMOS (NM3) die area with different voltage supplies.

We can find that reducing the shifting critical current density can make the magnetic nanowire longer. Furthermore, the thinner film can rebound to higher density whereas its effect is limited.

For example, one nanowire can store as many as 128 bits of data in 40 nm technology node with a critical current density 3.1 x 10^{11} A/m^2 and a thickness of 12.4 nm. Figure 5.19(b) illustrates the tradeoff between L_{Max} and the surface of shifting NMOS (NM3 shown in Figure 2.14). Enormous surface overhead can achieve more bits stored in racetrack memory. The voltage supply has a great impact on the storage capacity as well. For example, a voltage supply of 3 V allows two and half times of capacity of 2 V with a 20 F^2 NMOS.

In this design, aiming to tackle the bottleneck of feasibility of racetrack memory, we add the currents to induce magnetic field which assists to shift DWs. Nevertheless, it impacts overall power consumption, which is also of importance to racetrack memory. The energy dissipation during the DW shifting consists of the energy consumed by current and field, which can be described by Eq. 5.3:

$$E_{sh} = V_I \times I_I \times t + V_H \times I_H \times t \tag{5.3}$$

where V_I and V_H are the voltage supplies for shifting current I_I and current I_H generating magnetic field, t is the DW shifting duration. When no magnetic field is implemented for racetrack memory, the energy is only consumed by shifting current (the first term in the Eq. 5.3). However, as the voltage supply cannot increase unlimitedly, the storage capacity will easily reach the limitation. Applying magnetic field can make RM storage increase continuously.

Figure 5.20 shows the relation between storage capacity and power consumption for different assistance intensity of magnetic field. Especially, as larger storage capacity racetrack memory shift more bits at the same time, we thus consider the power consumption per bit in this analysis. Due to the limit of peripheral CMOS circuits and for the low-power purpose, V_I is fixed to 3 V in the case with magnetic field assistance and I_I is equal to critical current for the case without magnetic field. Considering the distance between two adjacent DWs is 40 nm, we find that there is a tradeoff when improving the storage capacity: more bits stored per racetrack memory require higher magnetic field and more energy consumption. For instance, 128 bits per racetrack memory necessitates a critical current as low as 3.4 x 10^{11} A/m^2 and consumes 11 pJ/bit for DW shifting. On the other hand, although the power consumption for generating magnetic field is much higher than that for DW shifting, it can still be controlled in an understandable extent (order of pJ/bit),

which benefits for this magnetic field assisting implementation. Depending on the different requirements (low power or high capacity), we can decide whether the magnetic field should be implemented.

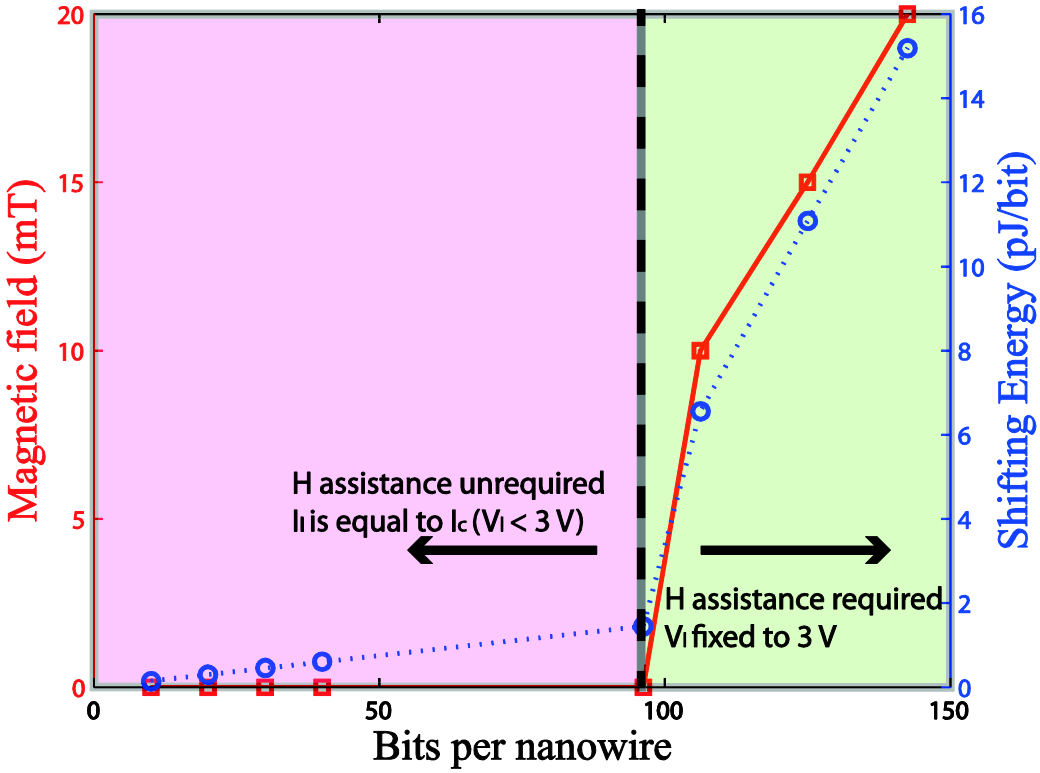


Figure 5.20 Dependence of magnetic field required and shifting energy versus different number of bits per nanowire in racetrack memory.

5.4 Conclusion

In this chapter, we proposed two design optimizations, which are the MLC for STT-MRAM and the magnetic field assistance for PMA racetrack memory.

The MLC design is based on the STT stochastic switching behavior, which allows high speed, low power and high storage density. Two types of structures, in parallel and in serial, were introduced and analyzed. By using a stochastic compact model of PMA STT MTJ and CMOS 40 nm design kit, these structures were validated by Monte-Carlo statistical simulations. The impacts of transistor area and switching pulse magnitude on the programming speed have been also investigated. This MLC cell has a great potential to achieve dense, fast and low-power non-volatile memory chips and neuromorphic computing systems.

By studying and analyzing the high critical current issue leading to the capacity bottleneck of racetrack memory, we found that it had difficulty in generating a high and stable current by peripheral CMOS circuits or reducing the resistivity by optimizing the materials. Recent experimental observations revealed that DW motion could be triggered below the critical current, which inspired us to design a racetrack memory with magnetic field assistance. We performed mixed simulations to validate this design's functionality and demonstrate its enhancing performance. The reduction of DW shifting current benefiting from magnetic field assistance can make racetrack memory more feasible to achieve ultra-denser storage. The additional consumption caused by magnetic field can be confined to a relatively acceptable extent whereas it has a tradeoff relation with storage capacity.

Chapter 6 Gerenal conclusions and perspectives

6.1 General conclusions

This thesis focused on two current-induced switching technologies: the PMA STT MTJ and the PMA racetrack memory. The work covers from compact modeling to hybrid circuit design and optimization. Through this work, the integration functionality of the current-induced spintronic deviecs based on PMA materials has been proven. The performance analyses of related hybrid logic and memory circuits demonstrated that PMA could provide various advantages compared with the in-plane magnetic anisotropy, such as high thermal stability, scalability, low switching current and high operation speed. Thanks to these, the application potential of PMA spintronic devices to achieve future low-power high-density high-speed electronic systems can be confirmed.

In order to better comprehend the related principles and the background of spintronics involved in this thesis, the state of the art of current research field was firstly presented. Since the GMR effect was discovered in 1980s, spintronics has revolutionized both storage and computing technologies. Beyond the spin valves that have been applied in the HDDs, the MTJ is one of the most important spintronic devices. Various switching approaches for MTJ were introduced. Among them, current-induced magnetization switching, i.e. STT, draws significant attention because of its excellent performances on power consumption, scalability and speed. On the other hand, DW motion in magnetic nanowire is an essential phenomenon to be used for storing and operating information. CIDW motion shows its advantages in many aspects compared with field-induced DW motion. Besides them, we also mentioned some other emerging spintronic devices; however they are not yet mature to be widely applied. Based on these spintronic devices, some representative hybrid spintronic circuits were demonstrated. For example, the MRAM is considered an excellent candidate for the next generation of universal memory; racetrack memory can be built to achieve ultra-high density and frequency; logic-in-memory is a potential architecture to replace the conventional Von-Neumann architecture, which can offer instant off/on, high speed and efficient area for computing logic systems. It is noteworthy that these spintronic devices and concepts are based on PMA materials, which can provide higher thermal stability in small size. This steers the trend of current research.

According to the state of the art, we concentrated our investigation on the PMA STT MTJ and the PMA racetrack memory. The compact models of them were firstly developed. In order to obtain good performance accuracy and agreement with experimental results, the compact modeling relied on numbers of theoretical equations and realistic parameters. Oxide barrier tunnel resistance model and bias-voltage-dependent TMR model were used to assess the resistance feature. STT switching static, dynamic and stochastic models were used to describe the STT switching process. In particular, Néel-brown model was used for the thermally assisted regime where the current is lower than the critical current; Sun model was used for the precessional regime where the current exceeds the critical current. 1D DW motion model was embedded to characterize the DW motions in racetrack memory. These compact models were programmed in Verilog-A language, provideing an easy configuration interface and allowing for SPICE compatibility in most of design platforms. By performing DC, transient and Monte-Carlo simulations in Cadence, the functionalities of these compact models were validated. In accordance to existing experiments, the critical currents for switching PMA STT MTJ compact model were 50 µA in average. The speed of CIDW can reach dozens of m/s with a current density in order of 10^8 A/cm^2. In addition, by taking the stochastic behavior into account, the magnetization switching of PMA STT MTJ and DW propagation in racetrack memory were proven to follow the specific distribution possibility.

By utilizing the compact models, certain hybrid logic and memory circuits based on PMA STT MTJ and PMA racetrack memory were designed and analyzed. PCSA was investigated as an example of sensing circuit to output the state of MTJ in logic value. 1-bit and 16k-bit PCSAs were respectively simulated. It was found that this sensing operation could provide the best tradeoff between energy (~10 fJ) and sensing disturbance (less than 10^{-10}%) while keeping a high speed (~100 ps). Its high performance is especially suitable for the logic circuits. A writing circuit to switch a couple of MTJs with complementary states was studied. As the writing circuit dominates the area efficiency compared with the sensing circuit, we analyzed its speed, power consumption and reliability as a function of the die area. It showed that, for a CoFeB/MgO PMA STT MTJ in 65 nm diameter, the writing circuit provided a good tradeoff among the area (~0.096 µm^2 or 30 F^2), power (1 pJ) and speed (~500 MHz) to build both logic and memory chips. Further, MFA was as an example of spintronic logic applications to be investigated. 1-bit MFA based on PMA STT MTJ and multi-bit MFA based on PMA racetrack memory were respectively

proposed. Compared with CMOS only full adders, they showed the considerable advantages in terms of static power and density thanks to non-volatility and 3D integration. Although their frequency and overall power were not very satisfactory, we believe that, with the shrinking of size and the improvement of structure, these drawbacks would be alleviated. The CAM based on PMA racetrack memory was designed, simulated and analyzed as an example of spintronic memory applications. As the comparison circuit requires the complementary states of MTJs, we proposed a dual-track structure. Benefiting from the CIDW propagation and sharing of CMOS peripheral circuits, this design increased the area and power efficiency compared with the conventional CAMs while keeping a comparable searching speed. However, programming of new data into the CAM still cost a lot of time and power, which should be improved in the future.

One of the most serious obstacles for wide application of STT-MRAM and racetrack memory is the low density. MLC based STT-MRAM was proposed to enhance the density. Serial and parallel structures were simulated and analyzed. Note that this design made use of STT stochastic behavior to achieve ultra-high speed. By strictly and properly controlling the switching process, for example applying "Self-Enable" switching circuit, this MLC can contribute to future implementation of STT-MRAM. In order to overcome the high critical current issue that limits the capacity of racetrack memory, racetrack memory with magnetic field assistance was proposed. The optimizations at the level of peripheral circuit and material were firstly discussed and compared. However, these optimizations had their shortcomings: the peripheral circuit was difficult to generate a current not only high but also stable for DW propagation; material optimization didn't meet the requirement and still depended strongly on the technology advancement. Due to the recent progress, magnetic field could assist to trigger the DW motion below the critical current. The structure and implementation of the racetrack memory with magnetic field assistance were presented. Its functionality was validated by mixed simulations, through which the enhancing performance was demonstrated. The tradeoff dependence between consumed energy and capacity was analyzed. It was shown that, although this alternative would increase the overall energy due to the generation of magnetic field, it could be confined to a relatively acceptable extent with regard to the large capacity (in the order of pJ/bit).

6.2 Perspectives

The emergence of spintronics is to achieve more efficient and reliable applications, which could overcome the issues of mainstream charge-based electronics. The term "efficient" here concerns many factors, which involve power, density and frequency, etc. This aim is the "beacon", which indicates the direction of the progress of spintronics. On this route, the innovative technologies are appearing ceaselessly, and an emerging mechanism would be replaced by a more emerging alternative. Here we briefly mention three effects which will be the continuance of our work.

1. High spin torque efficiency

PMA allows the MTJ size to scale down around ~40 nm. Beyond this bound, high spin torque efficiency was recently observed, which is occasionally 2-10 times of that predicted by macrospin models. Some investigation found that this phenomenon was related to sub-volume thermal excitation [43, 188], depending strongly on the strength of PMA and exchange stiffness. By optimizing PMA and exchange stiffness, the imporved spin torque efficiency could ensure a low critical current and a high thermal stability, which offers a promising solution for the further miniaturization of MTJs. As a result, the high spin troque efficiency should be considered in the future work involving the small-size MTJs and hybrid cirucits.

2. Spin orbit torque

As stated in the bibliographic chapter, spin orbit torque (SOT) is demonstrated to switch magnetization and nucleate DWs. Two main effects referred to SOT have been observed: spin hall effect (SHE) and Rashba effect. Compared with the STT switching mechanism, these effects are exhibited with assets in terms of power, speed and reliability. For example, three-terminal devices based on SHE can yield a more efficient spin torque which means to require a lower current. As a consequence, the power consumption can be further reduced. In addition, the current is not applied through the tunnel junction, which avoids the damage caused by the high current density. Furthermore, the separation of writing path and reading path can improve greatly the reliability performance.

3. Voltage-controlled effect

Whether STT or SOT, the mechanisms are based on the current. If there is a current, the power consumption will inevitably be produced. Voltage-controlled effect was recently observed, by which ultra-low power can be expected and then achieved. With regard to the racetrack memory that we have investigated in this thesis, voltage-controlled DW motion can be applied to carry out pinning or depinning operations. The common idea for DW pinning in racetrack memory is using geometric variations or material parameter modifications. However, they are permanent and normally lead to the increase of resistance, which is adverse to the flexibility and feasibility of racetrack memory. Voltage-controlled approach provides a possible pathway to create various new spintronic devices.

References

[1] C. Chappert, A. Fert and F. Nguyen Van Dau, "The emergence of spin electronics in data storage", *Nat Mater*, vol. 6, pp. 813-823, 2007.

[2] N. F. Mott and H. Jones, "Theory of the properties of metals and alloys". *Oxford University Press*, 1936.

[3] M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich and J. Chazelas. "Giant Magnetoresistance of (001)Fe/(001)Cr Magnetic Superlattices". *Physical Review Letters*, vol. 61, pp. 2472, 1988.

[4] G. Binasch, P. Grünberg, F. Saurenbach and W. Zinn, "Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange", *Phys. Rev. B*, vol. 39, pp. 4828-4830, 1989.

[5] I. Žutić, J. Fabian, et S. Das Sarma, "Spintronics: Fundamentals and applications", *Rev. Mod. Phys.*, vol. 76, pp. 323-410, 2004.

[6] R. H. Silsbee, "Novel method for the study of spin transport in conductors", *Bull. Magn. Reson.*, vol. 2, pp. 284-285, 1980.

[7] A. G. Aronov, "Spin injection in metals and polarization of nuclei", *Soviet Journal of Experimental and Theoretical Physics Letters*, vol. 24, pp. 32, 1976.

[8] B. Dieny, V. S. Speriosu, S. S. P. Parkin, B. A. Gurney, D. R. Wilhoit and D. Mauri, "Giant magnetoresistive in soft ferromagnetic multilayers", *Phys. Rev. B*, vol. 43, pp. 1297-1300, 1991.

[9] S. Parkin, X. Jiang, C. Kaiser, A. Panchula, K. Roche and M. Samant, "Magnetically engineered spintronic sensors and memory", *Proceedings of the IEEE*, vol. 91, pp. 661-680, 2003.

[10] A. Barthélémy, A. Fert, J.-P. Contour, M. Bowen, V. Cros, J. M. De Teresa, A. Hamzic, J. C. Faini, J. M. George, J. Grollier, F. Montaigne, F. Pailloux, F. Petroff and C. Vouille, "Magnetoresistance and spin electronics", *Journal of Magnetism and Magnetic Materials*, vol. 242-245, pp. 68-76, 2002.

[11] J. M. Rowell, M. Gurvitch, et J. Geerk, "Modification of tunneling barriers on Nb by a few monolayers of Al", *Phys. Rev. B*, vol. 24, pp. 2278-2281, 1981.

[12] J. S. Moodera, L. R. Kinder, T. M. Wong and R. Meservey, “Large Magnetoresistance at Room Temperature in Ferromagnetic Thin Film Tunnel Junctions”, *Phys. Rev. Lett.*, vol. 74, pp. 3273-3276, pp. L231-L234, 1995.

[13] T. Miyazaki and N. Tezuka, “Giant magnetic tunneling effect in Fe/Al2O3/Fe junction”, *Journal of Magnetism and Magnetic Materials*, vol. 139, 1995.

[14] M. Julliere, “Tunneling between ferromagnetic films”, *Physics Letters A*, vol. 54, pp. 225-226, 1975.

[15] S. Yuasa, T. Nagahama, A. Fukushima, Y. Suzuki and K. Ando, “Giant room-temperature magnetoresistance in single-crystal Fe/MgO/Fe magnetic tunnel junctions”, *Nat Mater*, vol. 3, pp. 868-871, 2004.

[16] S. Yuasa, A. Fukushima, H. Kubota, Y. Suzuki and K. Ando, “Giant tunneling magnetoresistance up to 410% at room temperature in fully epitaxial Co/MgO/Co magnetic tunnel junctions with bcc Co(001) electrodes”, *Applied Physics Letters*, vol. 89, pp. 042505, 2006.

[17] S. Ikeda, J. Hayakawa, Y. Ashizawa, Y. M. Lee, K. Miura, H. Hasegawa, M. Tsunoda, F. Matsukura and H. Ohno, “Tunnel magnetoresistance of 604% at 300K by suppression of Ta diffusion in CoFeB/MgO/CoFeB pseudo-spin-valves annealed at high temperature”, *Applied Physics Letters*, vol. 93, pp. 082508, 2008.

[18] W. S. Zhao, C. Chappert, V. Javerliac and J.-P. Noziere, “High Speed, High Stability and Low Power Sensing Amplifier for MTJ/CMOS Hybrid Logic Circuits”, *IEEE Trans. Magn.*, vol. 45, pp. 3784-3787, 2009.

[19] S. A. Wolf, D. D. Awschalom, R. A. Buhrman, J. M. Daughton, S. von Molnár, M. L. Roukes, A. Y. Chtchelkanova and D. M. Treger, “Spintronics: A Spin-Based Electronics Vision for the Future”, *Science*, vol. 294, pp. 1488-1495, 2001.

[20] W. J. Gallagher and S. S. P. Parkin, “Development of the magnetic tunnel junction MRAM at IBM: From first junctions to a 16-Mb MRAM demonstrator chip”, *IBM Journal of Research and Development*, vol. 50, pp. 5-23, 2006.

[21] I. L. Prejbeanu, M. Kerekes, R. C. Sousa, H. Sibuet, O. Redon, B. Dieny and J. P. Nozières, “Thermally assisted MRAM”, *J. Phys.: Condens. Matter*, vol. 19, pp. 165218, 2007.

[22] I. L. Prejbeanu, S. Bandiera, J. Alvarez-Hérault, R. C. Sousa, B. Dieny and J.-P. Nozières, “Thermally assisted MRAMs: ultimate scalability and logic functionalities”, *J. Phys. D: Appl. Phys.*, vol. 46, pp. 074002, 2013.

[23] I. L. Prejbeanu, W. Kula, K. Ounadjela, R. C. Sousa, O. Redon, B. Dieny and J.-P. Nozieres, "Thermally assisted switching in exchange-biased storage layer magnetic tunnel junctions", *IEEE Transactions on Magnetics*, vol. 40, n^{o} 4, p. 2625-2627, juill. 2004.

[24] J. C. Slonczewski, "Current-driven excitation of magnetic multilayers", *Journal of Magnetism and Magnetic Materials*, vol. 159, pp. L1-L7, 1996.

[25] L. Berger, "Emission of spin waves by a magnetic multilayer traversed by a current", *Phys. Rev. B*, vol. 54, pp. 9353-9358, 1996.

[26] D. C. Ralph and M. D. Stiles, "Spin transfer torques", *Journal of Magnetism and Magnetic Materials*, vol. 320, pp. 1190-1216, 2008.

[27] Z. Diao, Z. Li, S. Wang, Y. Ding, A. Panchula, E. Chen, L.-C. Wang, and Y. Huai, "Spin-transfer torque switching in magnetic tunnel junctions and spin-transfer torque random access memory", *J. Phys.: Condens. Matter*, vol. 19, pp. 165209, 2007.

[28] E. B. Myers, D. C. Ralph, J. A. Katine, R. N. Louie and R. A. Buhrman, "Current-Induced Switching of Domains in Magnetic Multilayer Devices", *Science*, vol. 285, pp. 867-870, 1999.

[29] J. A. Katine, F. J. Albert, R. A. Buhrman, E. B. Myers and D. C. Ralph, "Current-Driven Magnetization Reversal and Spin-Wave Excitations in Co /Cu /Co Pillars", *Phys. Rev. Lett.*, vol. 84, pp. 3149-3152, 2000.

[30] J. Z. Sun and D. C. Ralph, "Magnetoresistance and spin-transfer torque in magnetic tunnel junctions", *Journal of Magnetism and Magnetic Materials*, vol. 320, pp. 1227-1237, 2008.

[31] A. Kalitsov, M. Chshiev, I. Theodonis, N. Kioussis, and W. H. Butler, "Spin-transfer torque in magnetic tunnel junctions", *Phys. Rev. B*, vol. 79, pp. 174416, 2009.

[32] Y. M. Huai, "Spin-transfer torque MRAM (STT-MRAM): challenges and prospects", AAPPS Bulletin, vol. 18, pp. 33–40, 2008.

[33] T. Kawahara, K. Ito, R. Takemura, and H. Ohno, "Spin-transfer torque RAM technology: Review and prospect", *Microelectronics Reliability*, vol. 52, pp. 613-627, 2012.

[34] H. Xi, J. Stricklin, H. Li, Y. Chen, X. Wang, Y. Zheng, Z. Gao and M. X. Tang, "Spin Transfer Torque Memory With Thermal Assist Mechanism: A Case Study", *IEEE Transactions on Magnetics*, vol. 46, pp. 860-865, 2010.

[35] B. Dieny, R. C. Sousa, J. Herault, C. Papusoi, G. Prenat, U. Ebels, D. Houssameddine, B. Rodmacq, S. Auffret, L. D. Buda-Prejbeanu, M. C. Cyrille, B. Delaet, O. Redon, C. Ducruet and

J. P. Nozières, “Spin-transfer effect and its use in spintronic components”, *International Journal of Nanotechnology*, vol. 7, pp. 591-614, 2010.

[36] S. Ikeda, H. Sato, M. Yamanouchi, H. Gan, K. Miura, K. Mizunuma, S. Kanai, S. Fukami, F. Matsukura, N. Kasai and H. Ohno, “Recent progress of perpendicular anisotropy magnetic tunnel junctions for nonvolatile VLSI”, *SPIN*, vol. 2, pp. 1240003, 2012.

[37] N. Nishimura, T. Hirai, A. Koganei, T. Ikeda, K. Okano, Y. Sekiguchi and Y. Osada, “Magnetic tunnel junction device with perpendicular magnetization films for high-density magnetic random access memory”, *Journal of Applied Physics*, vol. 91, pp. 5246-5249, 2002.

[38] S. Ikeda, K. Miura, H. Yamamoto, K. Mizunuma, H. D. Gan, M. Endo, S. Kanai, J. Hayakawa, F. Matsukura and H. Ohno, “A perpendicular-anisotropy CoFeB–MgO magnetic tunnel junction”, *Nat Mater*, vol. 9, pp. 721-724, 2010.

[39] D. C. Worledge, G. Hu, D. W. Abraham, J. Z. Sun, P. L. Trouilloud, J. Nowak, S. Brown, M. C. Gaidis, E. J. O’Sullivan and R. P. Robertazzi, “Spin torque switching of perpendicular Ta|CoFeB|MgO-based magnetic tunnel junctions”, *Appl. Phys. Lett.*, vol. 98, pp. 022501, 2011.

[40] M. Gajek, J. J. Nowak, J. Z. Sun, P. L. Trouilloud, E. J. O’Sullivan, D. W. Abraham, M. C. Gaidis, G. Hu, S. Brown, Y. Zhu, R. P. Robertazzi, W. J. Gallagher and D. C. Worledge, “Spin torque switching of 20 nm magnetic tunnel junctions with perpendicular anisotropy”, *Applied Physics Letters*, vol. 100, pp. 132408, 2012.

[41] D. C. Worledge, M. Gajek, D. W. Abraham, S. Brown, M. C. Gaidis, G. Hu, J. Nowak, E. J. O’Sullivan, R. P. Robertazzi, J. Z. Sun, P. L. Trouilloud and W. J. Gallagher, “Recent Advances in Spin Torque MRAM”, *proceeding in 4th IEEE International Memory Workshop (IMW),* pp. 1-3, 2012.

[42] Y. Chen, H. Li, X. Wang, W. Zhu, W. Xu and T. Zhang, “A 130 nm 1.2 V/3.3 V 16 Kb Spin-Transfer Torque Random Access Memory With Nondestructive Self-Reference Sensing Scheme”, *IEEE Journal of Solid-State Circuits*, vol. 47, pp. 560-573, 2012.

[43] J. Z. Sun, R. P. Robertazzi, J. Nowak, P. L. Trouilloud, G. Hu, D. W. Abraham, M. C. Gaidis, S. L. Brown, E. J. O’Sullivan, W. J. Gallagher and D. C. Worledge, “Effect of subvolume excitation and spin-torque efficiency on magnetic switching”, *Phys. Rev. B*, vol. 84, pp. 064413, 2011.

[44] J. E. Hirsch, “Spin Hall Effect”, *Phys. Rev. Lett.*, vol. 83, pp. 1834-1837, 1999.

[45] L. Liu, C.-F. Pai, Y. Li, H. W. Tseng, D. C. Ralph and R. A. Buhrman, "Spin-Torque Switching with the Giant Spin Hall Effect of Tantalum", *Science*, vol. 336, pp. 555-558, 2012.

[46] T. Maruyama, Y. Shiota, T. Nozaki, K. Ohta, N. Toda, M. Mizuguchi, A. A. Tulapurkar, T. Shinjo, M. Shiraishi, S. Mizukami, Y. Ando and Y. Suzuki, "Large voltage-induced magnetic anisotropy change in a few atomic layers of iron", *Nat Nano*, vol. 4, pp. 158-161, 2009.

[47] K. L. Wang, J. G. Alzate and P. K. Amiri, "Low-power non-volatile spintronic memory: STT-RAM and beyond", *J. Phys. D: Appl. Phys.*, vol. 46, pp. 074003, 2013.

[48] Y. Shiota, T. Nozaki, F. Bonell, S. Murakami, T. Shinjo and Y. Suzuki, "Induction of coherent magnetization switching in a few atomic layers of FeCo using voltage pulses", *Nat Mater*, vol. 11, pp. 39-43, 2012.

[49] W.-G. Wang, M. Li, S. Hageman and C. L. Chien, "Electric-field-assisted switching in magnetic tunnel junctions", *Nat Mater*, vol. 11, pp. 64-68, 2012.

[50] A. Hubert and R. Schafer, "Magnetic Domains - The Analysis of Magnetic Microstructures", *Springer*, 2001.

[51] J. Shibata, G. Tatara and H. Kohno, "A brief review of field- and current-driven domain-wall motion", *J. Phys. D: Appl. Phys.*, vol. 44, pp. 384004, 2011.

[52] A. H. Bobeck, "Properties and Device Applications of Magnetic Domains in Orthoferrites", *Bell System Technical Journal*, vol. 46, pp. 1901-1925, 1967.

[53] J. C. Slonczewski, "Theory of domain-wall motion in magnetic films and platelets", *Journal of Applied Physics*, vol. 44, pp. 1759-1770, 2003.

[54] D. A. Allwood, G. Xiong, C. C. Faulkner, D. Atkinson, D. Petit, et R. P. Cowburn, "Magnetic Domain-Wall Logic", *Science*, vol. 309, pp. 1688-1692, 2005.

[55] L. Berger, "Low-field magnetoresistance and domain drag in ferromagnets", *Journal of Applied Physics*, vol. 49, pp. 2156-2161, 1978.

[56] L. Berger, "Exchange interaction between ferromagnetic domain wall and electric current in very thin metallic films", *Journal of Applied Physics*, vol. 55, pp. 1954-1956, 1984.

[57] P. P. Freitas and L. Berger, "Observation of s-d exchange force between domain walls and electric current in very thin Permalloy films", *Journal of Applied Physics*, vol. 57, pp. 1266-1269, 1985.

[58] L. Berger, "Motion of a magnetic domain wall traversed by fast-rising current pulses", *Journal of Applied Physics*, vol. 71, pp. 2721-2726, 1992.

[59] E. Salhi and L. Berger, “Current-induced displacements of Bloch walls in Ni-Fe films of thickness 120–740 nm”, *Journal of Applied Physics*, vol. 76, pp. 4787-4792, 1994.

[60] A. Yamaguchi, T. Ono, S. Nasu, K. Miyake, K. Mibu and T. Shinjo, “Real-Space Observation of Current-Driven Domain Wall Motion in Submicron Magnetic Wires”, *Phys. Rev. Lett.*, vol. 92, pp. 077205, 2004.

[61] M. Kläui, C. A. F. Vaz, J. A. C. Bland, W. Wernsdorfer, G. Faini, E. Cambril, L. J. Heyderman, F. Nolting and U. Rüdiger, “Controlled and Reproducible Domain Wall Displacement by Current Pulses Injected into Ferromagnetic Ring Structures”, *Phys. Rev. Lett.*, vol. 94, pp. 106601, 2005.

[62] M. Hayashi, L. Thomas, C. Rettner, R. Moriya and S. S. P. Parkin, “Direct observation of the coherent precession of magnetic domain walls propagating along permalloy nanowires”, *Nat Phys*, vol. 3, pp. 21-25, 2007.

[63] H. Ohno, “Making Nonmagnetic Semiconductors Ferromagnetic”, *Science*, vol. 281, pp. 951-956, 1998.

[64] O. Boulle, G. Malinowski and M. Kläui, “Current-induced domain wall motion in nanoscale ferromagnetic elements”, *Materials Science and Engineering: R: Reports*, vol. 72, pp. 159-187, 2011.

[65] D. Ravelosona, S. Mangin, Y. Lemaho, J. A. Katine, B. D. Terris, et E. E. Fullerton, “Domain Wall Creation in Nanostructures Driven by a Spin-Polarized Current”, *Phys. Rev. Lett.*, vol. 96, pp. 186604, 2006.

[66] D. Ravelosona, D. Lacour, J. A. Katine, B. D. Terris and C. Chappert, “Nanometer Scale Observation of High Efficiency Thermally Assisted Current-Driven Domain Wall Depinning”, *Phys. Rev. Lett.*, vol. 95, pp. 117203, 2005.

[67] T. A. Moore, I. M. Miron, G. Gaudin, G. Serret, S. Auffret, B. Rodmacq, A. Schuhl, S. Pizzini, J. Vogel, et M. Bonfim, “High domain wall velocities induced by current in ultrathin Pt/Co/AlOx wires with perpendicular magnetic anisotropy”, *Applied Physics Letters*, vol. 93, pp. 262504, 2008.

[68] C. Burrowes, A. P. Mihai, D. Ravelosona, J.-V. Kim, C. Chappert, L. Vila, A. Marty, Y. Samson, F. Garcia-Sanchez, L. D. Buda-Prejbeanu, I. Tudosa, E. E. Fullerton, et J.-P. Attané, “Non-adiabatic spin-torques in narrow magnetic domain walls”, *Nat Phys*, vol. 6, pp. 17-21, 2010.

[69] Y. Yoshimura, T. Koyama, D. Chiba, Y. Nakatani, S. Fukami, M. Yamanouchi, H. Ohno, et T. Ono, "Current-Induced Domain Wall Motion in Perpendicularly Magnetized Co/Ni Nanowire under In-Plane Magnetic Fields", *Appl. Phys. Express*, vol. 5, pp. 063001, 2012.

[70] N. Ohshima, T. Koyama, H. Tanigawa, M. Kotsugi, T. Ohkouchi, D. Chiba, T. Kinoshita and T. Ono, "Real space observation of current-induced magnetic domain wall displacement in Co/Ni nano-wire by photoemission electron microscopy", *J. Phys.: Condens. Matter*, vol. 23, pp. 382202, 2011.

[71] S. Fukami, T. Suzuki, Y. Nakatani, N. Ishiwata, M. Yamanouchi, S. Ikeda, N. Kasai, et H. Ohno, "Current-induced domain wall motion in perpendicularly magnetized CoFeB nanowire", *Applied Physics Letters*, vol. 98, pp. 082504, 2011.

[72] F.-T. Yuan, Y.-H. Lin, J. K. Mei, J.-H. Hsu, et P. C. Kuo, "Effect of thickness of MgO, Co-Fe-B, and Ta layers on perpendicular magnetic anisotropy of [Ta/Co60Fe20B20/MgO]5 multilayered films", *Journal of Applied Physics*, vol. 111, pp. 07C111, 2012.

[73] P. J. Metaxas, J. Sampaio, A. Chanthbouala, R. Matsumoto, A. Anane, A. Fert, K. A. Zvezdin, K. Yakushiji, H. Kubota, A. Fukushima, S. Yuasa, K. Nishimura, Y. Nagamine, H. Maehara, K. Tsunekawa, V. Cros, et J. Grollier, "High domain wall velocities via spin transfer torque using vertical current injection", *Sci. Rep.*, vol. 3, pp. 1829, 2013.

[74] A. J. Schellekens, A. van den Brink, J. H. Franken, H. J. M. Swagten and B. Koopmans, "Electric-field control of domain wall motion in perpendicularly magnetized materials", *Nat Commun*, vol. 3, pp. 847, 2012.

[75] D. Chiba, M. Kawaguchi, S. Fukami, N. Ishiwata, K. Shimamura, K. Kobayashi, et T. Ono, "Electric-field control of magnetic domain-wall velocity in ultrathin cobalt with perpendicular magnetization", *Nat Commun*, vol. 3, p. 888, 2012.

[76] U. Bauer, S. Emori, and G. S. D. Beach, "Electric field control of domain wall propagation in Pt/Co/GdOx films", *Applied Physics Letters*, vol. 100, pp. 192408, 2012.

[77] U. Bauer, M. Przybylski, J. Kirschner and G. S. D. Beach, "Magnetoelectric Charge Trap Memory", *Nano Lett.*, vol. 12, pp. 1437-1442, 2012.

[78] U. Bauer, S. Emori and G. S. D. Beach, "Voltage-gated modulation of domain wall creep dynamics in an ultrathin metallic ferromagnet", *Applied Physics Letters*, vol. 101, pp. 172403, 2012.

[79] U. Bauer, S. Emori and G. S. D. Beach, "Voltage-controlled domain wall traps in ferromagnetic nanowires", *Nat Nano*, vol. 8, pp. 411-416, 2013.

[80] I. M. Miron, T. Moore, H. Szambolics, L. D. Buda-Prejbeanu, S. Auffret, B. Rodmacq, S. Pizzini, J. Vogel, M. Bonfim, A. Schuhl and G. Gaudin, "Fast current-induced domain-wall motion controlled by the Rashba effect", *Nat Mater*, vol. 10, pp. 419-423, 2011.

[81] S. Emori, U. Bauer, S.-M. Ahn, E. Martinez and G. S. D. Beach, "Current-driven dynamics of chiral ferromagnetic domain walls", *Nat Mater*, vol. 12, pp. 611-616, 2013.

[82] Y. A. Bychkov and É. I. Rashba, "Properties of a 2D electron gas with lifted spectral degeneracy", *JETP*, vol. 39, p. 78, 1984.

[83] A. Manchon and S. Zhang, "Theory of spin torque due to spin-orbit coupling", *Phys. Rev. B*, vol. 79, pp. 094422, 2009.

[84] X. Wang and A. Manchon, "Diffusive Spin Dynamics in Ferromagnetic Thin Films with a Rashba Interaction", *Phys. Rev. Lett.*, vol. 108, pp. 117201, 2012.

[85] M. I. D'Yakonov and V. I. Perel', "Possibility of Orienting Electron Spins with Current", *JETP*, vol. 13, p. 657, 1971.

[86] S.-M. Seo, K.-W. Kim, J. Ryu, H.-W. Lee and K.-J. Lee, "Current-induced motion of a transverse magnetic domain wall in the presence of spin Hall effect", *Applied Physics Letters*, vol. 101, pp. 022405, 2012.

[87] J. Ryu, K.-J. Lee and H.-W. Lee, "Current-driven domain wall motion with spin Hall effect: Reduction of threshold current density", *Applied Physics Letters*, vol. 102, pp. 172404, 2013.

[88] K.-S. Ryu, L. Thomas, S.-H. Yang and S. Parkin, "Chiral spin torque at magnetic domain walls", *Nat Nano*, vol. 8, pp. 527-533, 2013.

[89] E. Martinez, S. Emori and G. S. D. Beach, "Current-driven domain wall motion along high perpendicular anisotropy multilayers: The role of the Rashba field, the spin Hall effect, and the Dzyaloshinskii-Moriya interaction", *Applied Physics Letters*, vol. 103, pp. 072406, 2013.

[90] S. Datta and B. Das, "Electronic analog of the electro-optic modulator", *Applied Physics Letters*, vol. 56, pp. 665-667, 1990.

[91] B. Dlubak, M.-B. Martin, C. Deranlot, B. Servet, S. Xavier, R. Mattana, M. Sprinkle, C. Berger, W. A. De Heer, F. Petroff, A. Anane, P. Seneor and A. Fert, "Highly efficient spin transport in epitaxial graphene on SiC", *Nat Phys*, vol. 8, pp. 557-561, 2012.

[92] H. Dery, H. Wu, B. Ciftcioglu, M. Huang, Y. Song, R. Kawakami, J. Shi, I. Krivorotov, I. Zutic, and L. J. Sham, "Nanospintronics Based on Magnetologic Gates", *IEEE Transactions on Electron Devices*, vol. 59, pp. 259-262, 2012.

[93] R. P. Cowburn and M. E. Welland, "Room Temperature Magnetic Quantum Cellular Automata", *Science*, vol. 287, pp. 1466-1468, 2000.

[94] A. Imre, G. Csaba, L. Ji, A. Orlov, G. H. Bernstein and W. Porod, "Majority Logic Gate for Magnetic Quantum-Dot Cellular Automata", *Science*, vol. 311, pp. 205-208, 2006.

[95] S. Breitkreutz, J. Kiermaier, I. Eichwald, C. Hildbrand, G. Csaba, D. Schmitt-Landsiedel, and M. Becherer, "Experimental Demonstration of a 1-Bit Full Adder in Perpendicular Nanomagnetic Logic", *IEEE Transactions on Magnetics*, vol. 49, pp. 4464-4467, 2013.

[96] R. Lavrijsen, J.-H. Lee, A. Fernández-Pacheco, D. C. M. C. Petit, R. Mansell, and R. P. Cowburn, "Magnetic ratchet for three-dimensional spintronic memory and logic", *Nature*, vol. 493, pp. 647-650, 2013.

[97] T. Schneider, A. A. Serga, B. Leven, B. Hillebrands, R. L. Stamps and M. P. Kostylev, "Realization of spin-wave logic gates", *Applied Physics Letters*, vol. 92, pp. 022505, 2008.

[98] A. Khitun, M. Bao and K. L. Wang, "Spin Wave Magnetic NanoFabric: A New Approach to Spin-Based Logic Circuitry", *IEEE Transactions on Magnetics*, vol. 44, pp. 2141-2152, 2008.

[99] A. Khitun and K. L. Wang, "Nano scale computational architectures with Spin Wave Bus", *Superlattices and Microstructures*, vol. 38, pp. 184-200, 2005.

[100] L. O. Chua, "Memristor-The missing circuit element", *IEEE Transactions on Circuit Theory*, vol. 18, pp. 507-519, 1971.

[101] D. B. Strukov, G. S. Snider, D. R. Stewart and R. S. Williams, "The missing memristor found", *Nature*, vol. 453, pp. 80-83, 2008.

[102] A. Chanthbouala, V. Garcia, R. O. Cherifi, K. Bouzehouane, S. Fusil, X. Moya, S. Xavier, H. Yamada, C. Deranlot, N. D. Mathur, M. Bibes, A. Barthélémy and J. Grollier, "A ferroelectric memristor", *Nat Mater*, vol. 11, pp. 860-864, 2012.

[103] X. Wang, Y. Chen, H. Xi, H. Li and D. Dimitrov, "Spintronic Memristor Through Spin-Torque-Induced Magnetization Motion", *IEEE Electron Device Letters*, vol. 30, pp. 294-297, 2009.

[104] K.-H. Kim, S. Gaba, D. Wheeler, J. M. Cruz-Albrecht, T. Hussain, N. Srinivasa, et W. Lu, "A Functional Hybrid Memristor Crossbar-Array/CMOS System for Data Storage and Neuromorphic Applications", *Nano Lett.*, vol. 12, pp. 389-395, 2012.

[105] J. M. Daughton, "Magnetic tunneling applied to memory", *Journal of Applied Physics*, vol. 81, pp. 3758-3763, 1997.

[106] W. Zhao, S. Chaudhuri, C. Accoto, J.-O. Klein, C. Chappert and P. Mazoyer, "Cross-Point Architecture for Spin-Transfer Torque Magnetic Random Access Memory", *IEEE Transactions on Nanotechnology*, vol. 11, pp. 907-917, 2012.

[107] T. M. Maffitt, J. K. DeBrosse, J. A. Gabric, E. T. Gow, M. C. Lamorey, J. S. Parenteau, D. R. Willmott, M. A. Wood and W. J. Gallagher, "Design Considerations for MRAM", *IBM J. Res. Dev.*, vol. 50, pp. 25–39, 2006.

[108] D. Apalkov, A. Khvalkovskiy, S. Watts, V. Nikitin, X. Tang, D. Lottis, K. Moon, X. Luo, E. Chen, A. Ong, A. Driskill-Smith and M. Krounbi, "Spin-transfer Torque Magnetic Random Access Memory (STT-MRAM)", *J. Emerg. Technol. Comput. Syst.*, pp. 13:1–13:35, 2013.

[109] X. Liu, W. Zhang, M. J. Carter and G. Xiao, "Ferromagnetic resonance and damping properties of CoFeB thin films as free layers in MgO-based magnetic tunnel junctions", *Journal of Applied Physics*, vol. 110, pp. 033910, 2011.

[110] T. Devolder, P.-H. Ducrot, J.-P. Adam, I. Barisic, N. Vernier, J.-V. Kim, B. Ockert and D. Ravelosona, "Damping of CoxFe80−xB20 ultrathin films with perpendicular magnetic anisotropy", *Applied Physics Letters*, vol. 102, pp. 022407, 2013.

[111] M. Konoto, H. Imamura, T. Taniguchi, K. Yakushiji, H. Kubota, A. Fukushima, K. Ando, et S. Yuasa, "Effect of MgO Cap Layer on Gilbert Damping of FeB Electrode Layer in MgO-Based Magnetic Tunnel Junctions", *Appl. Phys. Express*, vol. 6, pp. 073002, 2013.

[112] T. Kishi, H. Yoda, T. Kai, T. Nagase, E. Kitagawa, M. Yoshikawa, K. Nishiyama, T. Daibou, M. Nagamine, M. Amano, S. Takahashi, M. Nakayama, N. Shimomura, H. Aikawa, S. Ikegawa, S. Yuasa, K. Yakushiji, H. Kubota, A. Fukushima, M. Oogane, T. Miyazaki and K. Ando, "Lower-current and fast switching of a perpendicular TMR for high speed and high density spin-transfer-torque MRAM", *proceeding in IEDM 2008*, pp. 309-312, 2008.

[113] R. Takemura, T. Kawahara, K. Miura, H. Yamamoto, J. Hayakawa, N. Matsuzaki, K. Ono, M. Yamanouchi, K. Ito, H. Takahashi, S. Ikeda, H. Hasegawa, H. Matsuoka and H. Ohno, "A 32-Mb SPRAM With 2T1R Memory Cell, Localized Bi-Directional Write Driver and `1'/`0' Dual-

Array Equalized Reference Scheme", *IEEE Journal of Solid-State Circuits*, vol. 45, pp. 869-879, 2010.

[114] K. Tsuchida, T. Inaba, K. Fujita, Y. Ueda, T. Shimizu, Y. Asao, T. Kajiyama, M. Iwayama, K. Sugiura, S. Ikegawa, T. Kishi, T. Kai, M. Amano, N. Shimomura, H. Yoda and Y. Watanabe, "A 64Mb MRAM with clamped-reference and adequate-reference schemes", *proceeding in ISSCC* 2010, pp. 258-259, 2010.

[115] D. C. Worledge, G. Hu, P. L. Trouilloud, D. W. Abraham, S. Brown, M. C. Gaidis, J. Nowak, E. J. O'Sullivan, R. P. Robertazzi, J. Z. Sun and W. J. Gallagher, "Switching distributions and write reliability of perpendicular spin torque MRAM", *proceeding in IEDM 2010*, pp. 12.5.1-12.5.4, 2010.

[116] S. Chung, K.-M. Rho, S.-D. Kim, H.-J. Suh, D.-J. Kim, H. J. Kim, S. H. Lee, J.-H. Park, H.-M. Hwang, S.-M. Hwang, J.-Y. Lee, Y.-B. An, J.-U. Yi, Y.-H. Seo, D.-H. Jung, M.-S. Lee, S.-H. Cho, J.-N. Kim, G.-J. Park, G. Jin, A. Driskill-Smith, V. Nikitin, A. Ong, X. Tang, Y. Kim, J.-S. Rho, S.-K. Park, S.-W. Chung, J.-G. Jeong and S.-J. Hong, "Fully integrated 54nm STT-RAM with the smallest bit cell dimension for high density memory application", *proceeding in IEDM 2010*, pp. 12.7.1-12.7.4, 2010.

[117] C. J. Lin, S. H. Kang, Y. J. Wang, K. Lee, X. Zhu, W. C. Chen, X. Li, W. N. Hsu, Y. C. Kao, M. T. Liu, W. C. Chen, Y. Lin, M. Nowak, N. Yu and L. Tran, "45nm low power CMOS logic compatible embedded STT MRAM utilizing a reverse-connection 1T/1MTJ cell", *proceeding in IEDM 2009*, pp. 279282, 2009.

[118] Y. Huai, Y. Zhou, Z. Wang, R. Malmhall, R. Ranjan and J. Zhang, "STT MRAM: recent strong semiconductor industry traction", *proceeding in the 2nd CSIS International Symposium on Spintronics-based VLSIs and The 8th RIEC International Workshop on Spintronics*, pp. 19, 2012.

[119] S. Mangin, D. Ravelosona, J. A. Katine, M. J. Carey, B. D. Terris, et E. E. Fullerton, "Current-induced magnetization reversal in nanopillars with perpendicular anisotropy", *Nat Mater*, vol. 5, pp. 210-215, 2006.

[120] M. Hayashi, L. Thomas, R. Moriya, C. Rettner and S. S. P. Parkin, "Current-Controlled Magnetic Domain-Wall Nanowire Shift Register", *Science*, vol. 320, pp. 209-211, 2008.

[121] S. S. P. Parkin, M. Hayashi and L. Thomas, "Magnetic Domain-Wall Racetrack Memory", *Science*, vol. 320, pp. 190-194, 2008.

[122] International Roadmap for semiconductor (ITRS), 2010 ERD Update.

[123] W. S. Zhao, J. Duval, D. Ravelosona, J.-O. Klein, J. V. Kim and C. Chappert, "A compact model of domain wall propagation for logic and memory design", *Journal of Applied Physics*, vol. 109, pp. 07D501, 2011.

[124] A. J. Annunziata, M. C. Gaidis, L. Thomas, C. W. Chien, C.-C. Hung, P. Chevalier, E. J. O'Sullivan, J. P. Hummel, E. A. Joseph, Y. Zhu, T. Topuria, E. Delenia, P. M. Rice, S. S. P. Parkin and W. J. Gallagher, "Racetrack memory cell array with integrated magnetic tunnel junction readout", *proceeding in IEDM 2011*, pp. 24.3.1-24.3.4, 2011.

[125] M. Nakayama, T. Kai, N. Shimomura, M. Amano, E. Kitagawa, T. Nagase, M. Yoshikawa, T. Kishi, S. Ikegawa and H. Yoda, "Spin transfer switching in TbCoFe/CoFeB/MgO/CoFeB/TbCoFe magnetic tunnel junctions with perpendicular magnetic anisotropy", *Journal of Applied Physics*, vol. 103, pp. 07A710, 2008.

[126] W. Zhao, C. Chappert, V. Javerliac and J.-P. Noziere, "High Speed, High Stability and Low Power Sensing Amplifier for MTJ/CMOS Hybrid Logic Circuits", *IEEE Transactions on Magnetics*, vol. 45, pp. 3784-3787, 2009.

[127] W. H. Kautz, "Cellular Logic-in-Memory Arrays", *IEEE Transactions on Computers*, vol. C-18, pp. 719-727, 1969.

[128] N. S. Kim, T. Austin, D. Baauw, T. Mudge, K. Flautner, J. S. Hu, M. J. Irwin, M. Kandemir and V. Narayanan, "Leakage current: Moore's law meets static power", *Computer*, vol. 36, pp. 68-75, 2003.

[129] D. Patterson, T. Anderson, N. Cardwell, R. Fromm, K. Keeton, C. Kozyrakis, R. Thomas, and K. Yelick, "Intelligent RAM (IRAM): chips that remember and compute", *proceeding in ISSCC 1997*, pp. 224-225, 1997.

[130] T. Hanyu, K. Teranishi, et M. Kameyama, "Multiple-valued logic-in-memory VLSI based on a floating-gate-MOS pass-transistor network", *proceeding in ISSCC 1998*, pp. 194-195, 1998.

[131] R. Waser et M. Aono, "Nanoionics-based resistive switching memories", *Nat Mater*, vol. 6, pp. 833-840, 2007.

[132] W. Zhao, E. Belhaire, V. Javerliac, C. Chappert and B. Dieny, "A non-volatile flip-flop in magnetic FPGA chip", *proceeding in DTIS 2006*, pp. 323-326, 2006.

[133] N. Sakimura, T. Sugibayashi, R. Nebashi and N. Kasai, "Nonvolatile Magnetic Flip-Flop for Standby-Power-Free SoCs", *IEEE Journal of Solid-State Circuits*, vol. 44, pp. 2244-2250, 2009.

[134] S. Matsunaga, J. Hayakawa, S. Ikeda, K. Miura, H. Hasegawa, T. Endoh, H. Ohno and T. Hanyu, "Fabrication of a Nonvolatile Full Adder Based on Logic-in-Memory Architecture Using Magnetic Tunnel Junctions", *Applied Physics Express*, vol. 1, pp. 091301, 2008.

[135] J. Borghetti, G. S. Snider, P. J. Kuekes, J. J. Yang, D. R. Stewart and R. S. Williams, "'Memristive' switches enable 'stateful' logic operations via material implication", *Nature*, vol. 464, pp. 873-876, 2010.

[136] Y. Gang, W. Zhao, J.-O. Klein, C. Chappert and P. Mazoyer, "A High-Reliability, Low-Power Magnetic Full Adder", *IEEE Transactions on Magnetics*, vol. 47, pp. 4611-4616, 2011.

[137] B. Behin-Aein, D. Datta, S. Salahuddin and S. Datta, "Proposal for an all-spin logic device with built-in memory", *Nat Nano*, vol. 5, pp. 266-270, 2010.

[138] C. Mead, "Neuromorphic electronic systems", *Proceedings of the IEEE*, vol. 78, pp. 1629-1636, 1990.

[139] X. Wang and Y. Chen, "Spintronic memristor devices and application", *proceeding in DATE 2010*, pp. 667-672, 2010.

[140] M. Sharad, C. Augustine, G. Panagopoulos and K. Roy, "Spin-Based Neuron Model With Domain-Wall Magnets as Synapse", *IEEE Transactions on Nanotechnology*, vol. 11, pp. 843-853, 2012.

[141] W. F. Brinkman, R. C. Dynes and J. M. Rowell, "Tunneling Conductance of Asymmetrical Barriers", *Journal of Applied Physics*, vol. 41, pp. 1915-1921, 1970.

[142] W. Zhao, J. Duval, J.-O. Klein, et C. Chappert, "A compact model for magnetic tunnel junction (MTJ) switched by thermally assisted Spin transfer torque (TAS + STT)", *Nanoscale Research Letters*, vol. 6, pp. 368, 2011.

[143] K. Nepal, R. I. Bahar, J. Mundy, W. R. Patterson and A. Zaslavsky, "Designing MRF based Error Correcting Circuits for Memory Elements", *proceeding in DATE 2006*, pp. 792-793, 2006.

[144] S. Zhang, P. M. Levy, A. C. Marley and S. S. P. Parkin, "Quenching of Magnetoresistance by Hot Electrons in Magnetic Tunnel Junctions", *Phys. Rev. Lett.*, vol. 79, pp. 3744-3747, 1997.

[145] J. Z. Sun, "Spin-current interaction with a monodomain magnetic body: A model study", *Phys. Rev. B*, vol. 62, pp. 570-578, 2000.

[146] J. Xiao, A. Zangwill and M. D. Stiles, "Macrospin models of spin transfer dynamics", *Phys. Rev. B*, vol. 72, pp. 014446, 2005.

[147] T. L. Gilbert, "A phenomenological theory of damping in ferromagnetic materials", *IEEE Transactions on Magnetics*, vol. 40, pp. 3443-3449, 2004.

[148] G. D. Fuchs, I. N. Krivorotov, P. M. Braganca, N. C. Emley, A. G. F. Garcia, D. C. Ralph, and R. A. Buhrman, "Adjustable spin torque in magnetic tunnel junctions with two fixed layers", *Applied Physics Letters*, vol. 86, pp. 152509, 2005.

[149] J. C. Slonczewski, "Currents, torques, and polarization factors in magnetic tunnel junctions", *Phys. Rev. B*, vol. 71, pp. 024411, 2005.

[150] R. H. Koch, J. A. Katine and J. Z. Sun, "Time-Resolved Reversal of Spin-Transfer Switching in a Nanomagnet", *Phys. Rev. Lett.*, vol. 92, pp. 088302, 2004.

[151] R. Heindl, W. H. Rippard, S. E. Russek, M. R. Pufall and A. B. Kos, "Validity of the thermal activation model for spin-transfer torque switching in magnetic tunnel junctions", *Journal of Applied Physics*, vol. 109, pp. 073910, 2011.

[152] L.-B. Faber, W. Zhao, J.-O. Klein, T. Devolder and C. Chappert, "Dynamic compact model of Spin-Transfer Torque based Magnetic Tunnel Junction (MTJ)", *proceeding in DTIS 2009*, pp. 130-135, 2009.

[153] T. Devolder, J. Hayakawa, K. Ito, H. Takahashi, S. Ikeda, P. Crozat, N. Zerounian, J.-V. Kim, C. Chappert and H. Ohno, "Single-Shot Time-Resolved Measurements of Nanosecond-Scale Spin-Transfer Induced Switching: Stochastic Versus Deterministic Aspects", *Phys. Rev. Lett.*, vol. 100, pp. 057206, 2008.

[154] J. J. Nowak, R. P. Robertazzi, J. Z. Sun, G. Hu, D. W. Abraham, P. L. Trouilloud, S. Brown, M. C. Gaidis, E. J. O'Sullivan, W. J. Gallagher and D. C. Worledge, "Demonstration of Ultralow Bit Error Rates for Spin-Torque Magnetic Random-Access Memory With Perpendicular Magnetic Anisotropy", *IEEE Magnetics Letters*, vol. 2, pp. 3000204-3000204, 2011.

[155] Z. Wang, Y. Zhou, J. Zhang and Y. Huai, "Bit error rate investigation of spin-transfer-switched magnetic tunnel junctions", *Applied Physics Letters*, vol. 101, pp. 142406, 2012.

[156] H. Tomita, S. Miwa, T. Nozaki, S. Yamashita, T. Nagase, K. Nishiyama, E. Kitagawa, M. Yoshikawa, T. Daibou, M. Nagamine, T. Kishi, S. Ikegawa, N. Shimomura, H. Yoda and Y. Suzuki, "Unified understanding of both thermally assisted and precessional spin-transfer

switching in perpendicularly magnetized giant magnetoresistive nanopillars", *Applied Physics Letters*, vol. 102, pp. 042409, 2013.

[157] K. Lee and S. H. Kang, "Development of Embedded STT-MRAM for Mobile System-on-Chips", *IEEE Transactions on Magnetics*, vol. 47, pp. 131-136, 2011.

[158] Y. Kim, S. C. Oh, W. C. Lim, J. H. Kim, W. J. Kim, J. H. Jeong, H. J. Shin, K. W. Kim, K. S. Kim, J. H. Park, S. H. Park, H. Kwon, K. H. Ah, Y. Lee, S. O. Park, S. Choi, H.-K. Kang, and C. Chung, "Integration of 28nm MJT for 8~16Gb level MRAM with full investigation of thermal stability", *proceeding in 2011 Symposium on VLSI Technology DATE*, pp. 210-211, 2011.

[159] G. S. D. Beach, M. Tsoi and J. L. Erskine, "Current-induced domain wall motion", *Journal of Magnetism and Magnetic Materials*, vol. 320, pp. 1272-1281, 2008.

[160] A. Thiaville, Y. Nakatani, J. Miltat and Y. Suzuki, "Micromagnetic understanding of current-driven domain wall motion in patterned nanowires", *EPL*, vol. 69, pp. 990, 2005.

[161] H. Tanigawa, T. Koyama, G. Yamada, D. Chiba, S. Kasai, S. Fukami, T. Suzuki, N. Ohshima, N. Ishiwata, Y. Nakatani and T. Ono, "Domain Wall Motion Induced by Electric Current in a Perpendicularly Magnetized Co/Ni Nano-Wire", *Applied Physics Express*, vol. 2, pp. 053002, 2009.

[162] X. Jiang, L. Thomas, R. Moriya, M. Hayashi, B. Bergman, C. Rettner and S. S. P. Parkin, "Enhanced stochasticity of domain wall motion in magnetic racetracks due to dynamic pinning", *Nat Commun*, vol. 1, pp. 25, 2010.

[163] Verilog-A Manual, Agilent Technol., Palo Alto, CA, 2007.

[164] Virtuoso Spectre circuit simulator datasheet, Cadence.

[165] Eldo custom Design & Simulation simulator datasheet, Mentor-Graphics.

[166] Advanced Design System (ADS) simulator datasheet, Agilent.

[167] W. S. Zhao, T. Devolder, Y. Lakys, J. O. Klein, C. Chappert and P. Mazoyer, "Design considerations and strategies for high-reliable STT-MRAM", *Microelectronics Reliability*, vol. 51, pp. 1454-1458, 2011.

[168] CMOS065 Design Rule Manual, STMicroelectronics, Geneva, Switzerland, 2010.

[169] H. Meng, J. Wang and J.-P. Wang, "A spintronics full adder for magnetic CPU", *IEEE Electron Device Letters*, vol. 26, pp. 360-362, 2005.

[170] W. Kim, J. H. Jeong, Y. Kim, W. C. Lim, J.-H. Kim, J. H. Park, H. J. Shin, Y. S. Park, K. S. Kim, S. H. Park, Y. J. Lee, K. W. Kim, H. J. Kwon, H. L. Park, H. S. Ahn, S. C. Oh, J. E. Lee,

S. O. Park, S. Choi, H.-K. Kang and C. Chung, “Extended scalability of perpendicular STT-MRAM towards sub-20nm MTJ node”, *proceeding in IEDM 2011*, pp. 24.1.1-24.1.4, 2011.

[171] H. Yoda, S. Fujita, N. Shimomura, E. Kitagawa, K. Abe, K. Nomura, H. Noguchi and J. Ito, “Progress of STT-MRAM technology and the effect on normally-off computing systems”, *proceeding in IEDM 2012*, pp. 11.3.1-11.3.4, 2012.

[172] D. A. Patterson and J. L. Hennessy, “Computer Organization and Design: The Hardware/software Interface”, *Elsevier*, 2012.

[173] F. Ren and D. Markovic, “True Energy-Performance Analysis of the MTJ-Based Logic-in-Memory Architecture (1-Bit Full Adder)”, IEEE Transactions on Electron Devices, vol. 57, pp. 1023-1028, 2010.

[174] K. Pagiamtzis and A. Sheikholeslami, “Content-addressable memory (CAM) circuits and architectures: a tutorial and survey”, *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 712-727, 2006.

[175] R. Nebashi, N. Sakimura, Y. Tsuji, S. Fukami, H. Honjo, S. Saito, S. Miura, N. Ishiwata, K. Kinoshita, T. Hanyu, T. Endoh, N. Kasai, H. Ohno and T. Sugibayashi, “A content addressable memory using magnetic domain wall motion cells”, *proceeding in VLSIC 2011*, pp. 300-301, 2011.

[176] W. Zhao, D. Ravelosona, J. Klein and C. Chappert, “Domain Wall Shift Register-Based Reconfigurable Logic”, *IEEE Transactions on Magnetics*, vol. 47, pp. 2966-2969, 2011.

[177] S. Matsunaga, A. Katsumata, M. Natsui, S. Fukami, T. Endoh, H. Ohno and T. Hanyu, “Fully parallel 6T-2MTJ nonvolatile TCAM with single-transistor-based self match-line discharge control”, *proceeding in VLSIC 2011*, pp. 298-299, 2011.

[178] H. Kadota, J. Miyake, Y. Nishimichi, H. Kudoh and K. Kagawa, “An 8-kbit content-addressable and reentrant memory”, *IEEE Journal of Solid-State Circuits*, vol. 20, pp. 951-957, 1985.

[179] W. Wu, X. Zhu, S. Kang, K. Yuen and R. Gilmore, “Probabilistically Programmed STT-MRAM”, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, pp. 42-51, 2012.

[180] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder and W. Lu, “Nanoscale Memristor Device as Synapse in Neuromorphic Systems”, *Nano Lett.*, vol. 10, pp. 1297-1301, 2010.

[181] D. Querlioz, O. Bichler, P. Dollfus and C. Gamrat, "Immunity to Device Variations in a Spiking Neural Network With Memristive Nanodevices", *IEEE Transactions on Nanotechnology*, vol. 12, pp. 288-295, 2013.

[182] T. Koyama, K. Ueda, K.-J. Kim, Y. Yoshimura, D. Chiba, K. Yamada, J.-P. Jamet, A. Mougin, A. Thiaville, S. Mizukami, S. Fukami, N. Ishiwata, Y. Nakatani, H. Kohno, K. Kobayashi and T. Ono, "Current-induced magnetic domain wall motion below intrinsic threshold triggered by Walker breakdown", *Nat Nano*, vol. 7, pp. 635-639, 2012.

[183] G. Zhang, W. S. Hwang, S.-H. Lee, B.-J. Cho and W. J. Yoo, "Endurance Reliability of Multilevel-Cell Flash Memory Using a ZrO2/Si3N4 Dual Charge Storage Layer", *IEEE Transactions on Electron Devices*, vol. 55, pp. 2361-2369, 2008.

[184] A. Beck, J. G. Bednorz, C. Gerber, C. Rossel and D. Widmer, "Reproducible switching effect in thin oxide films for memory applications", *Applied Physics Letters*, vol. 77, pp. 139-141, 2000.

[185] K. Kinoshita, K. Tsunoda, Y. Sato, H. Noshiro, S. Yagaki, M. Aoki and Y. Sugiyama, "Reduction in the reset current in a resistive random access memory consisting of NiOx brought about by reducing a parasitic capacitance", *Applied Physics Letters*, vol. 93, pp. 033506, 2008.

[186] X. Lou, Z. Gao, D. V. Dimitrov and M. X. Tang, "Demonstration of multilevel cell spin transfer switching in MgO magnetic tunnel junctions", *Applied Physics Letters*, vol. 93, pp. 242502, 2008.

[187] B. Razavi, "Design of Analog CMOS Integrated Circuits". *Tata McGraw-Hill*, 2002.

[188] J. Sun, S. Brown, W. Chen, and E. Delenia, "Spin-torque switching efficiency in CoFeB-MgO based tunnel junctions," Phys. Rev. B, vol. 104426, pp. 1–9, 2013.

List of Publications

Articles published in referred pubilications (Journal)

[1] **Y. Zhang**, W.S. Zhao, D. Ravelosona, J.-O. Klein, J.-V. Kim and C. Chappert, "Compact Modeling of Perpenticular-Anistropy CoFeB/MgO Magnetic Tunnel Junction", *IEEE Transaction on Electron Device*, vol. 59, pp. 819-826, 2012.

[2] **Y. Zhang**, W.S. Zhao, D. Ravelosona, J.-O. Klein, J.-V. Kim and C. Chappert, "Perpendicular-Magnetic-Anisotropy CoFeB Racetrack Memory", *Journal of Applied Physics*, vol. 111, pp. 093925, 2012.

[3] **Y. Zhang**, W.S. Zhao, J.-O. Klein, D. Ravelosona and C. Chappert, "Ultra-High Density Content Addressable Memory based on Current Induced Domain Wall Motion in Magnetic Track", *IEEE Transaction on Magnetics*, vol. 48, pp. 3219-3222, 2012.

[4] **Y. Zhang**, W.S. Zhao, G. Prenat, T. Devolder, J.-O. Klein, C. Chappert, B. Dieny and D. Ravelosona, "Electrical Modeling of Stochastic Spin Transfer Torque Writing in Magnetic Tunnel Junctions for Memory and Logic Applications", *IEEE Transaction on Magnetics*, vol. 49, pp. 4375-4378, 2013.

[5] **Y. Zhang**, W.S. Zhao, J.-O. Klein, C. Chappert and D. Ravelosona, "Current Induced Perpendicular-Magnetic-Anisotropy Racetrack Memory with Magnetic Field Assistance", *Applied Physics Letter*, vol. 104, pp. 032409, 2014.

[6] **Y. Zhang**, W.S. Zhao, J.-O. Klein, C. Chappert and D. Ravelosona, "Implementation of Magnetic Field Assistance to Current-Induced Perpendicular-Magnetic-Anisotropy Racetrack Memory", *Journal of Applied Physics*, vol. 115, pp. 17D509, 2014.

[7] E. Deng, **Y. Zhang**, J.-O. Klein, C. Chappert, D. Ravelosona and W.S. Zhao, "Low Power Magnetic Full-Adder based on Spin Transfer Torque MRAM", *IEEE Transaction on Magnetics*, vol. 49, pp. 4982-4987, 2013.

[8] W.S. Zhao, **Y. Zhang**, T. Devolder, J.-O. Klein, D. Ravelosona, C. Chappert and P. Mazoyer, "Failure and Reliability Analysis of STT-MRAM", *Microelectronics Reliability*, vol. 52, pp. 1848-1852, 2012.

[9] Z.H. Wang, W.S. Zhao, W. Kang, **Y. Zhang**, J.-O. Klein, D. Ravelosona and C. Chappert, "Compact modelling of ferroelectric tunnel memristor and its use for neuromorphic simulation", *Applied Physics Letter*, vol. 104, pp. 053505, 2014.

[10] D. Chabi, W.S. Zhao, E. Deng, **Y. Zhang**, N. Ben Romdhane, J.-O. Klein and C. Chappert, "Ultra Low Power Magnetic Flip-Flop Based on checkpointing/Power Gatingand Self-Enable Mechanisms", *Accepted* by IEEE Transaction on Circuits and Systems I, to appear in 2014.

[11] W.S. Zhao, M. Moreau, E. Deng, **Y. Zhang**, J.-M. Portal, J.-O. Klein, M. Bocquet, H. Aziza, D. Deleruyelle, C. Muller, D. Querlioz, N. Ben Romdhane, C. Chappert and D. Ravelosona, "Synchronous Non-Volatile Logic Gate Design based on Resistive Switching Memories", *IEEE Transaction on Circuits and Systems I*, vol. 61, pp. 443-454, 2014.

[12] H.-P. Trinh, W.S. Zhao, J.-O. Klein, **Y. Zhang**, D. Ravelosona and C. Chappert, "Domain wall motion based Magnetic Adder", *Electronics Letters*, vol. 48, pp. 1049-1051, 2012.

[13] Y. Lakys, W.S. Zhao, T. Devolder, **Y. Zhang**, J.-O. Klein, D. Ravelosona and C. Chappert, "Self-Enabled "Error-Free" Switching Circuit for Spin Transfer Torque MRAM and Logic", *IEEE Transaction on Magnetics*, vol. 48, pp. 2403-2406, 2012.

[14] H.-P. Trinh, W.S. Zhao, J.-O. Klein, **Y. Zhang**, D. Ravelosona and C. Chappert, "Magnetic Adder based on racetrack memory", *IEEE Transactions on Circuits and Systems I*, vol. 60, pp. 1469-1477, 2012.

[15] W. Kang, E. Deng, J.-O. Klein, **Y. Zhang**, Y.G. Zhang, C. Chappert, D. Ravelosona and W.S. Zhao, "Separated Pre-Charge Sensing Amplifier for Deep Submicron MTJ/CMOS Hybrid Logic Circuits", *Accepted by IEEE Transaction on Magnetics*, to appear in 2014.

[16] W. Kang, W.S. Zhao, Z.H. Wang, **Y. Zhang**, J.-O. Klein, Y.G. Zhang, C. Chappert and D. Ravelosona, "DFSTT-MRAM: Dual Functional STT-MRAM Cell Structure for Reliability Enhancement and 3D MLC Functionality", *Accepted by IEEE Transaction on Magnetics*, to appear in 2014.

[17] W. Kang, W.S. Zhao, Z.H. Wang, **Y. Zhang**, J.-O. Klein, Y.G. Zhang, C. Chappert and D. Ravelosona, "A low-cost built-in error correction circuit design for STT-MRAM reliability improvement", *Microelectronics Reliability*, vol. 53, pp. 1224-1229, 2013.

[18] M. Bocquet, H. Aziza, W.S. Zhao, **Y. Zhang**, S. Onkaraiah, C. Muller, M. Reyboz, D. Deleruyelle, F. Clermidy and J.-M. Portal, "Compact modeling solutions for OxRAM memories", *Journal of Low Power Electronics and Applications*, vol. 4, pp. 1-14, 2014.

[19] W.S. Zhao, J.-M. Portal, W. Kang, M. Moreau, **Y. Zhang**, H. Aziza, J.-O. Klein, Z.H. Wang, D. Querlioz, D. Deleruyelle, M. Bocquet, D. Ravelosona, C. Muller and C. Chappert, "Design and analysis of Crossbar Architecture based on Complementary Resistive Switching non-volatile Memory Cells", *Accepted by Journal of Parallel and Distributed Computing*, to appear in 2014.

[20] Z.H. Wang, W.S. Zhao, W. Kang, A. Bouchenak-Khelladi, **Y. Zhang**, J.-O. Klein, D. Ravelosona and C. Chappert, "Write operation study of Co/BTO/LSMO ferroelectric tunnel junction", *Journal of Applied Physics*, vol. 114, pp. 044108, 2013.

[21] Z.H. Wang, W.S. Zhao, W. Kang, A. Bouchenak-Khelladi, **Y. Zhang**, J.-O. Klein, D. Ravelosona and C. Chappert, "A physics-based compact model of ferroelectric tunnel junction for memory and logic design", *Journal of Physics D: Applied Physics*, vol. 47, pp. 045001, 2014.

Articles published in referred international conference proceedings

[1] **Y. Zhang**, W.S. Zhao, J.-O. Klein, W. Kang, D. Querlioz, Y.G. Zhang, D. Ravelosona and C. Chappert, "Spintronics for low power computing", to be published in the Procs. Of IEEE DATE, 2014. (invited paper)

[2] **Y. Zhang**, E. Deng, J.-O. Klein, D. Querlioz, D. Ravelosona, C. Chappert, M. Moreau, J.-M. Portal, M. Bocquet, H. Aziza, D. Deleruyelle, C. Muller, W.S. Zhao, "Synchronous Full-Adder Based on Complementary Resistive Switching Memory Cells", *proceeding in IEEE NEWCAS 2013*, pp. 1-4, 2013.

[3] **Y. Zhang**, W.S. Zhao, J.-O. Klein, W. Kang, D. Querlioz, C. Chappert, D. Ravelosona "Multi-Level Cell Spin Transfer Torque MRAM Based on Stochastic Switching", *proceeding in IEEE NANO 2013*, pp. 233-236, 2013.

[4] W.S. Zhao, **Y. Zhang**, J.-O. Klein, D. Querlioz, D. Chabi, D. Ravelosona, C. Chappert, J.-M. Portal, M. Bocquet, H. Aziza, D. Deleruyelle, C. Muller, "Crossbar Architecture based on 2R Complementary Resistive Switching Memory Cell", *proceeding in ACM/IEEE Nanoarch 2012*, pp. 85-92, 2012.

[5] W.S. Zhao, **Y. Zhang**, H.-P. Trinh, J.-O. Klein, C. Chappert, R. Mantovan, A. Lamperti, R.P. Cowburn, T. Trypiniotis, M. Klaui, J. Heinen, B. Ocker and D.Ravelosona, "Magnetic Domain-Wall Racetrack Memory for high density and fast data Storage", *proceeding in IEEE ICSICT 2012*, pp. 1-4, 2012.

[6] J.-M. Portal, M. Moreau, **Y. Zhang**, E. Deng, J.-O. Klein, D. Querlioz, D. Ravelosona, C. Chappert, M. Bocquet, H. Aziza, D. Deleruyelle, C. Muller, W. Zhao, "Analytical Study of Complementary Memristive Synchronous Logic Gates", *proceeding in ACM/IEEE Nanoarch 2013*, pp. 70-75, 2013.

[7] W.S. Zhao, N. Ben Romdhane, **Y. Zhang**, J.-O. Klein and D. Ravelosona, "Racetrack Memory Based Reconfigurable Computing", *proceeding in IEEE Faible Tension Faible Consommation (FTFC) 2013*, pp. 1-4, 2013. (invited paper)

[8] W.S. Zhao, J.-O. Klein, Z.H. Wang, **Y. Zhang**, N. Ben Romdhane, D. Querlioz, D. Ravelosona and C. Chappert, "Spin-electronics Based Logic Fabrics", *proceeding in IEEE VLSI-SOC 2013*, pp. 174-179, 2013. (invited paper)

[9] Z.H. Wang, W.S. Zhao, A. Bouchenak-Khelladi, **Y. Zhang**, W.W. Lin, J.-O. Klein, D. Ravelosona and C. Chappert, "Compact Modelling for Co/BTO/ LSMO Ferroelectric Tunnel Junction", *proceeding in IEEE NANO 2013*, pp. 236-239, 2013.

[10] D. Chabi, W.S. Zhao, J.-O. Klein, **Y. Zhang** and C. Chappert, "Low Power Magnetic Flip-Flop Based on checkpointing and Self-Enable Mechanisms", *proceeding in IEEE NEWCAS 2013*, pp. 1-4, 2013.

[11] W. Kang, W.S. Zhao, J.-O. Klein, Z.H. Wang**, Y. Zhang**, Y.G. Zhang, D. Ravelosona and C. Chappert, "An Overview of Spin-based Integrated Circuits", *proceeding in IEEE ASP-DAC 2014*, pp. 676-683, 2014. (invited paper)

[12] W.S. Zhao, L.Torres, Y. Guillemenet, L. Vitorio Cargnini, **Y. Zhang**, Y. Lakys, J.-O. Klein, D. Ravelosona, G. Sassatelli and C. Chappert, "Embedded MRAM for High-speed Computing", *proceeding in IEEE VLSI-SOC 2011*, , pp. 37-42, 2011.

Book chapter

W.S. Zhao, L. Torres, Y. Guillemenet, L. Vitorio Cargnini, **Y. Zhang**, Y. Lakys, J.-O. Klein, D. Ravelosona, G. Sassatelli and C. Chappert, "High Performance SoC Design Using Magnetic Logic and Memory", *VLSI-SoC: Advanced Research for Systems on Chip, IFIP Advances in Information and Communication Technology*, by Springer, vol. 379, pp 10-33, 2012.

Appendix A List of Acronyms

1D	One Dimension
2D	Two Dimensions
3D	Three Dimensions
AFM	Anti-Ferromagnetic
AP	Anti-Parallel
ASL	All Spin Logic
BER	Bit Error Rate
BL	Bit Line
CAD	Computer Aided Design
CAM	Content Addressable Memory
CBRAM	Conductive Bridging Random Access Memory
CDF	Component Description Format
CIDW	Current Induced Domain Wall
CIP	Current In Plane
CMOS	Complementary Metal-Oxide-Semiconductor
CPP	Current Perpendicular to Plane
CPU	Central Processing Unit
DC	Direct Current
DMI	Dzyaloshinskii-Moriya Interaction
DRAM	Dynamic Random Access Memory
DW	Domain Wall
ECC	Error Correction Circuit
EDP	Energy-Delay Product
FET	Field Effect Transistor

FIMS	Field Induced Magnetic Switching
FM	Ferromagnetic
FPGA	Field Programmable Gate Array
GMR	Giant MagnetoResistance
HDD	Hard Disk Drive
ITRS	International Technology Roadmap for Semiconductor
LLG	Landau-Lifshitz-Gilbert
LUT	Look Up Table
MFA	Magnetic Full Adder
ML	Match Line
MLC	Multi-Level Cell
MOSFET	Metal-Oxide-Semiconductor Field Effect Transistor
MRAM	Magnetoreisistance Random Access Memory
MTJ	Magnetic Tunnel Junction
NM	Non-Magnetic
NML	Nanomagnetic Logic
OxRAM	Oxide Random Access Memory
P	Parallel
PCRAM	Phase-Change Random Access Memory
PCSA	Pre-Charge Sense Amplifier
PDF	Probability Density Function
PMA	Perpendicular Magnetic Anisotropy
RA	Resistance-Area Product

RAM	Random Access Memory
RM	Racetrack Memory
RRAM	Resistive Random Access Memory
SHE	Spin Hall Effect
SL	Source Line
SLC	Single-Level Cell
SOT	Spin Orbit Torque
SRAM	Static Random Access Memory
STT	Spin Transfer Torque
TAS	Thermally Assisted Switching
TMR	Tunnel MagnetoResistance
WL	Word Line

Appendix B Source code of PMA STT MTJ compact model

```
`resetall
`include "constants.vams"
`include "disciplines.vams"
`define explimit 85.0
`define exp(x) exp(min(max((x),-`explimit),`explimit))
`define sqrt(x) pow( (x), 0.5)

//Shape definition
`define rec 1
`define ellip 2
`define circle 3

/*-----------------------------------------
Electrical Constants
-----------------------------------------*/

/*----------Elementary Charge---------------*/
`define e 1.6e-19
/*----------Bohr Magnetron Costant----------*/
`define ub 9.27e-28
/*----------Boltzmann Constant------------- */
`define Kb 1.38e-23
/*----------Electron Mass------------- */
`define m 9.10e-31
/*----------Euler's constant---------------*/
`define C 0.577

module Model1(T1,T2,x,Ttrans,PI);

inout T1, T2, x, Ttrans, PI;
electrical T1, T2, x, Ttrans, PI;

/*-----------------------------------------
MTJ Technology Parameters
-----------------------------------------*/

/*----------Gilbert Damping Coefficient---------------*/
parameter real  alpha=0.027;
/*----------GyroMagnetic Constant in Hz/Oe---------------*/
parameter real  gamma=1.76e7;
/*----------Electron Polarization Percentage % ---------------*/
parameter real  P=0.52;
/*----------Out of plane Magnetic Anisotropy in Oersteds------------*/
parameter real  Hk=1433;
/*--------Saturation Field in the Free Layer in Oersteds---------*/
parameter real  Ms=15800;
```

```
/*-----The Energy Barrier Height for MgO in electron-volt----------*/
parameter real  PhiBas=0.4;
/*------Voltage bias when the TMR(real) is 1/2TMR(0) in Volt---*/
parameter real  Vh=0.5;
/*----------Current Pulse width in second---------------*/
parameter real  Pwidth=10e-3;


/*----------------------------------------
Device Parameters
----------------------------------------*/
/*----------Height of the Free Layer in nm---------------*/
parameter real  tsl=1.3e-9 from[0.7e-9:3.0e-9];
/*----------Length in nm---------------*/
parameter real  a=40e-9;
/*----------Width in nm---------------*/
parameter real  b=40e-9;
/*----------Radius in nm---------------*/
parameter real  r=20e-9;
/*----------Height of the Oxide Barrier in nm---------------*/
parameter real  tox=8.5e-10 from[8e-10:15e-10];
/*----------TMR(0) with Zero Volt Bias Voltage ---------------*/
parameter real  TMR=0.7;
/*----------Shape of MTJ---------------*/
parameter real   SHAPE=2  from[1:3];  //SQUARE
/*----------Neel-Brown model parameter ---------------*/
parameter real   tau0=8.7e-10;
/*----------Error probability Ps=1-Pr(t) ----------------*/
parameter real   Ps=0.999999;
/*----------Threshold for Neel-Brown model----------------*/
parameter real brown_threshold=0.0;

/*----------MTJ State Parameters----------------*/
/*---Initial state of the MTJ, 0 = parallele, 1 = anti-parallele---*/
parameter integer     PAP=0 from[0:1];
/*----------Room temperature in Kelvin----------------*/
parameter real        T= 300;//$random % 50 +323;
/*----------Resistance area product in ohmum2----------------*/
parameter real        RA=5 from[5:15];

//variables

//Polaristion constant for the two states of STT-MTJ
real PolaP;     //Polarization state parallel of STT-MTJ
real PolaAP;    //Polarization state anti-parallel of STT-MTJ

real surface;   //Surface of MTJ

//Critical current density for the two states of STT-MTJ
real gp;   //Critical current density for P state
real gap;  //Critical current density for AP state
```

```
real Em,EE;      //Variable of the Slonczewski model

//TMR real value for the two states of STT-MTJ
real TMRR;       //TMR real value for P state
real TMRRT;      //TMR real value for AP state

//Resistance of MTJ
real Ro;   //Resistance of MTJ when bias voltage = 0V
real Rap;  //Resistance value for AP state
real Rp;   //Resistance value for P state

//Voltage of MTJ
real Vb;   //V(T1,T2)
real Vc;   //V(T2,T1)

real Id;   //Current of MTJ

//critial current for the two states of STT-MTJ
real IcAP;       //Critial current for AP state
real IcP;  //Critial current for P state

integer ix;      //Current used to store the state of the MTJ

real tau;  //Probability parameter

real FA; //Factor for calculating the resistance based on RA


real durationstatic,duration;    //time needed to be sure that the
switching is effected

analog begin

      if (SHAPE==1)
      begin
            surface=a*b;     //SQUARE
      end
      else if (SHAPE==2)
      begin
            surface=`M_PI*a*b/4;  //ELLIPSE
      end
      else
      begin
            surface=r*r;     //ROUND
      end

      Vc=V(T2,T1);     //potential between T2 and T1
      Vb=V(T1,T2);     //potential between T2 and T1

//initial conditions
@(initial_step)
begin
```

```
      FA=3322.53/RA;   //initialization of resistance factor according
to RA product

      Ro=(toxreal*1.0e10/(FA*`sqrt(PhiBas)*surface*1.0e12))*exp(1.025*t
ox*1.0e10*`sqrt(PhiBas));   //resistance

      Em=Ms*tsl*surface*Hk/2;
      EE=Em/(`Kb*T*40*`M_PI);  //result of E/kbT

      duration=1.0;

      //States inititialisation
      ix=-PAP;
      I(Ttrans)<+PAP;

      if(I(Ttrans)==0) //Case which the magnetizations of the two layers
are parallel
      begin
            //TMR real
            TMRR=TMRreal/(1+Vb*Vb/(Vh*Vh));

            //Parallel resistance
            Rp=Ro;
            Id=Vb/Rp;

            PolaP=`sqrt(TMRreal*(TMRreal+2))/(2*(TMRreal+1));
            //Polarization state parallel

            gp=alpha*gamma*`e*Ms*tslreal*Hk/(40*`M_PI*(`ub*PolaP));
            //Critical current density


            IcP=gp*surface;       // Critical current

            if(Vb>=IcP*Rp)
            begin //Current higher than critical current

            //STT-MTJ dynamic behavior : Sun model
      duration=(`C+ln(`M_PI*`M_PI*(Em/(`Kb*T*40*`M_PI))/4))*`e*1000*Ms*
surface*tslreal*(1+P*P)/(4*`M_PI*2*`ub*P*10000*abs(Id-IcP));

                  if(duration<Pwidth)
                  begin //Switching occurs

                        ix=-1.0;    //state changes

                  end
                  else
                  begin
                        ix=0.0;     //state maintains
                  end
```

```
            end
            else
            begin //Current smaller than critical current

                  ix=0.0;     //state maintains
                  //STT-MTJ dynamic behavior : Neel-Brown model
                  tau=tau0*exp(Em*(1-abs(Id/IcP))/(`Kb*T*40*`M_PI));
                  Id=Vb/Rp;

                  if(Vb>brown_threshold)
                  begin
                        if(Vb<0.8*IcP*Rp)
                        begin
                              duration=tau;
                              if(Pwidth > duration)
                              begin
                                    ix=-1.0;    //state changes

                              end
                              else
                              begin
                                    ix=0.0;  //state maintains

                              end
                        end
                  end
            end
      end

      if(I(Ttrans)!=0)      //Case which the magnetizations of the two
layers are antiparallel
      begin
            //TMR real
            TMRRT=TMRreal/(1+Vb*Vb/(Vh*Vh));

            // Antiparallel resistance
            Rap=Ro*(1+TMRRT);
            Rp=Ro;
            Id=Vb/(Rap);

            PolaAP=`sqrt(TMRreal*(TMRreal+2))/(2*(TMRreal+1));
            //Polarization state anti parallel

            gap=alpha*gamma*`e*Ms*tslreal*Hk/(40*`M_PI*(`ub*PolaAP));
            //Critical current density

            IcAP=gap*surface;     // Critical current

            if(Vc>=(IcAP*Rap))
            begin //Current higher than critical current

            //STT-MTJ dynamic behavior : Sun model
```

```
      duration=(`C+ln(`M_PI*`M_PI*(Em/(`Kb*T*40*`M_PI))/4))*`e*1000*Ms*
surface*tslreal*(1+P*P)/(4*`M_PI*2*`ub*P*10000*abs(-Id-IcAP));

                  if((duration<Pwidth))
                  begin //Switching occurs

                        ix=0.0;    //state changes

                  end
                  else
                  begin
                        ix=-1.0;    //state maintains
                  end
            end
            else
            begin //Current smaller than critical current

                  ix=-1.0;    //state maintains
                  //STT-MTJ dynamic behavior : Neel-Brown model
                  tau=tau0*exp(Em*(1-abs(Id/IcAP))/(`Kb*T*40*`M_PI));
                  Id=Vb/Rp;

                  if(Vc>brown_threshold)
                  begin
                        if(Vc<0.8*IcAP*Rap)
                        begin
                              duration=tau;
                              if(duration<Pwidth)
                              begin
                                    ix=0.0;    //state changes
                              end
                              else
                              begin
                                    ix=-1.0;

                              end
                        end
                  end
            end
      end

I(x)<+ix;  //Actualisation of the state of x with the value calculated


//Ttrans has the same function than x but it includes the time effects
V(Ttrans)<+transition(-ix,duration,1e-12,1e-12);

I(T1,T2)<+Id;    //Actualisation of the current of MTJ with the value
calculated

end
```

```
end

endmodule
```

Appendix C Source code of PMA racetrack memory compact model

```
`resetall
`include "constants.vams"
`include "disciplines.vams"
`define explimit 85.0
`define exp(x) exp(min(max((x),-`explimit),`explimit))
`define sqrt(x) pow( (x), 0.5)

//Shape definition
`define rec 1
`define ellip 2
`define circle 3

module model_layer_Domain_Wall(x,T1, T2, NextState, Ttrans);

inout T1, T2, NextState, x, Ttrans;
electrical T1, T2, NextState, x, Ttrans;

/*----------------LAYER PARAMETERS ----------*/

/*----------Length in nm---------------*/
parameter real  c=1.0e-6;
/*----------Width in nm---------------*/
parameter real  b=65.0e-9;
/*----------Thickness of the Oxide Barrier in nm---------------*/
parameter real  tox=8.5e-10 from[8e-10:15e-10];
/*----------Thickness of the Free Layer  in nm---------------*/
parameter real  thick_f = 1.3e-9;

/*----------MTJ Initial State ----------------*/
parameter real PAP=0 from[0:1]; //Initial state of the MTJ, 0 =
parallele, 1 = anti-parallele

/*----------Shape of MTJ---------------*/
parameter real SHAPE=2  from[1:2];

/*----------Critical current density to propagate DWs in the free layer
in A/(m*m)---------------*/
parameter real  Jc0=0.62e12;

/*---------Resistivity ----------------------------*/
parameter real rau=1e-7;

/*---------Polarization rate----------------------------*/
parameter real P=0.72;

/*-------------Factor for velocity in m*m*m/C---------*/
```

```
parameter real Fa=3e-11;


/*-----------------LAYER VARIABLES---------------------------*/
real Vwrite;          //V(T1,T2)
real Vwriteinv;       //V(T1,T2)
real Vnextstate;      //V(nextstate)
real VTtrans;         //V(Ttrans)
real Vx;              //V(x)

real Ishift;     //Current of the layer

real state_layer;     //State of the free layer of the MTJ ;
0=Parallel ; 1=Antiparallel

real duration_shift;  //Time to shift in second

real Ic0;  //Critical current to propagate DWs in the free layer

real surface_layer_shift;
real vol_layer_shift;
real res_layer_shift;
real speed_shifting;
real J;

analog begin

//initial conditions
@(initial_step)
     begin

          state_layer=PAP;
          Ishift=0;
          Vwrite=0;
          Vwriteinv=0;
          Vnextstate=0;
          VTtrans=0;
          Vx=0;

     end

     Vwrite=V(T1,T2);
     Vwriteinv=V(T2,T1);

     Vnextstate=V(NextState);

     VTtrans=V(Ttrans);
     Vx=V(x);

     surface_layer_shift=b*thick_f;
     res_layer_shift=rau*(c/(surface_layer_shift));
```

```
        Ic0=surface_layer_shift*Jc0;

        Ishift = Vwrite/res_layer_shift;
        J=Ishift/surface_layer_shift;

        if( abs ( Vwrite ) >= abs (Ic0*res_layer_shift) )
        begin //Current higher than critical current => DW propagation

              speed_shifting = (Ishift/surface_layer_shift)*P*Fa;
              duration_shift = abs(1 / ( speed_shifting /c));

              if ( Vwrite > 0 )
              begin

                    if ( Vnextstate == 0 )
                    begin
                          state_layer=0;
                    end

                    else
                    begin
                          state_layer=1;
                    end

              end

              if ( Vwrite < 0 )
              begin //The DW is shifted in the other direction

                    if ( Vnextstate == 0 )
                    begin
                          state_layer=0;
                    end

                    else
                    begin
                          state_layer=1;
                    end

              end

        end

//Static state of MTJ
I(x) <+ -state_layer;

//Ttrans has the same function than x but it includes the time effects
V(Ttrans) <+ transition(V(x),duration_shift,1e-11,1e-11);

//Currents of MTJ
I(T1,T2) <+ Ishift;
```

```
end

endmodule
```

Appendix D Résumé en français

R.1 Introduction

Durant les six dernières décennies, la manipulation de la charge de l'électron a dominé le monde de l'électronique. En particulier, depuis l'avènement de la technologie CMOS (complementary metal-oxide-semiconductor), la croissance de la densité d'intégration des circuits intégrés suit la loi de Moore. Néanmoins, selon les prévisions de l'ITRS (International Technology Roadmap for Semiconductors) illustrée sur la figure R.1, cette augmentation exponentielle atteindra sa limite autour de 2020 alors que la miniaturisation du nœud technologique passera en dessous de 90 nm. Ceci est principalement dû à la consommation statique élevée causée par les courants de fuite. Cette limite pousse les équipes de recherche industrielle et académique explorer de nouvelles technologies pour compléter, voire remplacer l'électronique traditionnelle à base de charges.

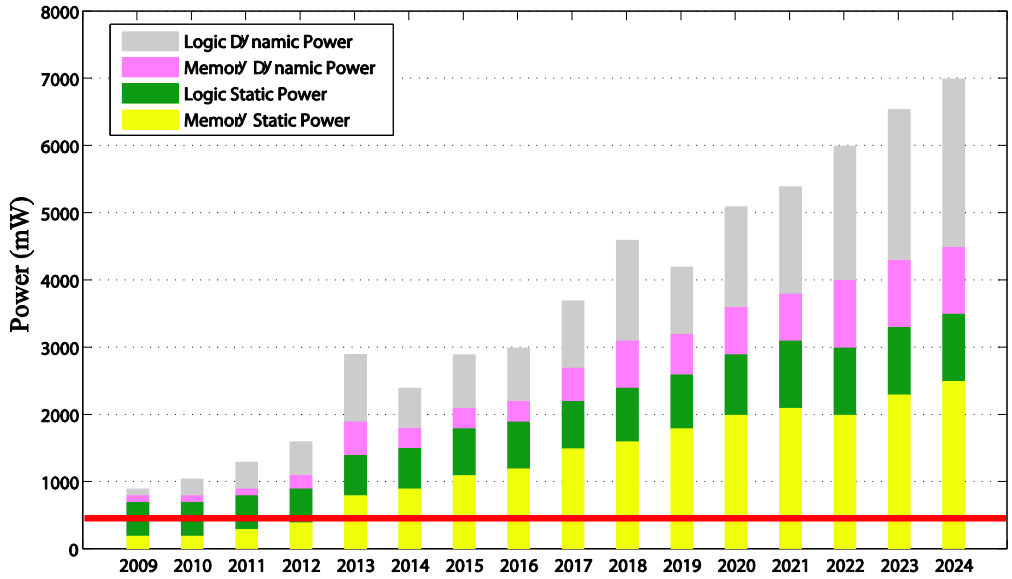


Figure R.1 Prévisions ITRS 2011, jusqu'à l'horizon 2024. (Ligne rouge: exigence de puissances statique et dynamique)

Dans ce contexte, la spintronique ou l'électronique de spin nous constitue un candidat pour résoudre le problème de puissance consommation précédamment exposé. Cette technique utilise non seulement la charge de l'électron mais également le spin, une autre propriété au même titre

que sa charge et sa masse. Parmi tous les effets concernant la spintronique, la découverte de la magnétorésistance géante (GMR) par Fert et Grünberg a eu le plus important impact sur le développement de ce domaine émergent.

Depuis la découverte de la GMR, la spintronique a permis l'émergence de dispositifs et systèmes avancés. Ces applications possèdent de nombreux avantages. Le premier résulte de propriété de non-volatilité permettant de conserver des informations en l'absence d'alimentation externe. Ce caractère peut considérablement diminuer la consumation d'énergie au total. De plus, grâce à l'intégration 3D avec la partie de CMOS, en utilisant les dispositifs spintroniques la distance entre la mémoire et les éléments logiques peut se raccourcir. Comme la puissance dynamique correspond à cette distance, celle-ci peut être diminuée également.

Le but de cette thèse sera d'étudier les dispositifs spintroniques tels que la jonction tunnel magnétique (JTM) et la mémoire racetrack. La JTM est un élément fondamental pour construire la mémoire magnétique (MRAM). Grâce à l'effet tunnel, la JTM permet une plus grande différence de résistance que la GMR. Nous avons choisi de concentrer notre étude sur les méthodes d'écriture attisant le retournement de l'aimantation par courant polarisé en spin (STT). Cette approche permet de simplifier la procédure de la commutation et d'économiser l'énergie comportée à l'écriture par champ magnétique induit (FIMS). La mémoire racetrack est un concept basé sur les parois de domaine propagées par le courant dans les pistes magnétiques. Puisque la distance entre des parois de domaine peut être extrêmement petite, ce concept autorise une très grande densité. Afin de conserver une haute stabilité thermique en diminuant la taille des dispositifs, les matériaux à l'anisotropie magnétique perpendiculaire (AMP) attirent davantage l'intérêt par rapport à l'anisotropie magnétique planaire.

La thèse est organisée en six chapitres, y compris l'introduction et la conclusion. Le chapitre R.2 présente un état de l'art relatif aux dispositifs et circuits hybrides à base de la spintronique. Le chapitre R.3 se concentre sur la modélisation compacte de la JTM avec APM commutée par STT et la mémoire racetrack. De nombreux modèles physiques et paramètres expérimentaux sont intégrés dans ces modèles compacts. Le chapitre R.4 se centre sur les circuits hybrides pour la logique et la mémoire. En utilisant les modèles compacts de la JTM et la mémoire racetrack, un additionneur complet et une mémoire adressable par contenu (CAM) sont conçus, afin d'évaluer les avantages en termes de consommation, vitesse et densité. Au sein du chapitre R.5, deux

améliorations pour augmenter la densité de MRAM et mémoire racetrack sont proposées. Enfin, la thèse est conclue dans le chapitre R.6.

R.2 Etat de l'art

R.2.1 La spintronique

Comme introduit, la spintronique est une inter-discipline émergente dont l'idée principale est de contrôler les degrés de liberté de spin des électrons pour l'électronique. L'effet de magnétorésistance géante (GMR) est réellement le point de départ de la spintronique. Cet effet était observé dans une structure composée des multicouches ferromagnétiques et non-magnétiques. Les différentes orientations relatives des aimantations des couches magnétiques produisent une différence de résistance à cause de la diffusion magnétique dépendante du spin. Ainsi, si les orientations des aimantations des couches sont parallèles, les électrons avec le spin parallèle à l'aimantation peuvent passer à travers le dispositif plus facilement, conduisant à une faible résistance. De l'autre côté, dans le cas où les orientations sont antiparallèles, tous les électrons sont diffusés, il en résulte une forte résistance. La différence des résistances étant importante, cet effet a été nommé GMR. La figure R.2 présente l'effet GMR dans un cas simplifié avec deux couches ferromagnétiques séparées par une couche non-magnétique métallique. Ce type de structure est à la base des vannes de spin utilisées largement pour les disques durs.

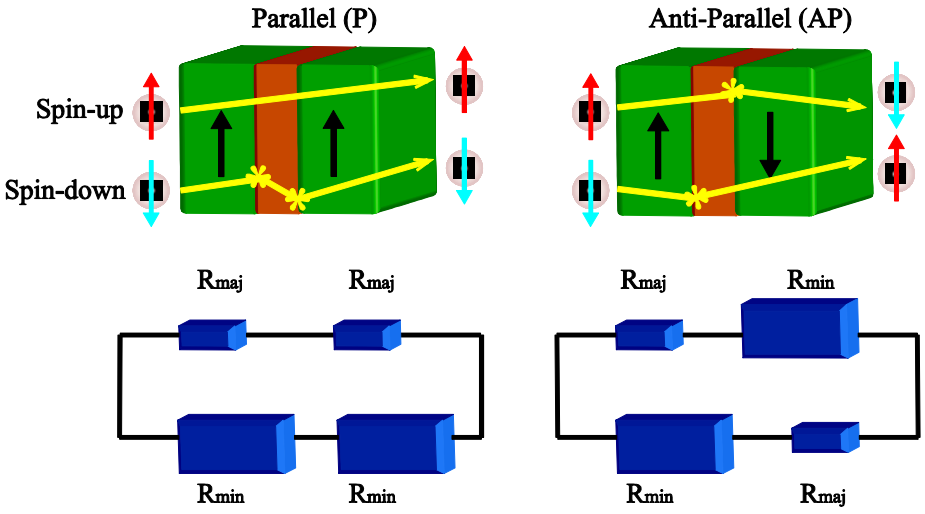


Figure R.2 Effet GMR dans la vanne de spin.

R.2.2 Dispositifs spintroniques

R.2.2.1 Jonction tunnel magnétique (JTM)

R.2.2.1.1 Structure de JTM

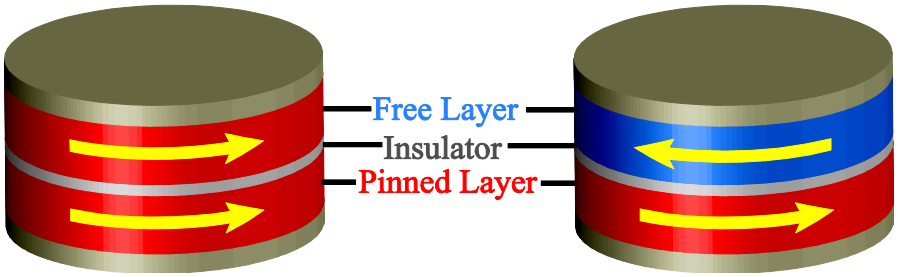


Figure R.3 Structure de la JTM

A la différence des dispositifs basés sur la GMR, les deux couches ferromagnétiques de la JTM sont séparées par une fine couche isolante. Comme l'illustre sur la figure R.3, une JTM possède une couche ferromagnétique à aimantation fixée dans une orientation spécifique et une autre à aimantation libre. La JTM basé sur l'effet magnétorésistance tunnel (TMR) permet une différence de résistance plus élevée qu'avec une vanne de spin. Normalement, on utilise le rapport de TMR donné par l'Eq. R1 pour décrire la magnétorésistance d'une JTM.

$$TMR = \frac{\Delta R}{R_P} = \frac{R_{AP} - R_P}{R_P} = \frac{G_P - G_{AP}}{G_{AP}} \qquad \text{(R1)}$$

L'utilisation d'une barrière cristalline MgO qui peut fournir un grand TMR a constitué une avancée majeure dans le domaine des JTMs. A ce jour, le record de TMR des JTMs avec MgO peut atteindre jusqu'à 600% à la température ambiante. Ces résultats revêtent une grande importance non seulement pour assurer l'adaptation avec la technologie CMOS et tolérer les variations de paramètres, mais aussi pour miniaturiser la surface des amplificateurs de lecture.

R.2.2.1.2 Couple de transfert de spin (STT)

Il existe plusieurs approches de commutation pour les JTMs. La première approche consiste à utiliser le champ magnétique produit par une ligne de courant. Toutefois cette approche consomme beaucoup d'énergie et de surface. Une autre approche est fondée sur l'effet de

transfert de spin (STT) prédit théoriquement par Berger and Slonczewski en 1996. Ils ont trouvé qu'un courant électrique polarisé en spin portant un moment cinétique pourrait transférer ce moment cinétique à l'aimantation d'une couche magnétique. Lorsque l'amplitude du courant est supérieure à la valeur de seuil (le courant critique), l'aimantation peut être renversée. Dans un premier temps, cet effet était utilisé pour des vannes de spin. Ensuite il a aussi été observé dans des JTMs. Dans une telle structure, une couche ferromagnétique fonctionne comme un polariseur en spin du courant électrique qui transfère le moment magnétique en appliquant un couple sur l'aimantation de la deuxième couche ferromagnétique. Cette approche simplifie considérablement le processus de commutation car il nécessite seulement un courant bidirectionnel. De plus, l'amplitude du courant pour la STT est normalement d'un ordre de grandeur plus faible que celui destiné à générer un champ magnétique. En conséquence, l'approche de commutation STT est largement considérée comme l'une des plus prometteuses pour les MRAMs futures.

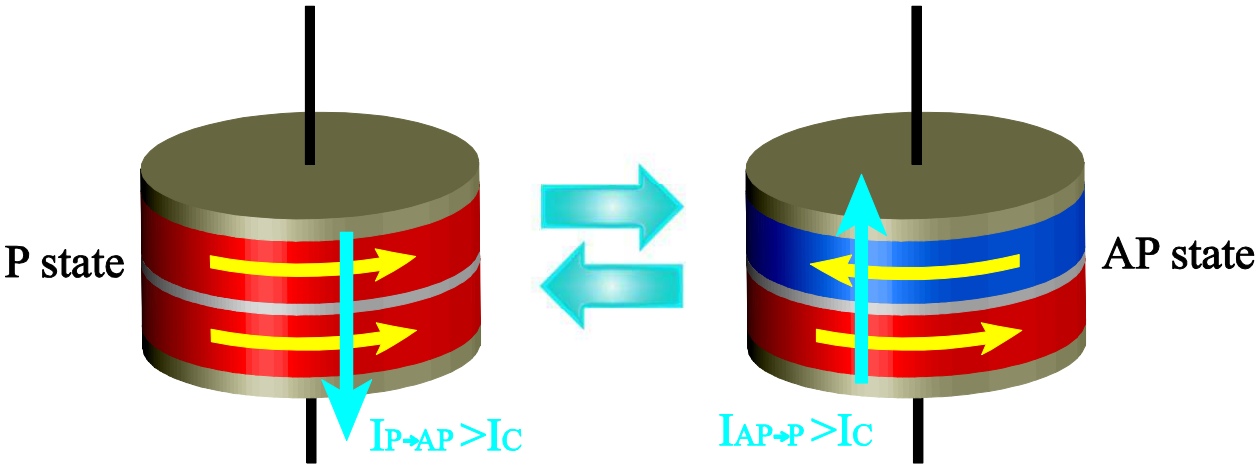


Figure R.4 Commutation de STT

R.2.2.1.3 Anisotropie magnétique perpendiculaire (APM)

Pour réaliser la JTM avec de hautes performances pour les futures applications logiques et de mémoire, il y a généralement cinq critères d'évaluation: petite taille, haute TMR, faible courant de commutation, comptabilité avec les procédés CMOS et haute stabilité thermique. Au fur et à mesure que sa taille diminue, il est de plus en plus difficile d'atteindre ces critères avec la JTM conventionnelle à anisotropie magnétique planaire. Les progrès récents des matériaux ont montré que la JTM avec APM pourrait offrir un plus faible courant critique, une plus grande vitesse de

commutation et une plus haute stabilité thermique que la JTM avec l'anisotropie magnétique planaire. Ceci est dû au champ de démagnétisation et à l'anisotropie magnétique plus importante qui réduisent le courant critique tout en conservant une stabilité thermique relativement élevée. Une variété des matériaux a été tentée, parmi lesquelles, Ta/CoFeB/MgO démontre une caractéristique excellente grâce à l'anisotropie d'interface entre les matériaux. Il offre un bon compromis entre la taille (40 nm), le courant critique (~50 µA), la stabilité thermique (40 k_BT) et le ratio TMR (> 100%).

R.2.2.2 Paroi de domaine magnétique

Le terme « paroi de domaine magnétique » est utilisé pour décrire les zones de transition entre différents domaines magnétiques, dans lesquels les vecteurs d'aimantation tournent graduellement. L'utilisation de champ magnétique est une approche pour faire propager une paroi de domaine magnétique. Cependant, elle soulève plusieurs inconvénients critiques pour des applications pratiques, par exemple, leur vitesse est faible et leur consommation d'énergie est importante à cause de la génération de champ magnétique. Dans ce contexte, l'utilisation d'un courant polarisé en spin pour propager les parois de domaine peut ouvrir de nouvelles perspectives d'applications.

R.2.2.2.1 Propagation de parois de domaine magnétique sous courant

Le fait que les parois de domaine magnétique puissent être poussées par un courant a été d'abord proposé par Berger en 1978. Depuis ce moment-là, beaucoup d'efforts et de travaux de recherche ont porté sur ce phénomène. En 1996, dû à la prédiction de STT, ce phénomène était relié au « transfert de spin » qui est largement utilisé pour l'instant pour décrire la dynamique de parois de domaine.

Grâce aux progrès des techniques de nano-fabrication et de mesure, des nano-pistes de taille inférieure à 100 nm peuvent être gravés. De nombreux matériaux ont été utilisés pour fabriquer ces nano-pistes. Dans cette thèse, nous étudions essentiellement le mouvement de parois de domaine induits par courant dans les pistes magnétiques à AMP, qui s'est révélé exiger moins de densité de courant pour la propagation que ceux à l'anisotropie magnétique planaire.

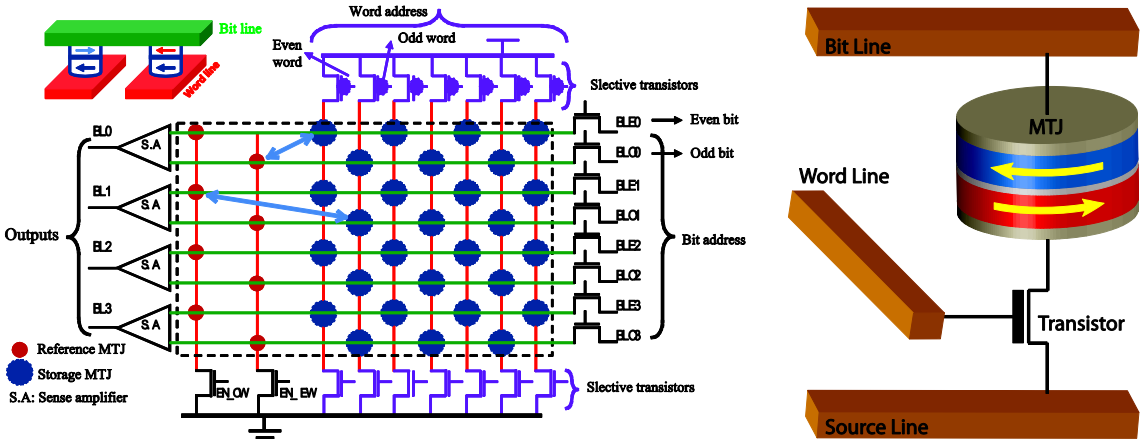
R.2.3 Circuit logique et de la mémoire hybride à base de la spintronique

R.2.3.1 MRAM

La mémoire magnétique (MRAM) est une des applications spintroniques les plus importantes. Elle constitue une voie prometteuse pour construire une mémoire universelle grâce à ses nombreux avantages, tels que la non-volatilité, l'endurance infinie, le faible temps d'accès et la compatibilité avec les procédés CMOS.

Du point de vue architectural, deux types de cellules ont été proposés pour réaliser des MRAMs. L'architecture « cross-point » permet une très haute densité, cependant il souffre de plusieurs problèmes techniques de conception, par exemple, les courants parasites (« sneak path ») et la faible vitesse d'accès. Afin d'éliminer ces défauts, un transistor est toujours ajouté dans chaque cellule pour créer l'architecture « 1T/1JTM ». Ces deux architectures sont illustrées dans la figure R.5.

En fonction des différentes approches de commutation de la JTM, la MRAM a connu plusieurs générations. Parmi celles-ci, la MRAM basée sur l'effet de transfert de spin (STT-MRAM) démontre la meilleure performance en termes de densité, temps d'accès et consommation d'énergie.



(a) (b)

Figure R.5 (a) Schéma de l'architecture cross-point pour la MRAM. (b) Schéma de l'architecture de 1T/1JTM pour la MRAM.

R.2.3.2 Mémoire racetrack

La mémoire racetrack est un concept émergent basé sur la propagation de parois de domaine sous courant. Il a été proposé par Parkin en 2008. Deux JTMs sont utilisées comme têtes d'écriture et de lecture. Une piste magnétique est utilisée pour stocker et transférer les données, comme l'illustre la figure R.6. La tête d'écriture injecte les domaines magnétiques dans la piste. Un courant polarisé en spin entraine ensuite les parois de domaine de la tête d'écriture à la tête de lecture. Comme la distance entre les parois adjacentes peut être extrêmement petite, ce concept est considéré comme ayant un fort potentiel pour atteindre une ultra-haute densité. Grâce à l'intégration de JTMs, la mémoire racetrack peut être compatible avec le CMOS. Des exemples de circuits d'écriture et lecture sont présentés sur la figure R.6. Ils seront présentés en détails dans la partie suivante. Le premier prototype a été démontré par IBM en 2011, cependant il s'est basé sur l'anisotropie magnétique planaire et sa capacité était très limite (256 bit). Dans cette thèse on focalise sur la mémoire racetrack basée sur les matériaux à AMP (e.g. CoFeB/MgO, CoNi).

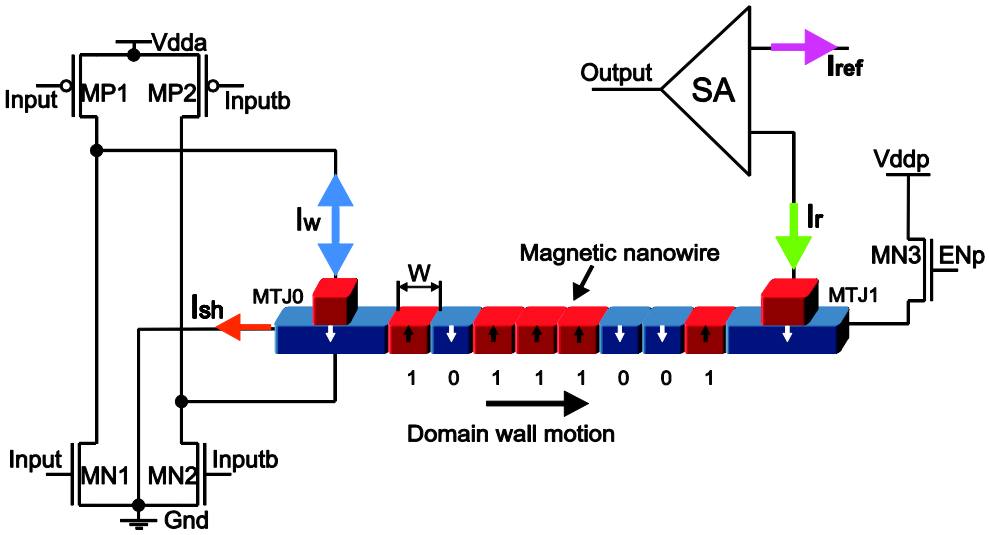


Figure R.6 Mémoire racetrack basée sur la propagation des parois de domaine sous courant.

R.2.3.3 Logique-en-mémoire

Le concept « logic-in-memory » correspond à une architecture dans laquelle des éléments de mémoire sont répartis dans les unités arithmétiques, permettant l'amélioration des performances de vitesse et d'énergie par rapport à l'architecture classique de Von-Neumann. Depuis

l'apparition des mémoires non-volatiles, plusieurs unités logiques ont été conçues ou fabriquées à base de logic-in-memory, tel que l'additionneur complet et le flip-flop.

La figure R.7 présente une architecture générale de logique-en-mémoire à base de STT-MRAM qui est composé de trois parties : un circuit de lecture (e.g. pre-charge sense amplifier (PCSA)) évaluant les résultats logiques sur les sorties ; un bloc d'écriture programmant la STT-MRAM ; un bloc de commande logique pour les données. Selon l'état des MOS dans l'arbre logique et l'état d'éléments de STT-MRAM, les courants de décharge sont différents dans les deux branches et le circuit de lecture produit une valeur logique.

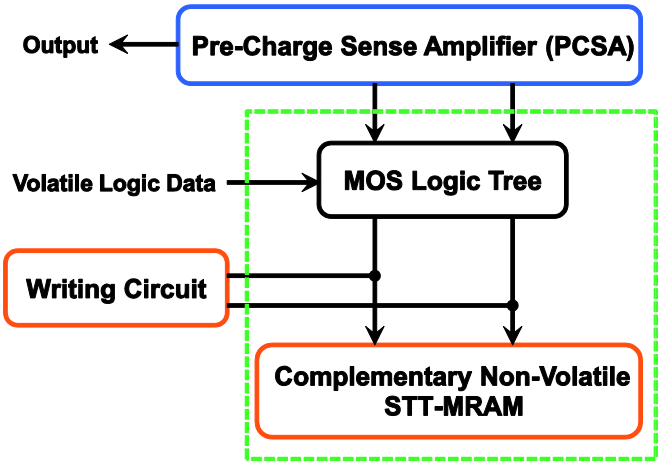


Figure R.7 Architecture générale de logique-en-mémoire à base de STT-MRAM.

R.3 Modélisation compacte de la STT JTM à AMP et la mémoire racetrack à AMP

R.3.1 Modèle physique de la STT JTM à AMP

De nombreux paramètres expérimentaux et modèles physiques ont été intégrés dans le modèle pour obtenir une bonne cohérence avec les mesures expérimentales. En particulièr, les modèles physiques relatifs aux résistances de JTM doivent être d'abord inclus, par exemple, le modèle de résistance de barrière tunnel et le modèle de TMR dépendant de la tension. Ensuite, les modèles concernant le comportement de la commutation STT doivent être intégrés, y compris les modèles statique, dynamique et stochastique. Le modèle statique est utilisé pour calculer le courant

critique. Le modèle dynamique définit la relation entre le courant et la durée nécessaire pour la commutation. Le modèle stochastique prend en compte les influences des fluctuations thermiques. La figure R.8 illustre la hiérarchie des modèles physiques intégrés dans le modèle compacte de la STT JTM à AMP.

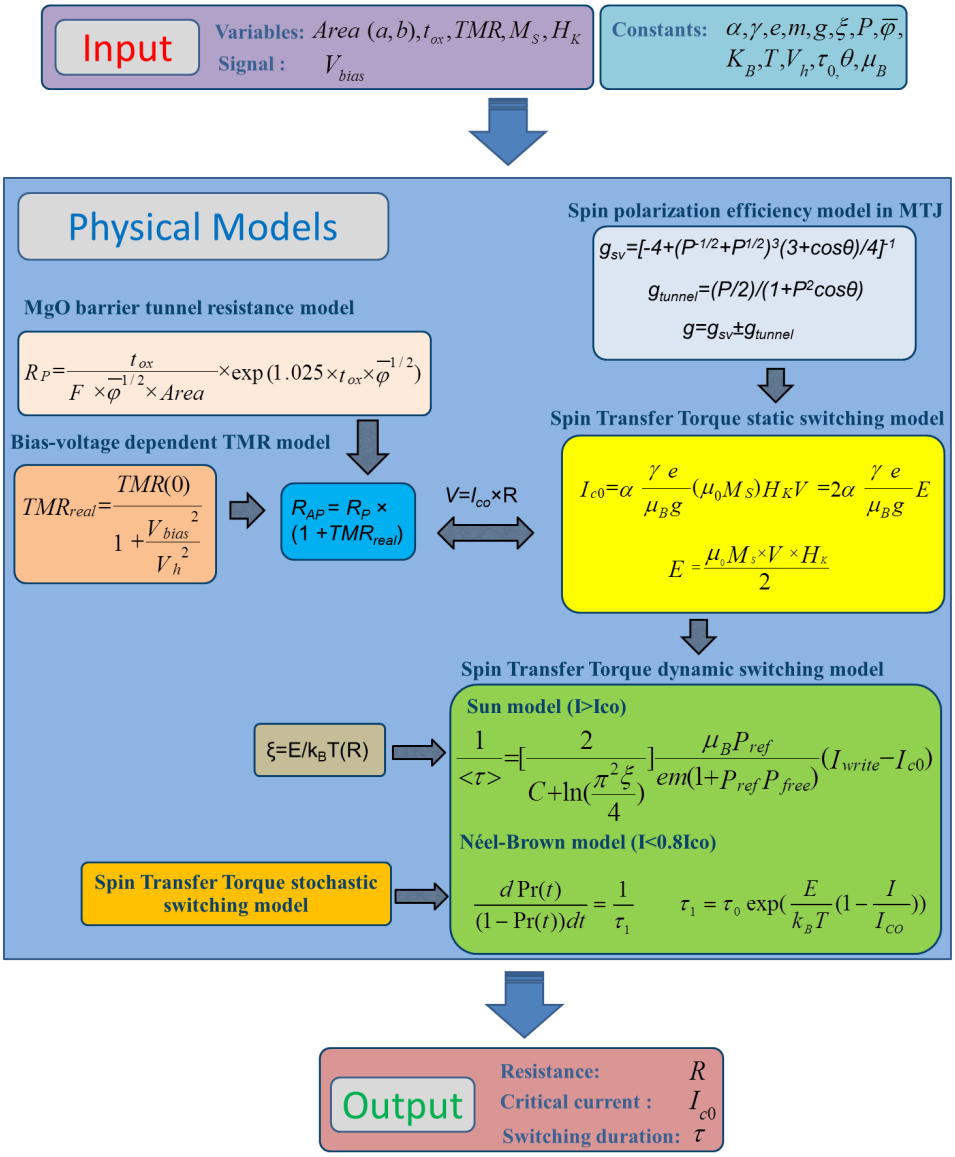


Figure R.8 Hiérarchie d'intégration des modèles physiques dans la STT JTM à AMP.

R.3.2 Modèle physique de la mémoire racetrack à AMP

Comme l'illustre la figure R.6, la nucléation et la détection des domaines magnétiques s'exécutent par les têtes de JTMs, alors les modèles physiques concernant ces comportements

sont les mêmes que ceux de JTM. Afin de compléter la modélisation de la mémoire racetrack, il faut ajouter un modèle pour décrire la propagation des parois de domaine. Comme la vitesse est cruciale pour les dispositifs basés sur les parois de domaine, nous avons décidé d'intégrer un modèle 1D dans le modèle compact pour calculer la vitesse de propagation des parois de domaine. Considérant la seule influence du courant, la vitesse en fonction de la densité de courant s'écrit sous la forme :

$$V_j = u = \frac{\mu_B P j_p}{e M_S} \qquad \text{(R2)}$$

En comparant un certain nombre de langages de modélisation, tels que Verilog-A, C et VHDL-AMS, ces modèles nous avons décidé de programmer avec le langage Verilog-A, qui est compatible avec les outils de CAO (e.g. la plate-forme Cadence) et fournit une interface simple de configuration. Les utilisateurs peuvent modifier les variables via l'interface de paramètres pour les adapter à leurs besoins spécifiques.

R.3.3 Validation de la STT JTM à AMP

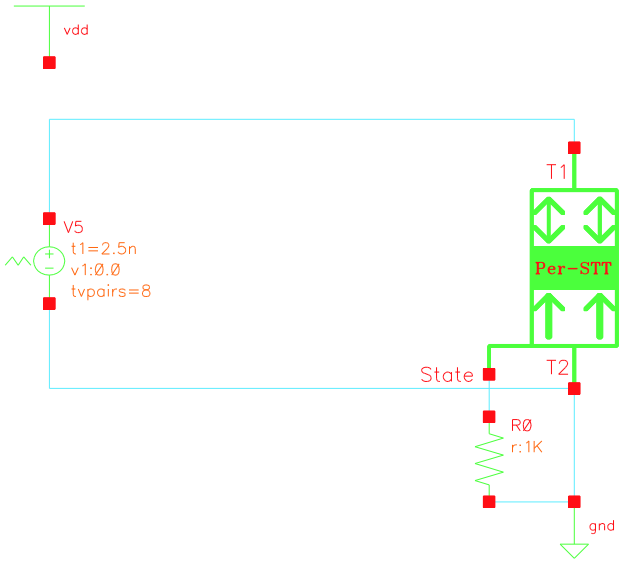


Figure R.9 Schéma de simulation de la STT JTM à AMP.

Au-delà de l'intégration des modèles physiques, nous avons effectué leur validation au travers des diverses simulations. La figure R.9 montre le schéma pour les simulations de la STT JTM à

AMP. La simulation DC est d'abord faite pour tester la caractéristique statique de commutation STT. La figure R.10 illustre que les courants critiques sont les mêmes que les résultats expérimentaux.

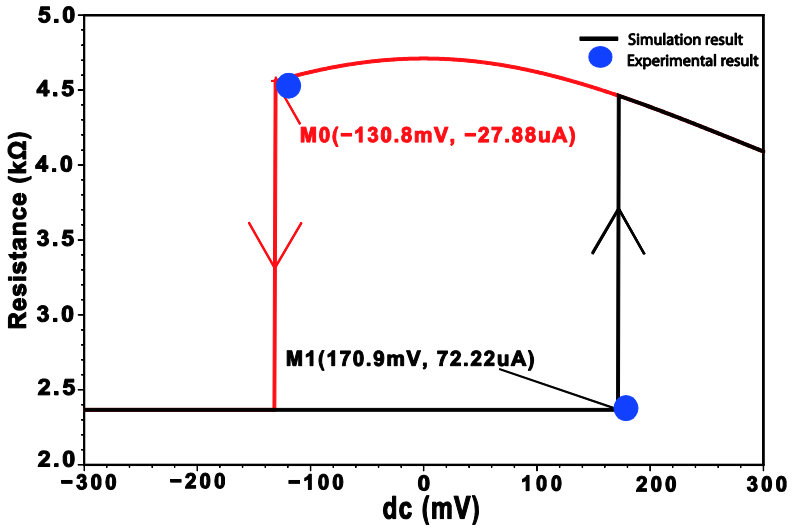


Figure R.10 Simulation DC de la STT JTM à AMP.

La simulation transitoire est ensuite effectuée pour valider le comportement dynamique de ce modèle compact. La figure R.11 montre que le délai de commutation est inversement proportionnel au courant d'écriture.

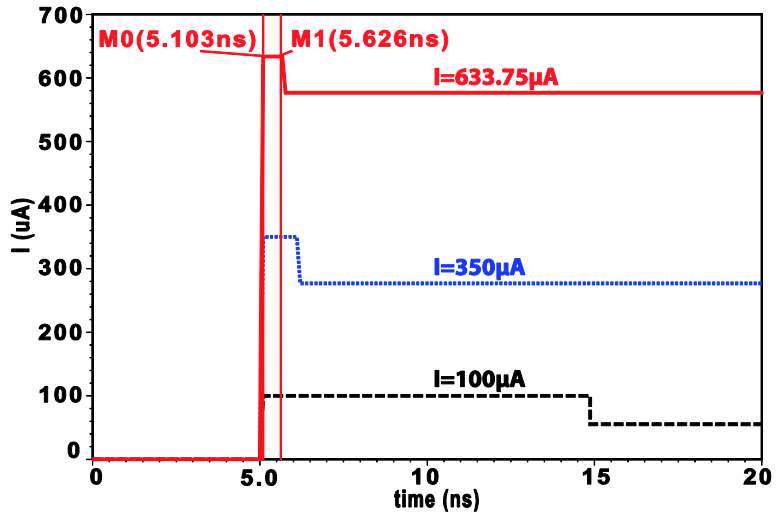


Figure R.11 Simulation transitoire de la STT JTM AMP.

Finalement, l'effet stochastique s'observe par simulation Monte-Carlo. Comme le montre la figure R.8, l'effet stochastique est divisé en deux régimes décrivant sa dynamique : le régime « Néel-Brown » pour la lecture (où le courant est toujours inférieur au courant critique) et le régime « Sun » pour l'écriture (où le courant est supérieur au courant critique). Les simulations correspondant aux deux régimes sont présentées respectivement, par les figures R.12 et R.13.

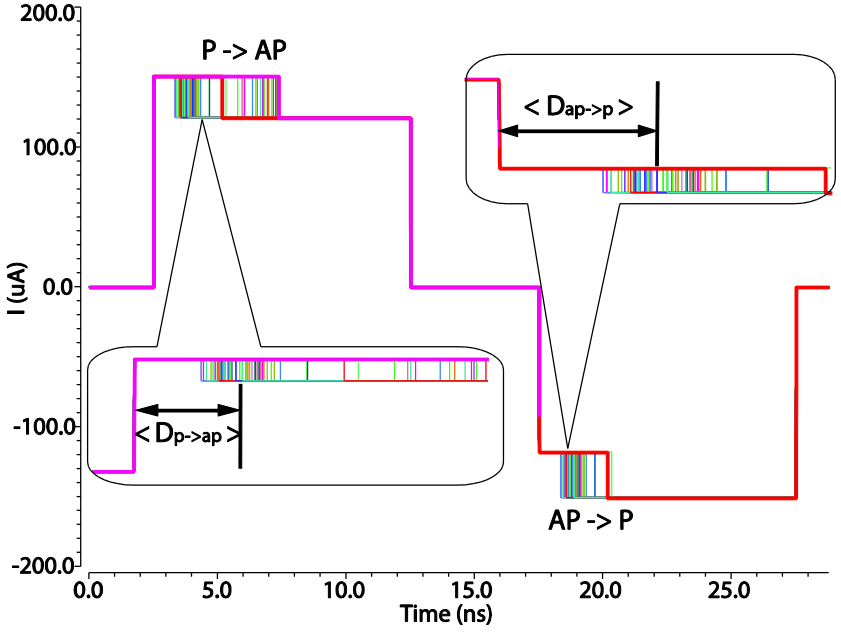


Figure R.12 Simulation Monte-Carlo de la STT JTM à AMP.

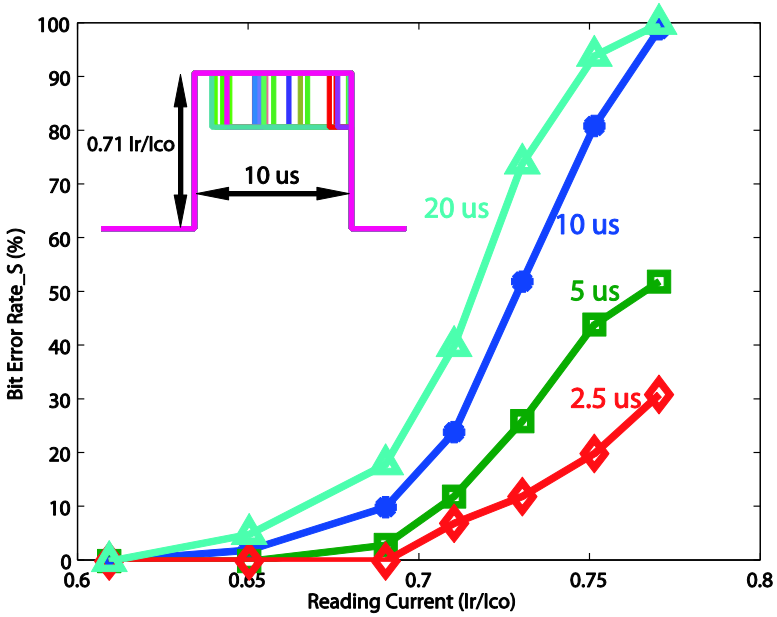


Figure R.13 Dependance de taux d'erreur de lecture versus current de lecture pour différente durées d'impulsion.

R.3.4 Validation de la mémoire racetrack à AMP

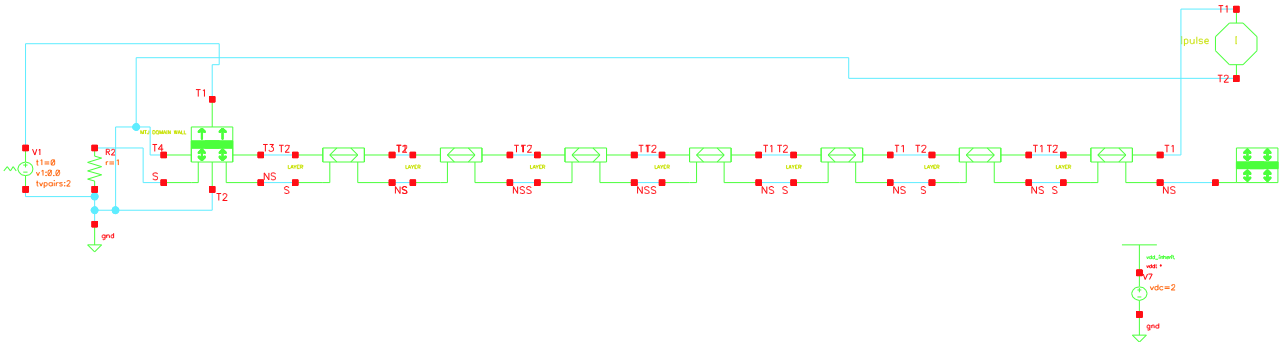


Figure R.14 Schéma de simulation de la mémoire racetrack à AMP.

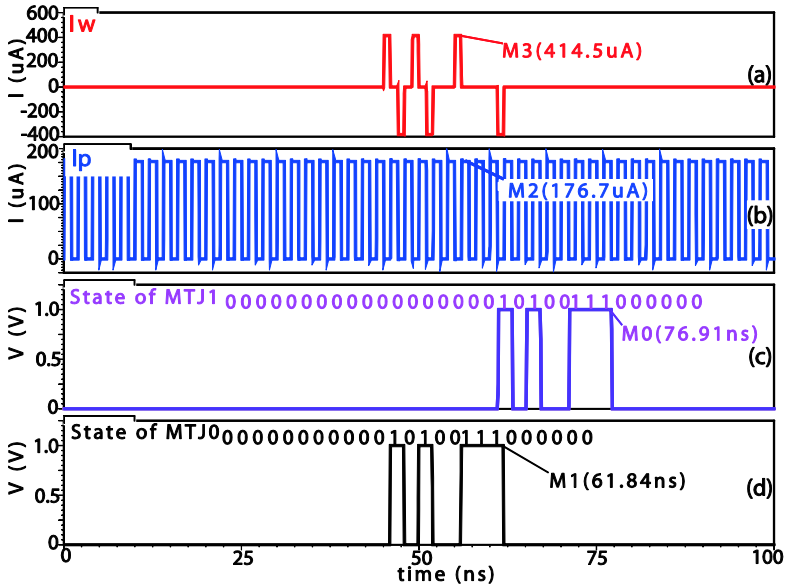


Figure R.15 Simulation transitoire de la mémoire racetrack à AMP.

Ensuite, une mémoire racetrack à AMP de 8 bits est simulée pour valider la fonctionnalité de son modèle compact. La figure R.14 présente le schéma, y compris la tête d'écriture, la tête de lecture, une piste magnétique et les circuits périphériques. La figure R.15 montre les résultats de simulation transitoire en utilisant un courant de propagation à la fréquence de 500 MHz. Grâce à la haute vitesse de commutation de l'aimantation dans les matériaux à AMP, la durée de la propagation de parois de domaine peut être égale à celle de la nucléation. Ainsi, le signal carré, étant généré plus facilement et utilisé plus largement, peut être utilisé pour entrainer les parois de domaine dans la mémoire racetrack. Par ailleurs, en prenant en compte l'effet stochastique, des simulations Monte-Carlo ont été réalisées également pour cette mémoire racetrack. Les cas avec

différentes capacités sont simulés et comparés. Les résultats des simulations montrent la relation de compromis entre la capacité, la vitesse et la fiabilité : 1) la mémoire racetrack avec une plus grande capacité de stockage est moins fiable que celle avec une plus faible capacité, à fréquence identique ; 2) la haute fiabilité s'obtient en réduisant la vitesse : une plus grande capacité nécessite une plus grande durée d'impulsion de courant pour atteindre la même fiabilité.

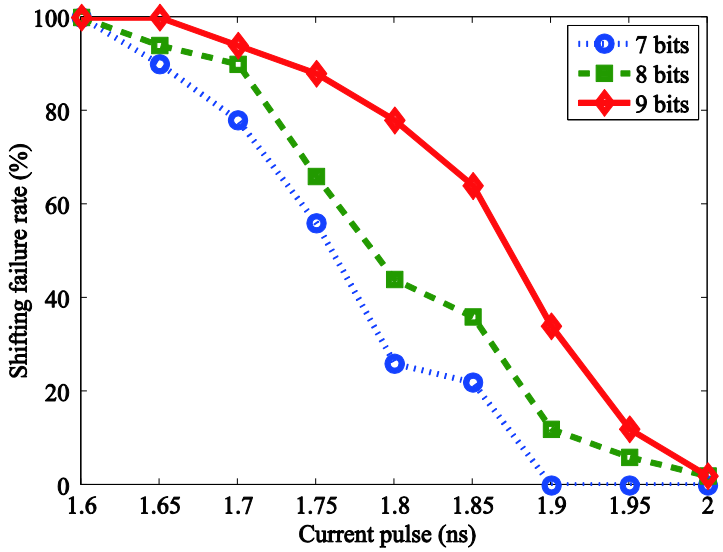


Figure R.16 Dependance de taux d'erreur de propagation versus durées de courant de propagation et capacité de mémoire racetrack.

R.4 Design de circuits hybrides spintronique/CMOS

R.4.1 Additionneur complet magnétique (ACM)

Afin de surmonter le problème de l'augmentation de la consommation statique liée aux courants de fuite et de la consommation dynamique, un processeur magnétique fondé sur des dispositifs spintroniques est proposé. Comme l'addition constitue l'opération élémentaire de l'unité arithmétique et logique de tous les processeurs, l'additionneur complet magnétique (ACM) intéresse spécialement les groupes de recherche académiques et industriels. Ici, nous présentons un ACM 1-bit basé sur la STT JTM à AMP (STT ACM) et un ACM multi-bit basé sur la mémoire racetrack à AMP.

R.4.1.1 1-bit ACM basé sur la STT JTM à AMP (STT-ACM)

Comme l'illustre la figure R.17, le STT-ACM est à base de la structure générale introduite dans le chapitre R.1. L'évaluation de la fonction logique utilise le circuit pre-charge sense amplifier (PCSA) pour assurer une meilleure performance en termes de fiabilité de lecture et une consommation réduite tout en conservant une haute vitesse. Les entrées sont « A », « B » et « C_i », et les sorties sont « Sum » et « C_0 ». Parmi celles, l'entrée « B » se rapporte à une mémoire non-volatile.

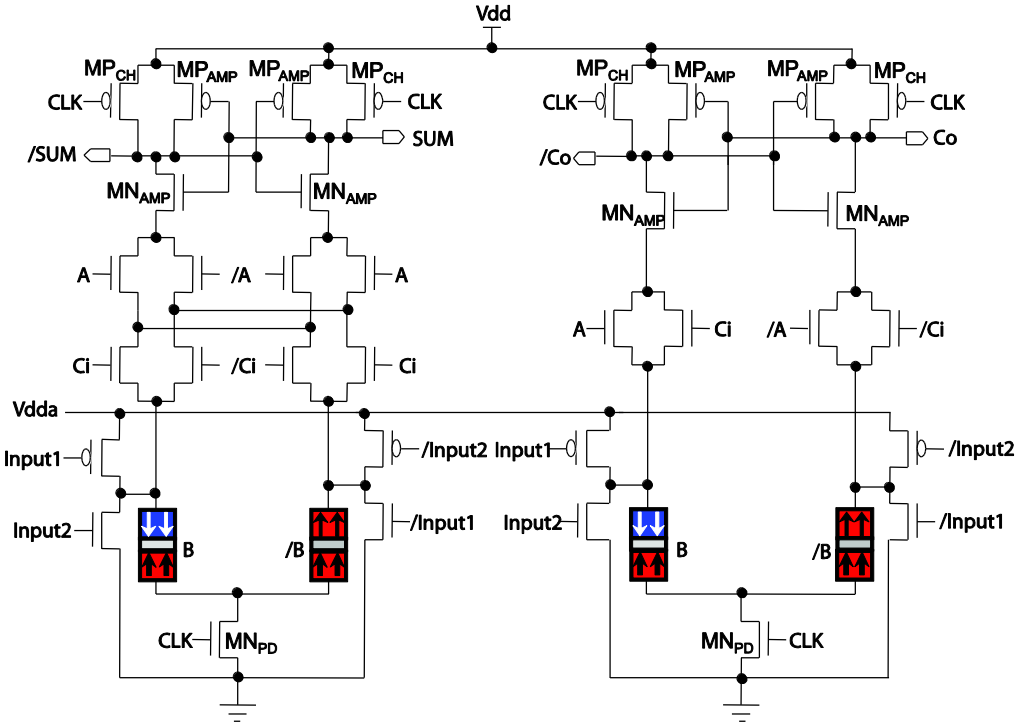


Figure R.17 Architecure de 1-bit STT ACM.

La fonction d'addition de ce STT-ACM est illustrée par la simulation transitoire sur la figure R.18. Par exemple, tant que « A »= '1', « B »= '0' et « C_i »= '0', le résultat « Sum » est '1' et il n'y a pas de retenue ; tant que « A »= '1', « B »= '0' et « C_i »= '1', le résultat « Sum » est '0' et il y a une retenue.

Nous comparons le STT-ACM proposé avec l'additionneur complet basé sur CMOS en termes de durée, puissance consumée et taille. Grâce à l'architecture « Logic-in-memory » utilisée pour ce nouveau STT-ACM, il est possible d'économiser de façon très importante le temps et la

consommation dynamique associée au transfert de données. De plus, l'intégration 3D de STT-MRAM avec les procédés CMOS permet de diminuer la surface totale. Néanmoins, le produit énergie-délai (EDP) est supérieur à celui d'un additionneur complet CMOS d'environ 10%, parce qu'il nécessite un temps supplémentaire pour l'opération de lecture avec le PCSA. Ce nouveau design, grâce à la non-volatilité de STT-MRAM, permet d'éteindre totalement le circuit, réduisant ainsi la consommation statique jusqu'à 0,75 nW. Ainsi, le STT-ACM pourrait réduire la consommation complète d'un système de calcul, en particulier pour ceux qui sont la plupart du temps à l'état OFF.

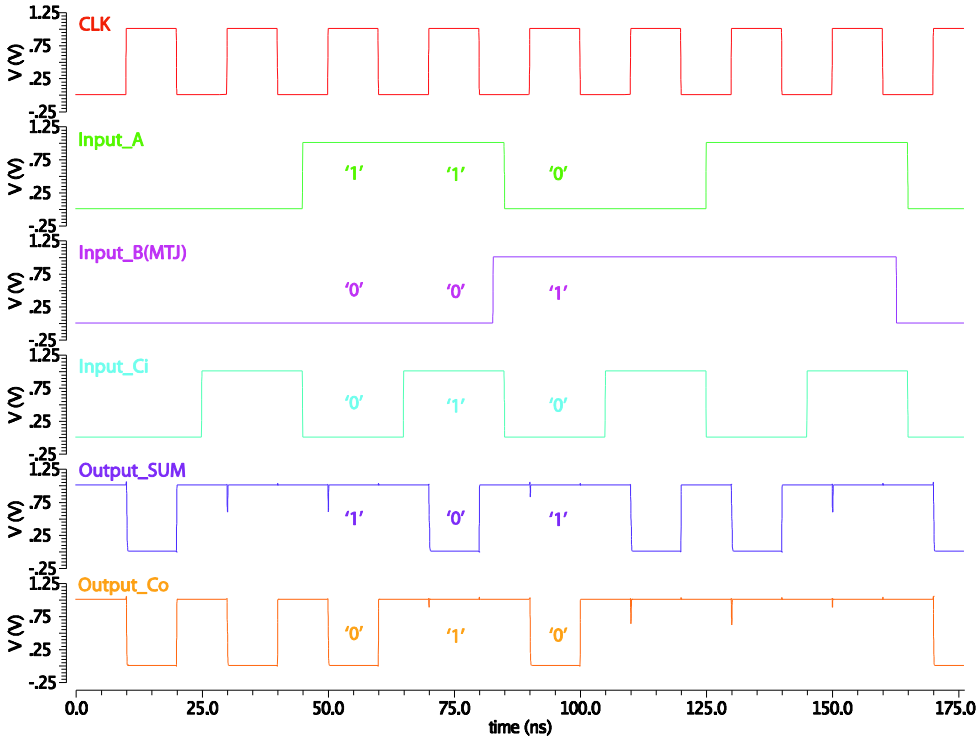


Figure R.18 Simulation transitoire de 1-bit STT-ACM.

Table R.1 Comparaison entre le 1-bit STT ACM proposé et l'additionneur complet de CMOS.

Performance	CMOS full adder (40 nm)	STT-MFA
Delay time	75 ps	87.4 ps
Dynamic power @500 MHz	2.17 µW	1.98 µW
Standby power	71 nW	<1 nW
Data transfer energy	>1 pJ/bit	<1 fJ/bit
Die area	46 MOS	38 MOS + 4 MTJs

R.4.1.2 Multi-bit ACM basé sur la mémoire racetrack à AMP

La mémoire racetrack se distingue par le fait qu'il peut stocker et transférer plusieurs bits de données en propageant des parois de domaine dans une nano-piste magnétique. Cette caractéristique avantageuse permet de concevoir un ACM multi-bit rapide et compact.

La figure R.19 présente le schéma détaillé du circuit permettant de calculer la retenue y compris les circuits périphériques (circuit PCSA, circuit d'écriture et circuit de propagation). Chaque bit de données est stocké par une paire de pistes magnétiques de configurations complémentaires. Cela permet de minimiser l'effet de la variabilité lorsque le même courant de propagation est utilisé pour propager les parois de domaine. A noter que la propagation des parois de domaine est mise en œuvre dans la phase de pré-charge de PCSA afin d'éviter de perturber l'évaluation de la sortie.

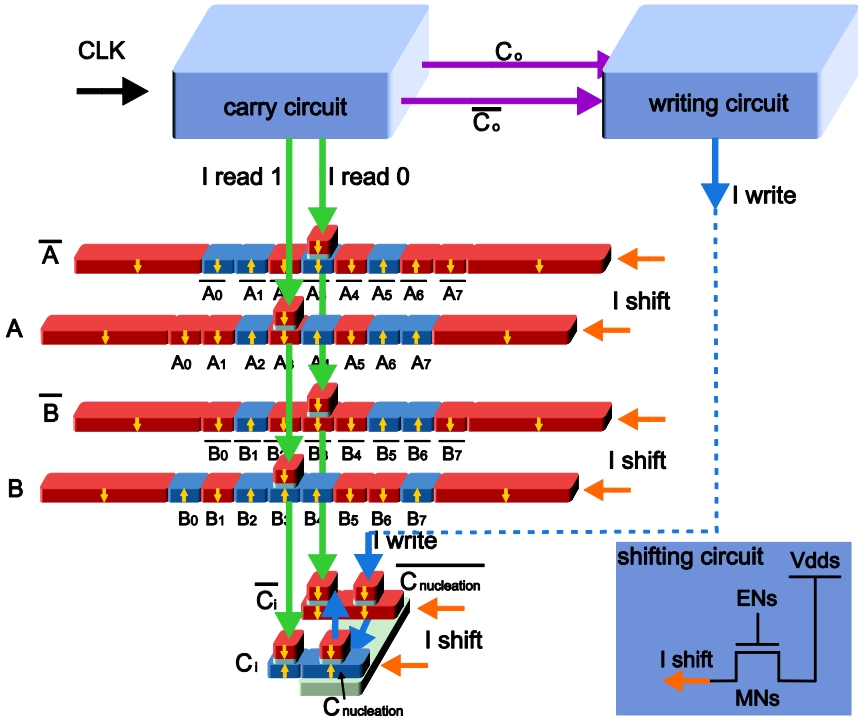


Figure R.19 Circuit de retenue de multi-bit ACM base sur la mémoire racetrack à AMP.

La simulation transitoire a été faite à l'aide du modèle compact pour valider le fonctionnement correct de cet ACM multi-bit. La figure R.20 montre l'opération d'addition de deux mots 8-bit arbitraires: « A »= « 01110011 » et « B »= « 01011010 ». L'addition s'effectue en série poids

faibles en tête. Les résultats de simulation « Sum »= « 11001101 » et « C_0 »= « 01110010 » confirment le bon fonctionnement.

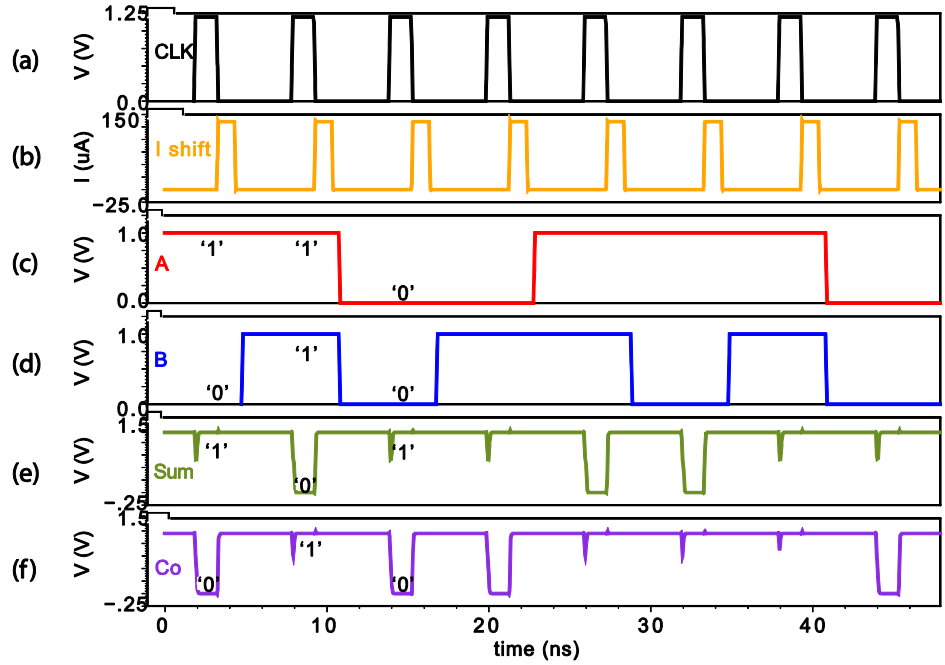


Figure R.20 Simulation transitoire du 8-bit ACM proposé.

Table R.2 Comparaison entre le 8-bit ACM proposé et l'additionneur complet de CMOS.

Performance	CMOS full adder (65 nm)	Proposed MFA
Write time	200 ps	2 ns
Write energy	16 fJ/8 bits	(21.39+29) pJ/8 bits
Transfer time	~ ns	0
Transfer energy	8 pJ/mm (for 8 bits)	~0
Die area	310 MOS	23 MOS + 18 MTJs

Afin de comprendre les avantages et les inconvénients de cet ACM multi-bit, nous comparons ses performances avec celles d'un additionneur CMOS. Au travers les résultats de comparaison illustrés dans le tableau R.2, nous trouvons que la surface de cet ACM basé sur la mémoire racetrack est considérablement réduite par rapport à son équivalent CMOS. Par contre, l'ACM proposé consomme six fois plus d'énergie dynamique que sa version CMOS puisque l'énergie nécessaire pour la nucléation et la propagation des parois de domaine est encore très élévée avec la technologie actuelle. Cependant, nous n'avons pas encore considéré l'énergie statique dans

cette comparaison. Comparé au circuit CMOS, l'ACM proposé peut éliminer l'énergie statique nécessaire au maintien des données mémorisées : grâce à sa non-volatilité, ce qui réduit l'énergie totale pour le système entier.

R.4.2 CAM basé sur l'APM mémoire racetrack

La mémoire adressable par contenu (CAM) est une mémoire dont la particularité est de produire sur sa sortie l'adresse de la donnée recherchée. Son fonctionnement parallèle la rend très rapide. Par conséquence, il est largement utilisé pour les téléphones portables, les routeurs d'internet et les processeurs afin de fournir un accès rapide aux données et une ultra-haute densité. Les CAMs traditionnelles sont composées de grande capacité de SRAMs volatiles qui conduisent à une forte consommation statique et une importante surface.

Dans ce contexte, nous proposons une conception de la CAM à base d'une paire de mémoire racetrack à états complémentaires. Sa caractéristique non-volatile réduit l'énergie statique. Les circuits pour la nucléation et la propagation de parois de domaine sont globalement partagés pour améliorer l'efficacité de surface. La structure de pistes complémentaires permet la détection locale et une plus grande vitesse lors de la recherche de données.

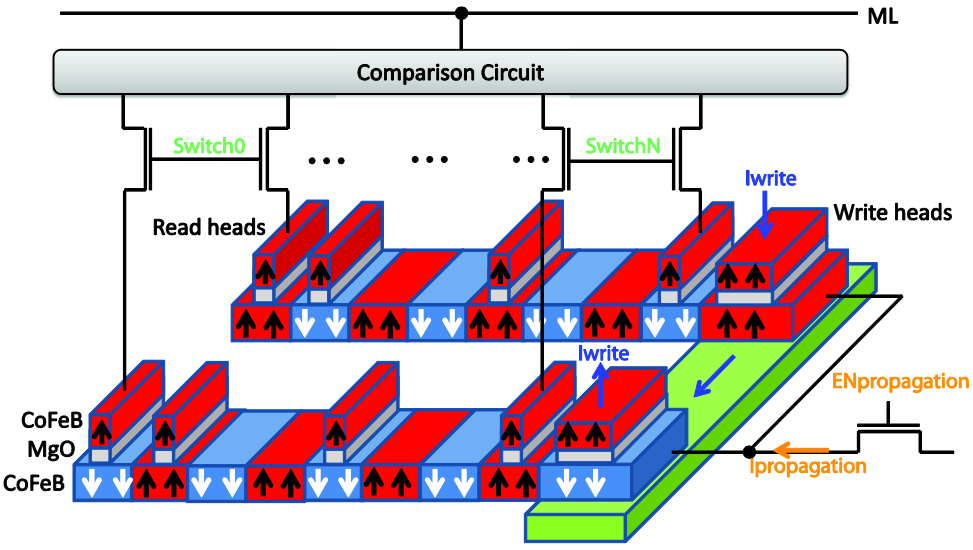


Figure R.21 Structure de CAM basée sur l'AMP mémoire racetrack.

Comme le montre la figure R.21, la CAM proposée consiste en un circuit de comparaison, des mémoires racetrack et des circuits pour la nucléation et la propagation de parois de domaine.

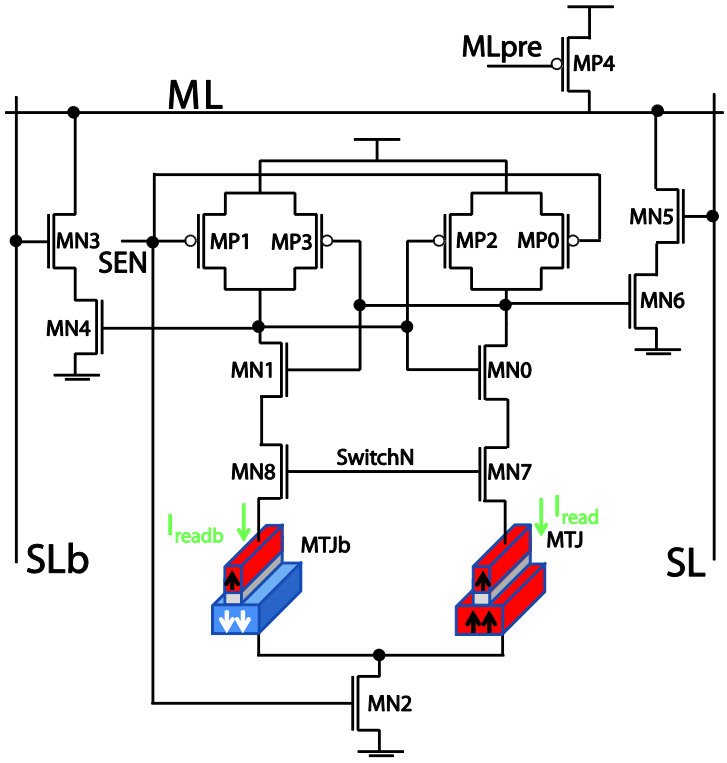


Figure R.22 Schéma de circuit de comparaison.

Le circuit de comparaison, comme présenté sur la figure R.22, est composé de deux parties : un PCSA, détectant les aimantations complémentaires des têtes de lecture par deux impulsions de courant, produisant une valeur logique ; les transistors MN3-MN6 réalisent une CAM classique de NOR type. Le signal « MLpre » est utilisé pour pré-charger la ligne de correspondance (ML). Dans le cas où la ligne de recherche « SL » (dont « SLb » est le complémentaire) correspond aux données stockées, il n'y a pas de chemin pour mettre à zéro le signal ML qui sera ainsi maintenu. Dans le cas contraire, le signal ML sera forcé à zéro.

Afin de réduire surface occupée, chaque paire de pistes magnétiques partage le circuit de comparaison. En même temps, le même circuit d'écriture est aussi partagé pour chaque piste magnétique. En considérant une piste extrêmement longue, la partie CMOS pour chaque cellule de stockage pourrait être ignorée.

Des simulations transitoires de la CAM de 8 x 8 bits ont été effectuées en utilisant le modèle compact de mémoire racetrack à AMP. Dans un premier temps, nous ignorons les propagations de parois de domaine, considérant que les données que nous voulons chercher sont stockées dans

la CAM. Le résultat de simulation pour ce cas est présenté sur la figure R. 24(a). Ainsi, nous montrons qu'une opération de recherche ne prend que 0,45 ns, ce qui est plus rapide que les CAMs traditionnelles. De plus l'énergie consommée par la recherche qui n'excède pas ~12 fJ/bit/recherche, peut être encore réduite par la diminution du taux d'activité grâce à la segmentation de la ligne ML.

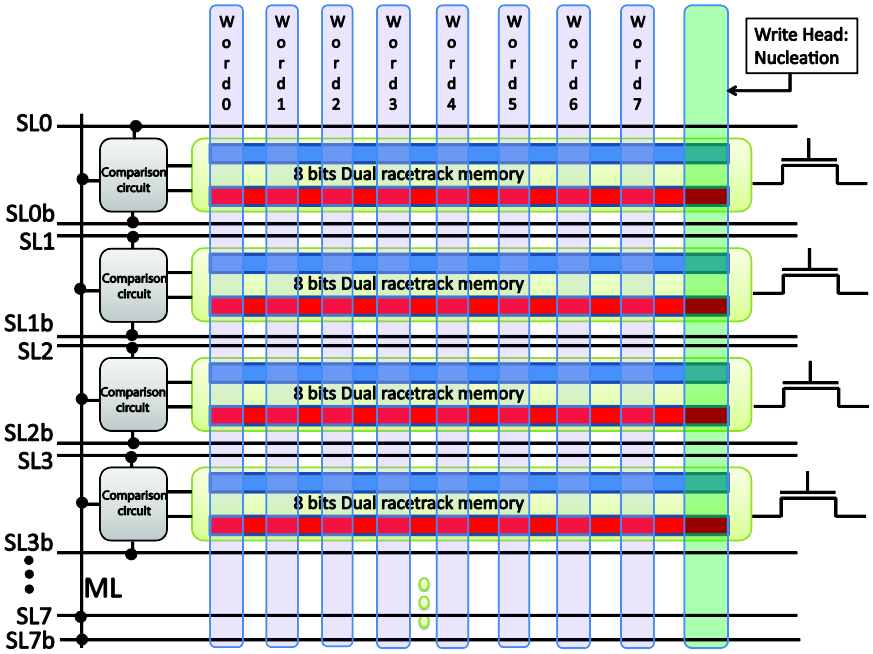
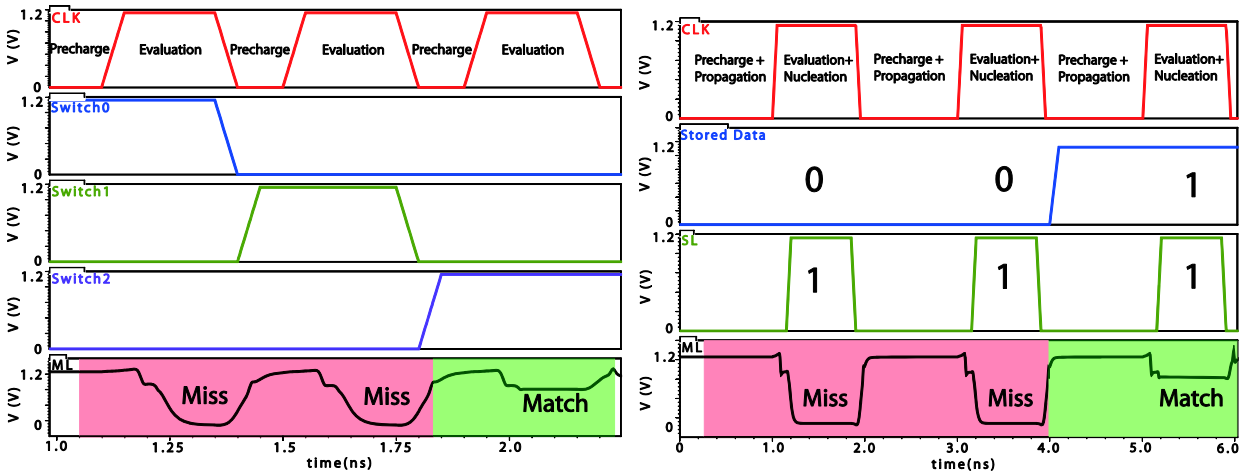


Figure R.23 Schéma de 8 x 8 bits CAM proposée.

Dans le cas où aucune donnée mémorisée ne correspond au mot recherché, un nouveau mot sera inséré et propagé dans les pistes magnétiques pour la prochaine phase de recherche. La figure R.24(b) montre les résultats de la simulation transitoire dans le pire cas : l'échec de 1 bit. Cela veut dire que les 7 autres bits du mot recherché correspondent aux données de stockage, un seul bit est différent. Comme le montre la figure R.24(b), le bit de recherche est '1', si aucune correspondance n'est trouvée, l'impulsion de courant de propagation va commencer à entrainer les parois de domaine, jusqu'à ce que la correspondance se produise. Nous montrons que l'ensemble de l'opération, consistant en « pré-charge », « propagation » et « évaluation », ne prend que ~2 ns.



(a) (b)

Figure R.24 Simulations transitoires de CAM proposée : (a) sans la propagation de parois de domaine ; (b) avec la propagation de parois de doamine.

R.5 Optimisation de conception de la STT-MRAM et la mémoire racetrack

Bien que la STT-MRAM et la mémoire racetrack aient démontré de nombreux avantages, l'objectif de haute densité reste confronté à plusieurs défis. Dans ce chapitre, nous allons proposer deux optimisations de conception pour la STT-MRAM et pour la mémoire racetrack : la cellule multi niveaux (MLC) et la mémoire racetrack assistée par champ magnétique.

R.5.1 Cellule à multi niveaux (MLC)

La cellule multi niveaux (MLC) constitue un concept nouveau par rapport à une cellule binaire. Elle est capable de stocker plusieurs bits de données dans une seule cellule. Par conséquence, il s'agit d'une technologie prometteuse pour améliorer la densité des MRAMs et mettre en œuvre le calcul neuromorphique.

Du point de vu de leur structure, il y a deux types de MLC : série et parallèle (illustré par la figure R.25). La MLC utilise généralement un transistor de sélection pour permettre les processus de commutation, ce qui est similaire à la structure 1T/1MTJ utilisée dans la MRAM. D'ailleurs,

le nombre de niveaux est déterminé par le nombre de MTJs, par exemple, N MTJs dans la MLC permettent d'obtenir N+1 niveaux, correspondant à $\log_2(N+1)$ bits de données.

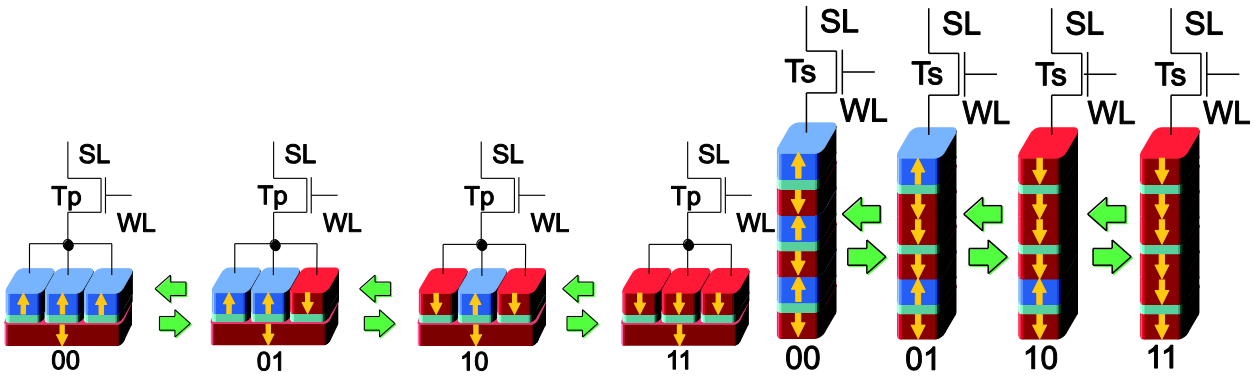


Figure R.25 Structures de 2-bit MLC : (a) série ; (b) parallèle.

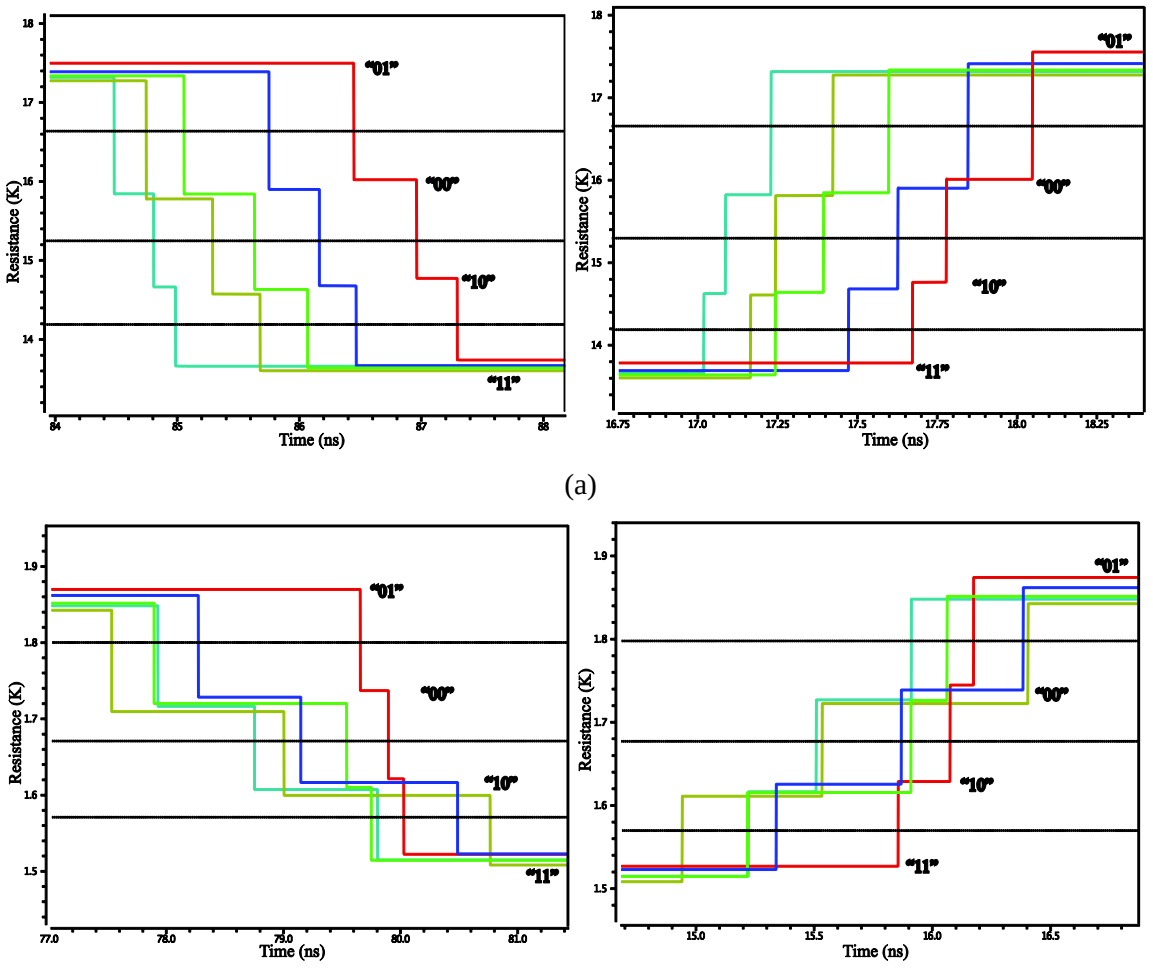


Figure R.26 Simulations Monte-Carlo de 2-bit MLC : (a) série; (b) parallèle.

Les fonctionnalités de ces deux structures sont validées par leurs simulations de Monte-Carlo. Les variations des paramètres clés de la JTM, tel que le TMR, les épaisseurs de la couche libre et la barrière d'oxyde, sont également considérées. La figure R.26 montre les résultats de simulation, dans lesquels chaque processus a effectivement été exécuté 100 fois sous une impulsion de 10 ms. Cependant, nous ne montrons que 5 d'entre eux pour des raisons de lisibilité. Nous pouvons constater que les états de MLC varient aléatoirement par palier, ce qui correspond à la fonctionnalité que nous avions prévue. A noter que ces simulations ont été faites dans le régime de précession (modèle de Néel-Brown). Lorsqu'on s'attend à accélérer la vitesse de fonctionnement, un courant plus élevé peut être appliqué pour forcer la MLC à fonctionner dans le régime assisté thermiquement (modèle de Sun). A cette condition, la durée de commutation peut atteindre la nanoseconde. Toutefois la performance multi-niveau se dégrade en conséquence. En comparant ces deux résultats de simulation, la différence entre deux niveaux de résistance de la MLC en série est plus grande que celle en parallèle, ce qui prouve l'avantage de la structure en série pour la lecture.

R.5.2 Mémoire racetrack assistée par champs magnétique

Etant donné que le transfert de données et sa vitesse dépendent fortement du courant de propagation des parois de domaine, une source de courant est nécessaire. Elle devrait fournir un courant suffisamment élevé et stable pour assurer la propagation des parois de domaine et maintenir une vitesse constante. Cependant, en raison de la haute résistivité des matériaux et de la limitation des circuits CMOS, cette contrainte sur le courant devient le principal obstacle pour la réalisation d'une grande capacité de la mémoire racetrack.

Les progrès récents montrent que les parois de domaine peuvent être propagées en dessous du courant critique en présence d'un champ magnétique externe. Ce phénomène contraire à l'intuition est attribué à l'effet « Walker breakdown ». Cette découverte ouvre un chemin pour atteindre l'objectif consistant à réduire le courant critique et encourage l'utilisation du champ magnétique dans la conception de la mémoire racetrack afin d'améliorer la capacité de stockage et la faisabilité.

Figure R.27 illustre la structure proposée de la mémoire racetrack assistée par champ magnétique. Le courant traversant un fil sur les niveaux de métaux supérieurs génère le champ magnétique.

Une paire de mémoires racetracks partagent un fil métallique pour économiser la surface. Comme la distance entre le fil métallique et la piste magnétique est constante et la taille de la piste est relativement petite, le champ magnétique local peut être considéré comme un champ global pour chaque piste.

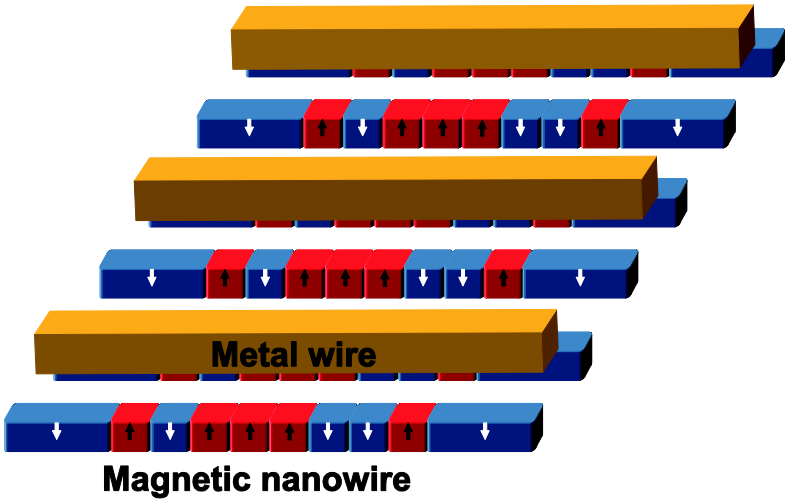


Figure R.27 Structure de l'AMP mémoire racetrack assisté par champ magnétique.

La génération de champ magnétique est un élément crucial pour cette conception. Il existe de nombreuses méthodes pour générer le champ magnétique, par exemple des bobines, l'aimant permanent et le courant passant à travers un fil métallique. Cependant, les bobines coûtent plus cher en consommation d'énergie et en surface ; l'aimant permanent est difficile à intégrer avec une technologie de pointe et à mettre en œuvre au niveau du nanomètre. Comme la mémoire racetrack est un système linéaire, l'installation des fils est plus facile pour la réalisation et la miniaturisation. Néanmoins, elle reste soumise à certaines contraintes : d'une part, puisqu'un courant relativement élevé est nécessaire, il faut utiliser des métaux épais pour éviter les problèmes d'électromigration ; d'autre part, la génération de champ magnétique cohérent avec l'orientation de l'aimantation nécessite un intervalle constant entre le fil métallique et la piste magnétique. La figure R.28 présente la vue en coupe de la structure de la mémoire racetrack assisté par champ magnétique. On peut déposer une couche d'isolant (e.g. MgO) pour séparer les pistes magnétiques et métalliques.

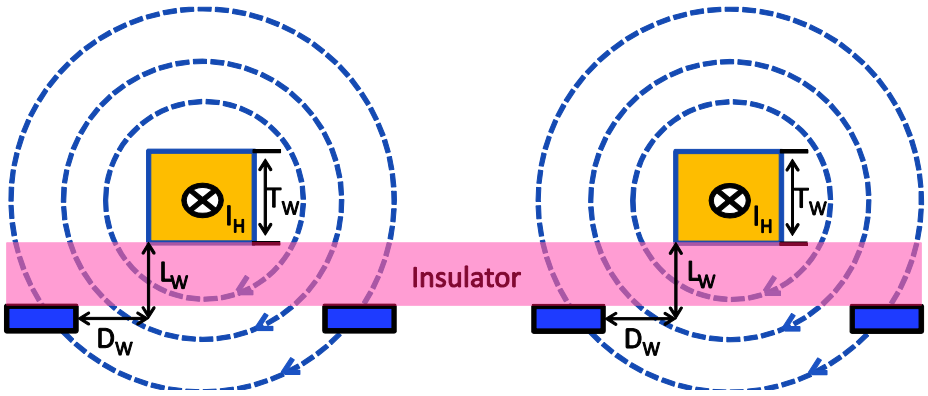


Figure R.28 Vue en coupe de la structure de l'AMP mémoire racetrack à l'aide de champ magnétique.

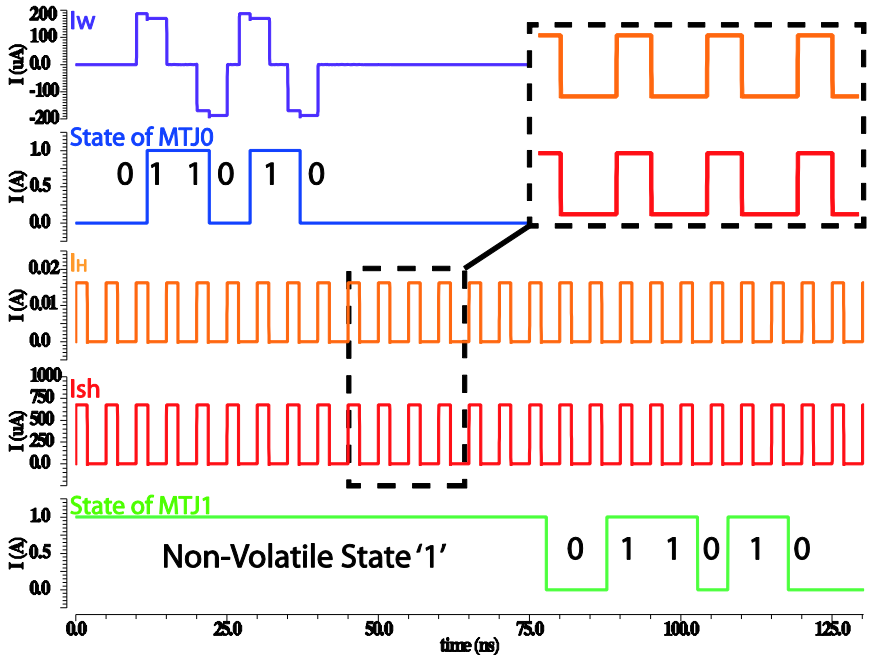


Figure R.29 Simulation transitoire de mémoire racetrack 16-bit à AMP à l'aide de champ magnétique.

La figure R.29 présente les résultats de simulation transitoire d'une mémoire racetrack 16-bit de APM matériau Co/Ni assisté par champ magnétique. Dans cette simulation, nous utilisons un courant de 20 mA pour générer un champ magnétique de 19 mT. Le courant critique dans ce cas-là est 3.1 x 10^{11} A/m^2 par rapport au courant critique intrinsèque de 4.7 x 10^{11} A/m^2. Comme le montrent les résultats de simulation, lorsqu'on injecte un courant supérieur au seuil, la tête de lecture lit les mêmes données émises par la tête d'écriture après 16 impulsions de courant. La figure R.30 montre la relation entre la capacité de stockage et la consommation d'énergie pour différentes intensités du champ magnétique. Nous constatons qu'il y a un compromis lors de l'amélioration de la capacité de stockage : plus de bits stockés dans la mémoire racetrack exigent

un champ élevé, plus la consommation d'énergie est importante. D'autre part, bien que la consommation d'énergie pour générer le champ magnétique soit beaucoup plus élevée que celle nécessaire à la propagation des parois de domaine, elle est toujours limitée à une mesure compréhensible (de l'ordre de pJ/bit), ce qui bénéficie à la mise en œuvre de cette conception. Selon les différentes contraintes du contexte applicatif (e.g. faible énergie ou haute capacité), nous pouvons décider si le champ magnétique doit être utilisé ou non.

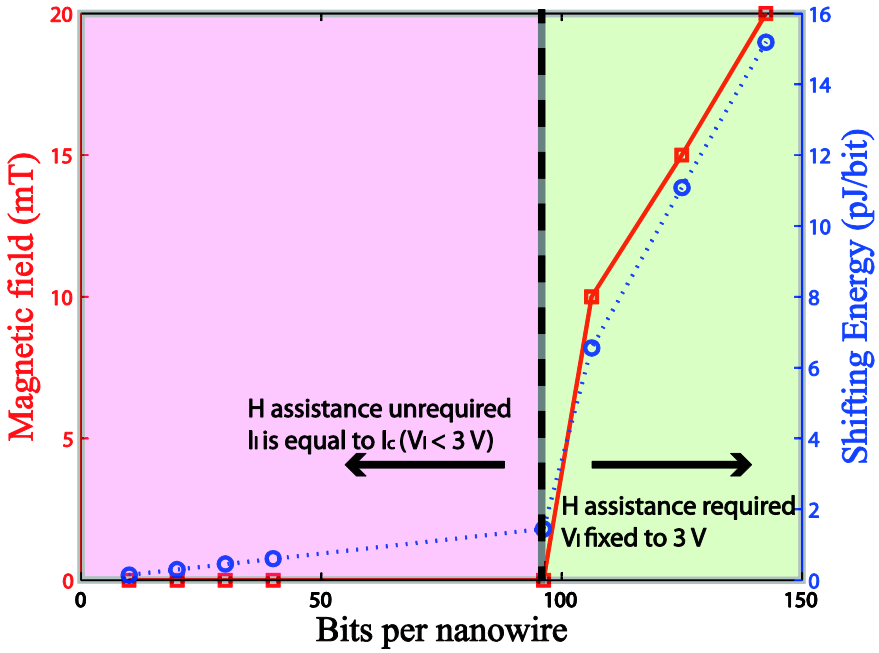


Figure R.30 Dependance de champ magnétique et énergie de propagation versus nombre de bits par piste dans la mémoire racetrack.

R.6 Conclusion

Cette thèse s'est concentrée sur les dispositifs et les circuits logiques et de la mémoire hybride basé sur la spintronique, en particulier concernant deux concepts utilisant la commutation par un courant : la STT JTM à AMP et la mémoire racetrack à AMP. Les travaux comprennent en la modélisation compacte, la conception des circuits et l'optimisation de la conception. En comparant avec les technologies conventionnelles, les applications spintroniques se révèlent être une alternative prometteuse pour surmonter la limitation d'énergie imposée à la miniaturisation des circuits CMOS. Par ailleurs, grâce à la compatibilité avec la technologie CMOS et

l'intégration 3D, ils ont également démontré leurs avantages en termes de vitesse et de densité. En outre, en prenant en compte le comportement stochastique de STT, la fiabilité peut s'étudier au travers de simulations de Monte-Carlo.

Afin de mieux comprendre les principes relatifs et le contexte de la spintronique impliquées dans cette thèse, l'état de l'art de la recherche actuelle a été présenté. La JTM et les différentes approches de commutation pour JTM ont été introduites. Parmi eux, la commutation d'aimantation induite par le courant, soit STT, attire une attention considérable en raison de ses excellentes performances de consommation d'énergie, de densité et de vitesse. D'autre part, la propagation de parois de domaine dans les nano-pistes magnétiques est un phénomène essentiel pour le stockage et la manipulation d'information. La propagation par courant montre ses avantages de nombreuses façons par rapport à celle induite par champ magnétique. Sur la base de ces dispositifs spintroniques, certains circuits hybrides ont été démontrés. Par exemple, la MRAM est considérée comme un excellent candidat pour la prochaine génération de mémoire universelle ; la mémoire racetrack se positionne pour atteindre une très haute densité et une haute fréquence ; la logic-in-memory est une architecture potentielle pour remplacer celle classique de Von Neumann. A noter que ces dispositifs et concepts spintroniques à base de matériaux APM peuvent fournir une stabilité thermique plus élevée pour de petite taille, ce qui domine la tendance des recherches en cours.

Les modèles compacts de la STT JTM à AMP et la mémoire racetrack à AMP ont d'abord été développés. Afin d'obtenir une bonne précision de la performance et un bon accord avec les résultats expérimentaux, la modélisation compacte s'est appuyée sur de nombreux équations théoriques et paramètres réalistes. Le modèle de résistance d'oxyde et le modèle de TMR dépendant de la tension ont été utilisés pour évaluer la caractéristique de résistance. Les modèles en termes statique, dynamique et stochastique ont été utilisés pour décrire le comportement de commutation par STT. D'autre part, un modèle 1D a été intégré pour caractériser la propagation des parois de domaine dans la mémoire racetrack. Ces modèles compacts ont été programmés avec le langage Verilog-A, qui fournit une interface simple de configuration et permet la compatibilité de SPICE dans de nombreuses plates-formes de conception. En effectuant des simulations DC, transitoire et Monte-Carlo en Cadence, les fonctionnalités de ces modèles compacts ont été validées.

En utilisant les modèles compacts, des circuits logiques et de mémoire hybrides à base de la STT JTM à AMP et la mémoire racetrack à AMP ont été conçus et analysés. L'ACM est un exemple d'applications logiques spintroniques à étudier. L'ACM 1-bit basé sur la STT JTM à AMP et l'ACM multi-bit basé sur la mémoire racetrack ont été respectivement proposés. Par rapport aux additionneurs complets purement CMOS, ils ont montré des avantages en termes de puissance statique et de densité grâce à la non-volatilité et l'intégration 3D. Bien que la fréquence et la puissance ne soient pas très satisfaisantes, nous pensons que, avec la diminution de taille et l'amélioration de structure, ces inconvénients seront atténués. Une CAM basée sur une mémoire racetrack a été conçue, simulée et analysée comme exemple d'applications de mémoire spintroniques. Comme le circuit de comparaison nécessite des JTMs avec les états complémentaires, nous avons proposé une structure à double pistes. Profitant de la propagation de parois de domaine sous courant et du partage des circuits périphériques, cette conception optimise de la surface et la puissance par rapport aux CAMs conventionnelles tout en conservant une haute vitesse de recherche. Cependant, la programmation de nouvelles données dans la CAM coûte encore beaucoup de temps et d'énergie, ce qui devrait être amélioré à l'avenir.

L'un des obstacles les plus sérieux pour une large application de la STT-MRAM et la mémoire racetrack est la faible densité. STT-MRAM à base de MLC a été proposée pour augmenter la densité. Les structures en série et en parallèle ont été simulées et analysées. A noter que cette conception utilise le comportement stochastique des STT pour atteindre une ultra-haute vitesse. D'autre part, en raison de progrès récents, le champ magnétique peut aider à déclencher le mouvement de parois de domaine sous le courant critique. Afin de surmonter le problème de courant critique élevé qui limite la capacité de la mémoire racetrack, la mémoire racetrack assisté par champ magnétique a été proposée. La structure et la mise en œuvre ont été présentées. Sa fonctionnalité a été validée par des simulations mixtes, au travers de laquelle les performances améliorées ont été démontrées. Le compromis entre la consommation d'énergie et la capacité a aussi été analysé. Il montre que, bien que la consommation d'énergie de cette conception reste trop importante à cause de la génération du champ magnétique, celle-ci pourrait se confiner à un degré relativement acceptable (de l'ordre de pJ/bit) en ce qui concerne la grande capacité.