



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

Synthèse en français

De 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
Co-directeur de thèse Weisheng ZHAO

Directeur de recherche, CNRS
Chargé de recherche, CNRS

Composition du jury :

Rapporteur Ian O'CONNOR
Rapporteur Julie GROLLIER

Professeur, École Centrale de Lyon
Chargée de recherche, HDR, Unité Mixte de
Physique CNRS/Thales

Examineur Russell COWBURN
Examineur Lionel TORRES
Invité Jacques-Olivier KLEIN

Professeur, University of Cambridge
Professeur, Université Montpellier 2
Professeur, Université Paris-Sud

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.

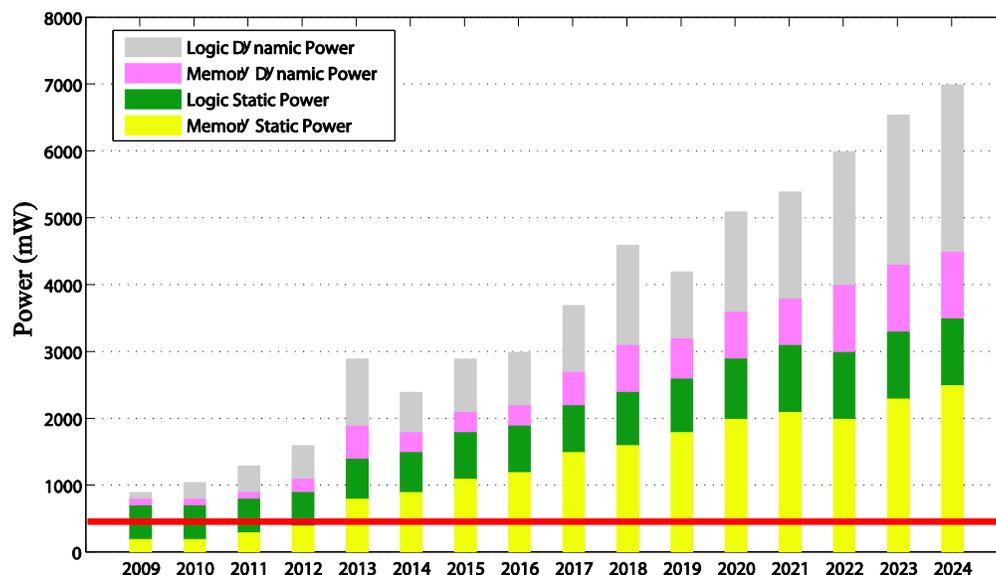


Figure R.1 Prévisions ITRS 2011, jusqu'à l'horizon 2024. (Ligne rouge: exigence de puissances statique et dynamique)

Dans ce contexte, la spintronique ou l'électronique de spin nous constitue un candidat pour résoudre le problème de puissance consommation précédemment 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 consommation 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 utilisant 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 comparée 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.

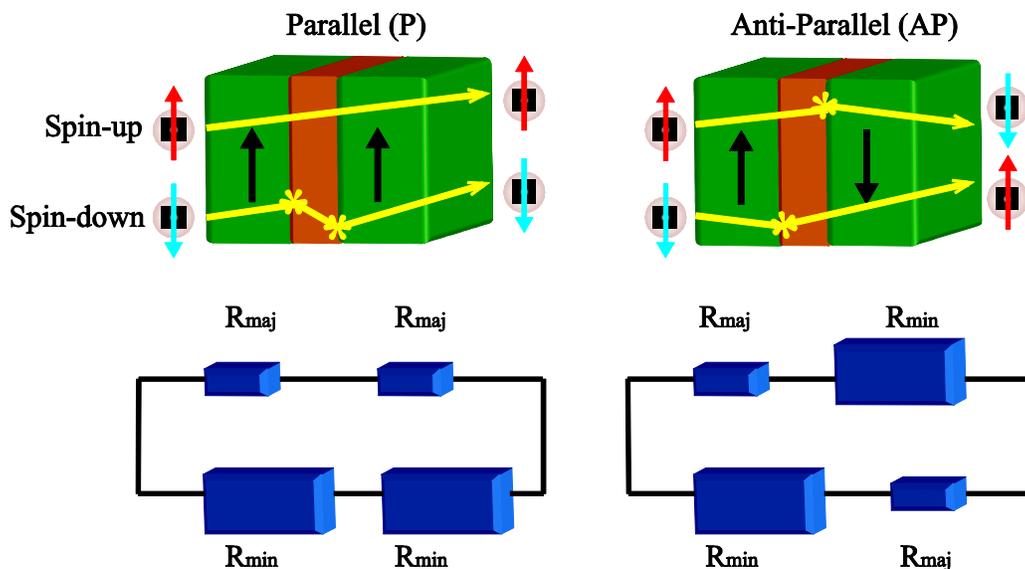


Figure R.2 Effet GMR dans la vanne de spin.

R.2.2 Dispositifs spintroniques

R.2.2.1 Jonction tunnel magnétique (JTM)

R.2.2.1.1 Structure de JTM

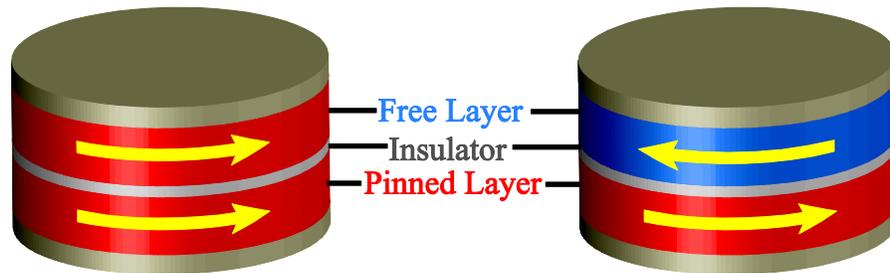


Figure R.3 Structure de la JTM

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.

$$TMR = \frac{\Delta R}{R_P} = \frac{R_{AP} - R_P}{R_P} = \frac{G_P - G_{AP}}{G_{AP}} \quad (R1)$$

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.

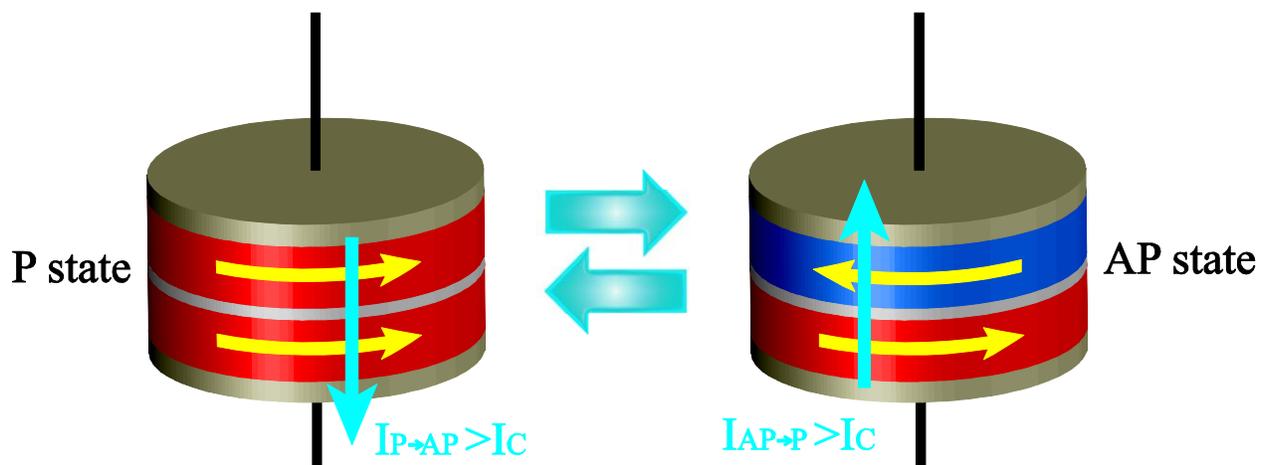


Figure R.4 Commutation de STT

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, compatibilité 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 ($\sim 50 \mu\text{A}$), la stabilité thermique ($40 \text{ k}_B\text{T}$) 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.

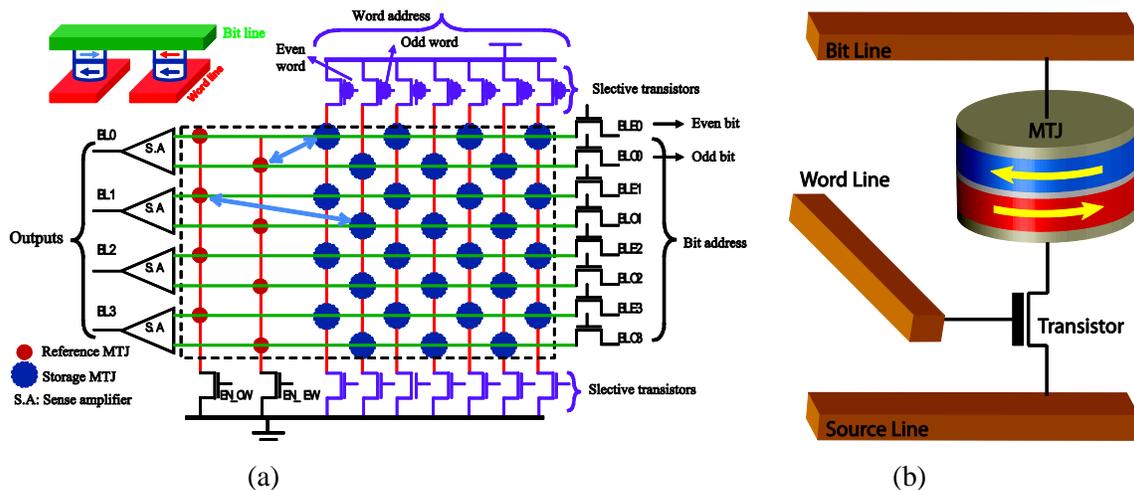


Figure R.5 (a) Schéma de l'architecture cross-point pour la MRAM. (b) Schéma de l'architecture de 1T/1JTM pour la MRAM.

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 entraîne 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 de 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 limitée (256 bit). Dans cette thèse on focalise sur la mémoire racetrack basée sur les matériaux à AMP (e.g. CoFeB/MgO, CoNi).

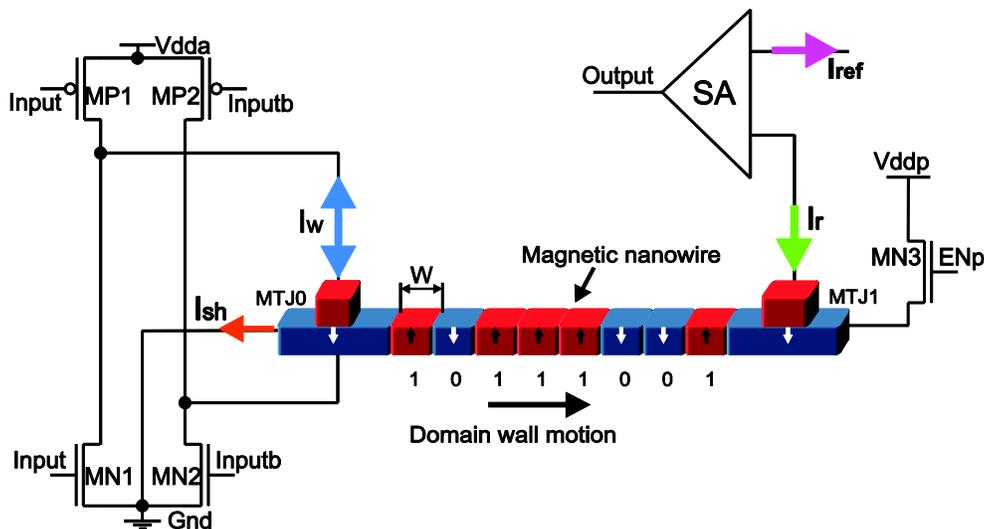


Figure R.6 Mémoire racetrack basée sur la propagation des parois de domaine sous courant.

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.

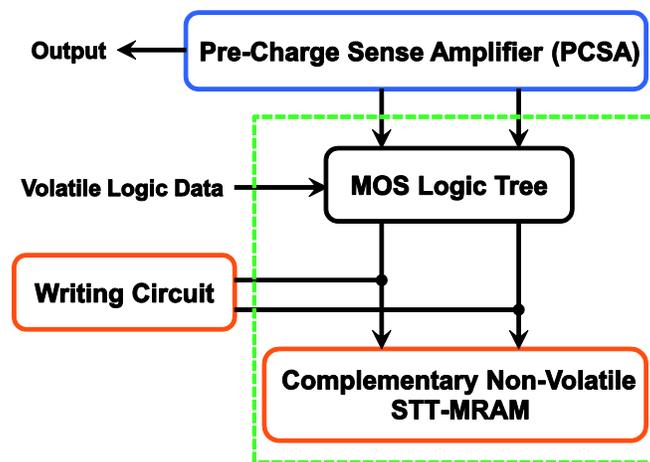


Figure R.7 Architecture générale de logique-en-mémoire à base de STT-MRAM.

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 particulier, 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.

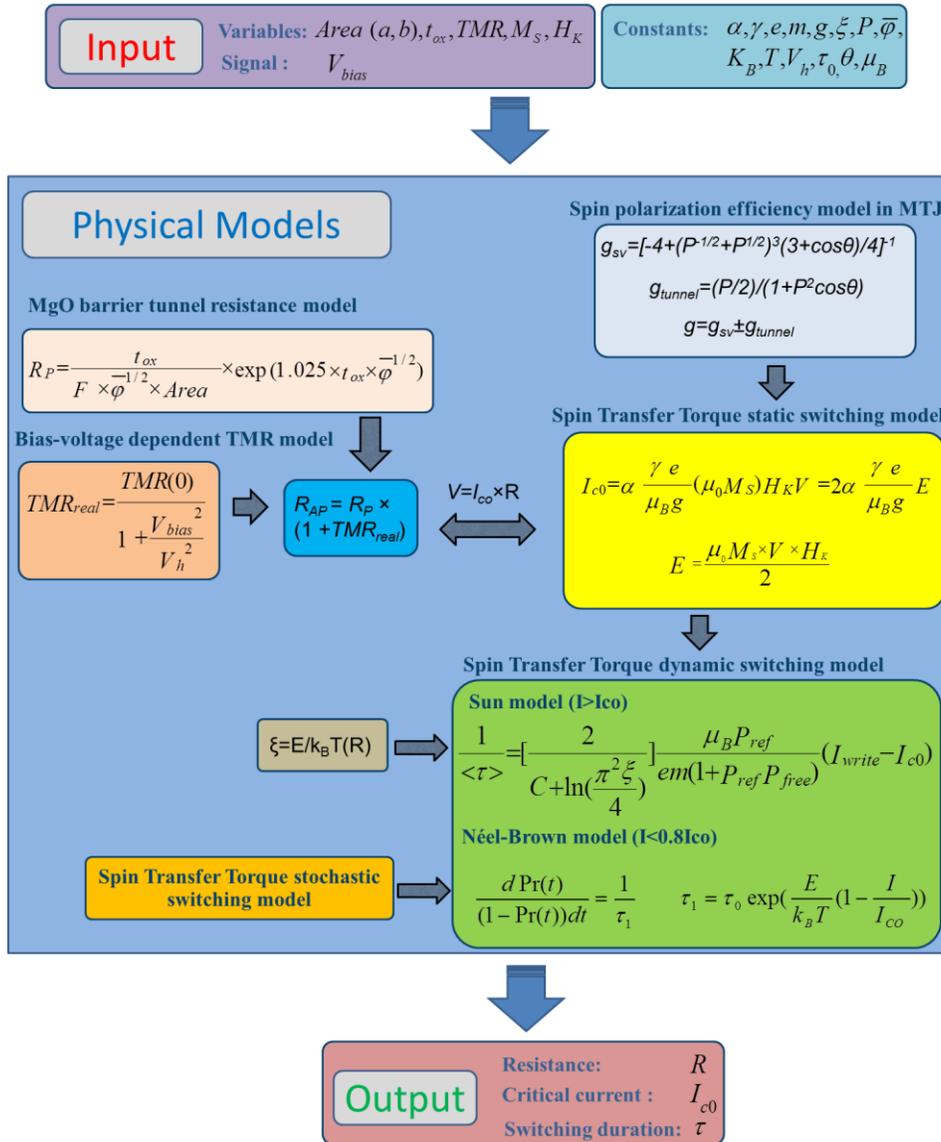


Figure R.8 Hiérarchie d'intégration des modèles physiques dans la STT JTM à AMP.

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 :

$$V_j = u = \frac{\mu_B P j_p}{e M_s} \quad (R2)$$

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

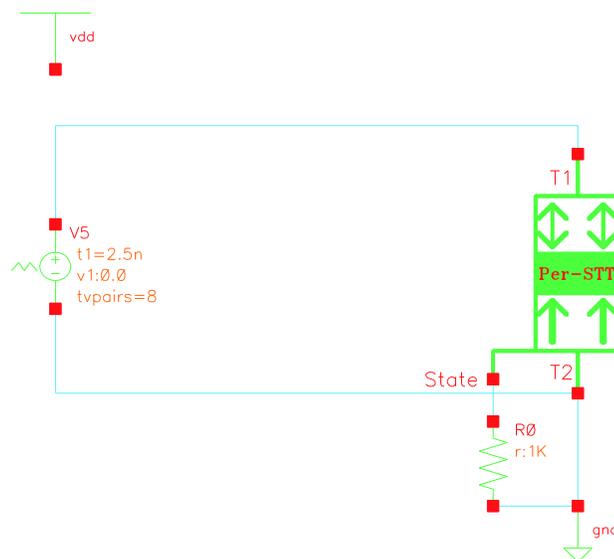


Figure R.9 Schéma de simulation de la STT JTM à AMP.

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.

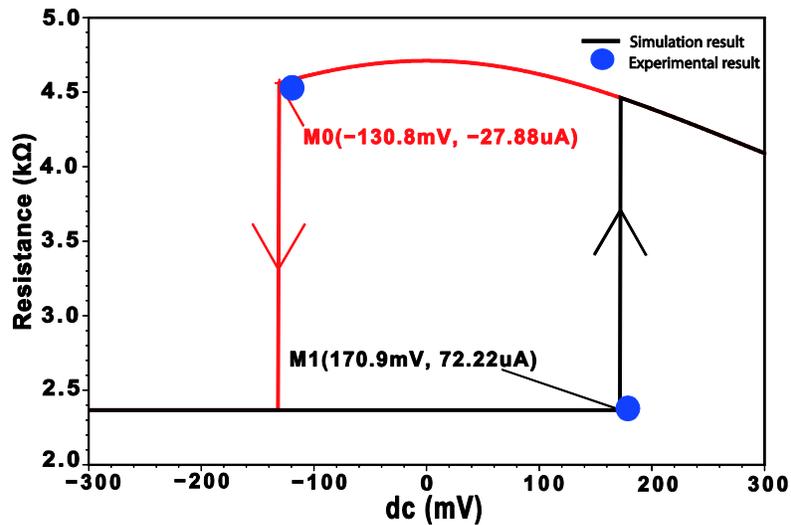


Figure R.10 Simulation DC de la STT JTM à AMP.

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.

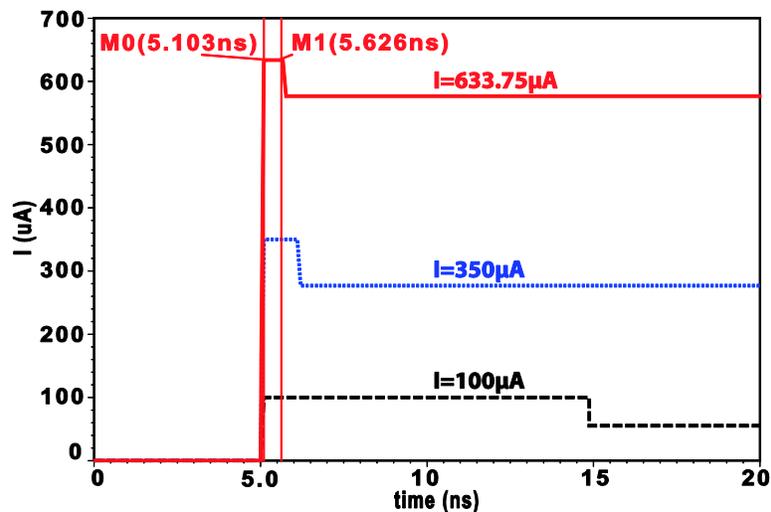


Figure R.11 Simulation transitoire de la STT JTM AMP.

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.

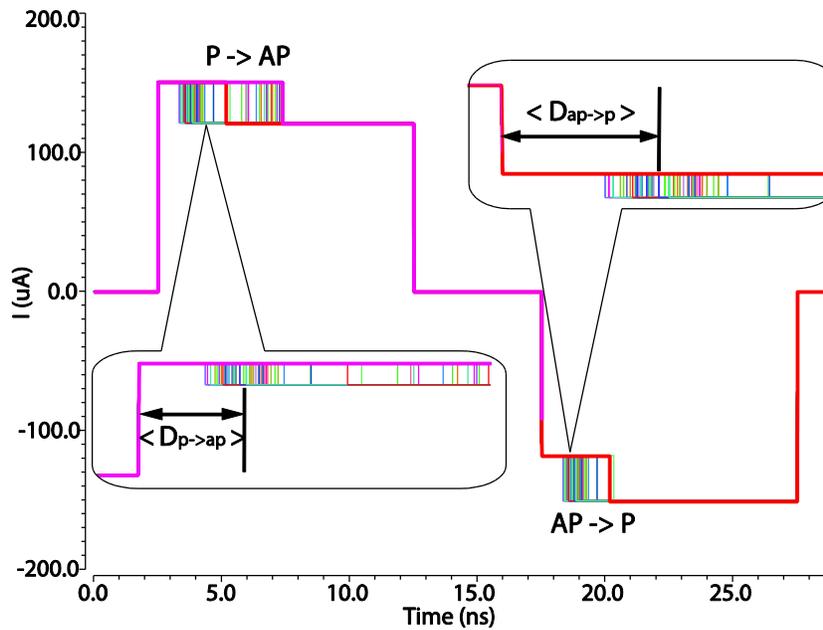


Figure R.12 Simulation Monte-Carlo de la STT JTM à AMP.

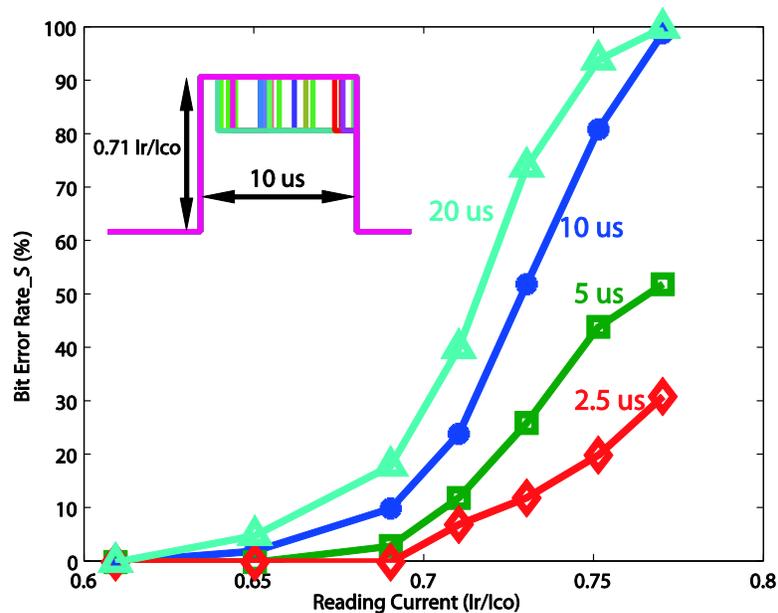


Figure R.13 Dépendance de taux d'erreur de lecture versus current de lecture pour différentes durées d'impulsion.

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.

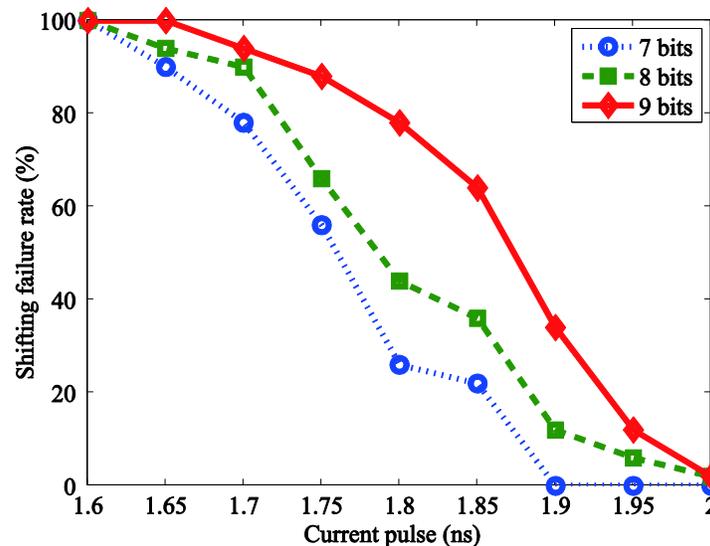


Figure R.16 Dependance de taux d'erreur de propagation versus durées de courant de propagation et capacité de mémoire racetrack.

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 « C₀ ». Parmi celles, l'entrée « B » se rapporte à une mémoire non-volatile.

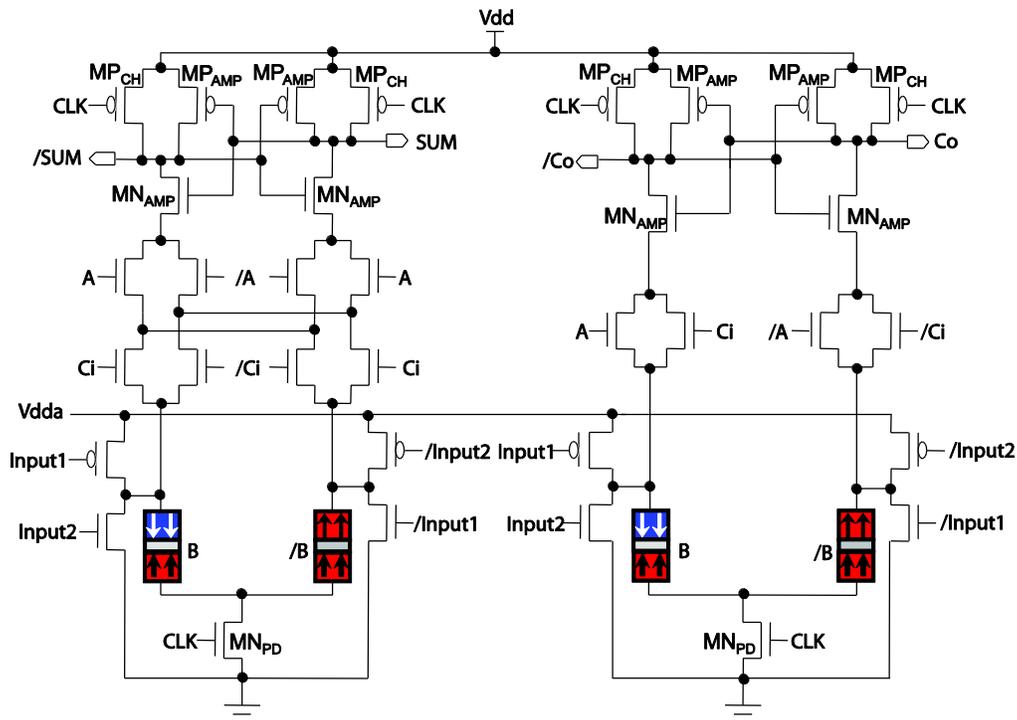


Figure R.17 Architecture de 1-bit STT ACM.

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 consommée 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.

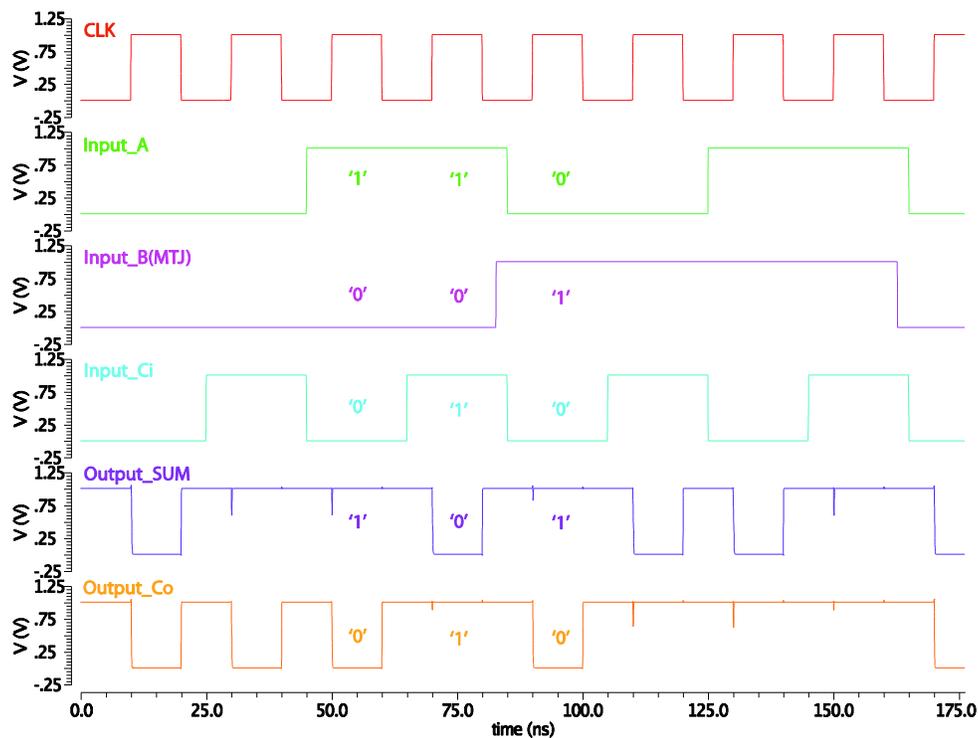


Figure R.18 Simulation transitoire de 1-bit STT-ACM.

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
<i>Delay time</i>	75 ps	87.4 ps
<i>Dynamic power @500 MHz</i>	2.17 μ W	1.98 μ W
<i>Standby power</i>	71 nW	<1 nW
<i>Data transfer energy</i>	>1 pJ/bit	<1 fJ/bit
<i>Die area</i>	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.

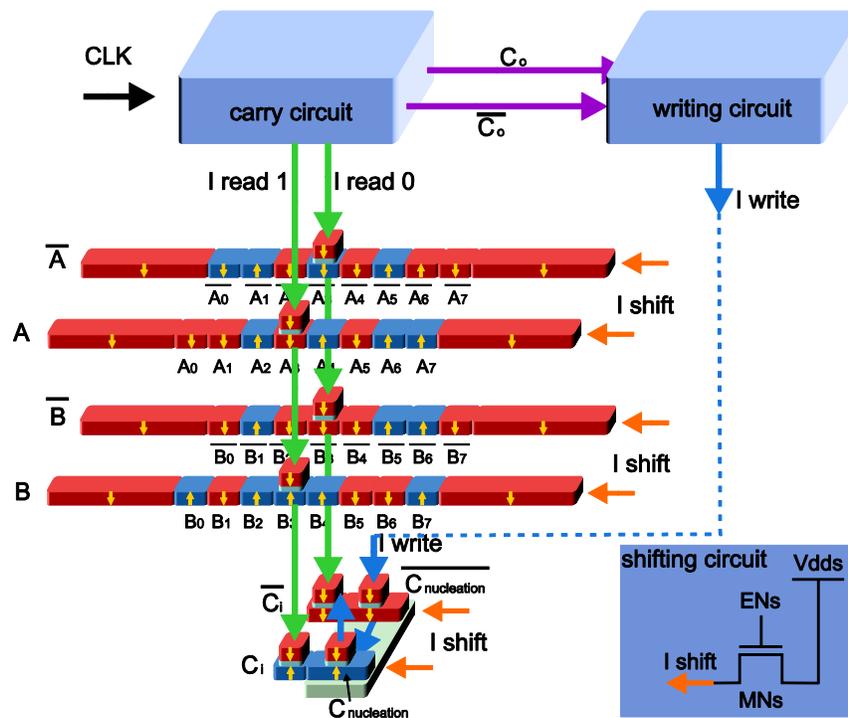


Figure R.19 Circuit de retenue de multi-bit ACM base sur la mémoire racetrack à AMP.

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 « C₀ »= « 01110010 » confirment le bon fonctionnement.

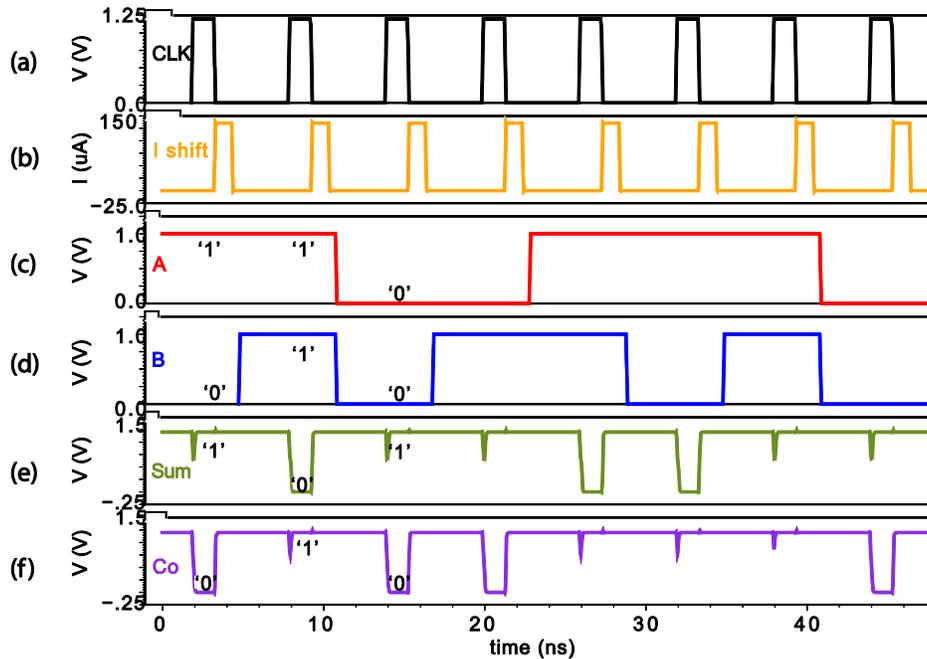


Figure R.20 Simulation transitoire du 8-bit ACM proposé.

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 élevée 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 conséquent, 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.

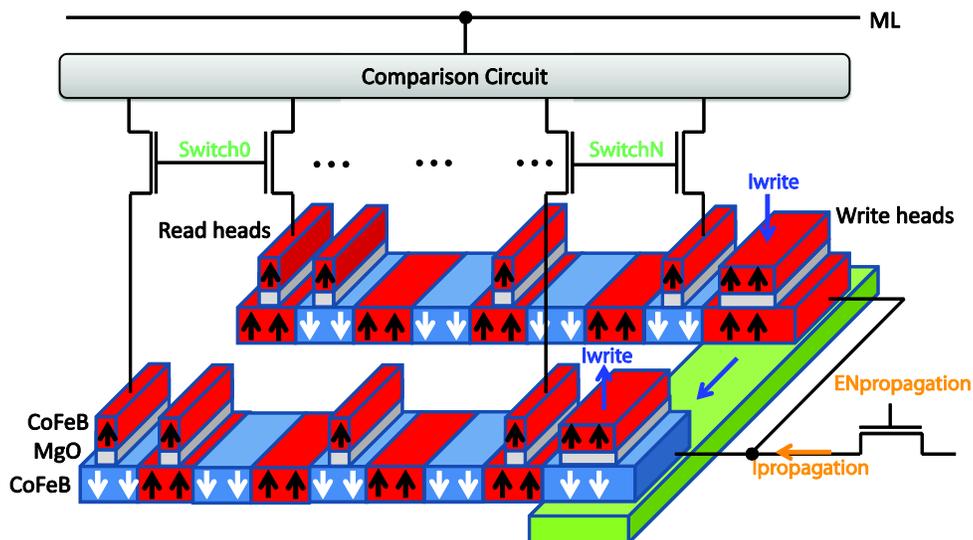


Figure R.21 Structure de CAM basée sur l'AMP mémoire racetrack.

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.

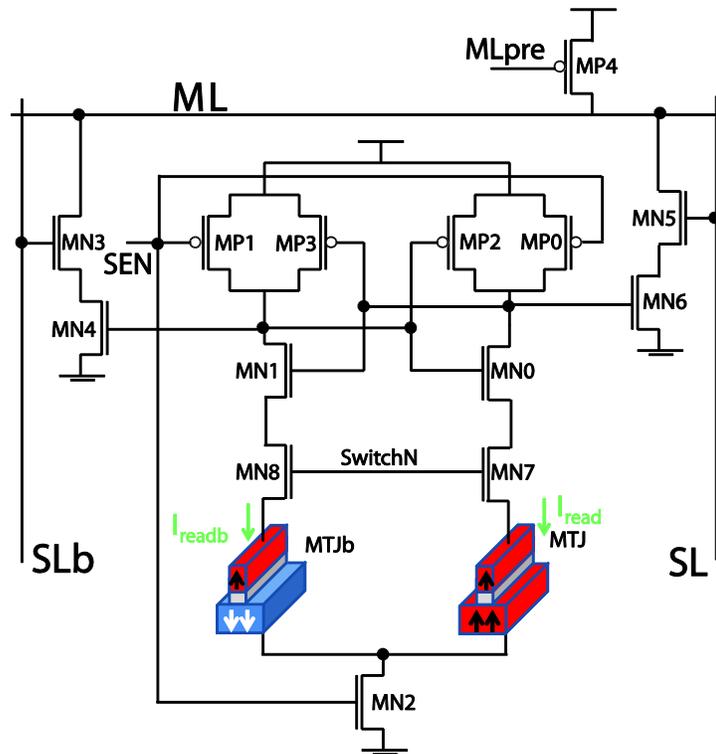


Figure R.22 Schéma de circuit de comparaison.

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.

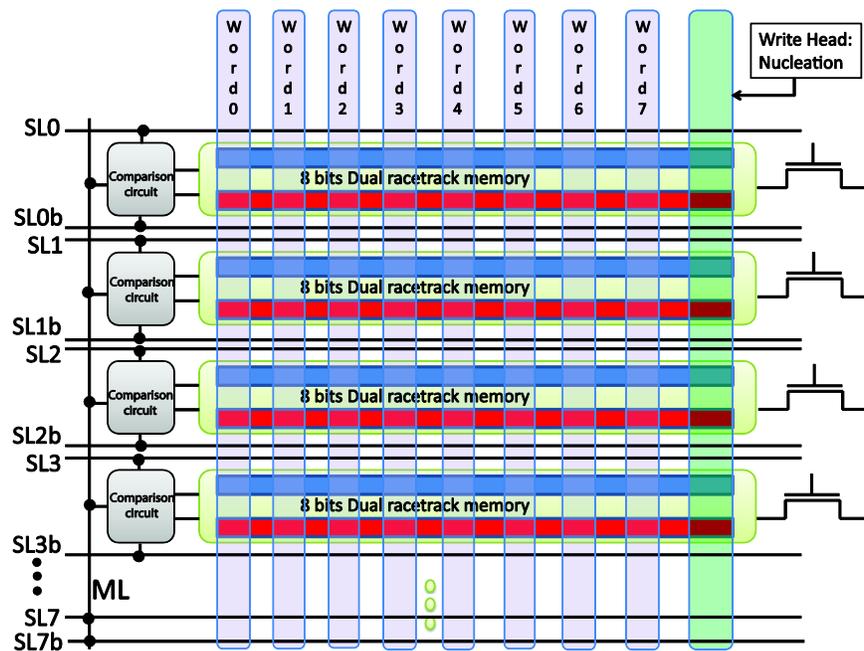


Figure R.23 Schéma de 8 x 8 bits CAM proposée.

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 entraîner 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.

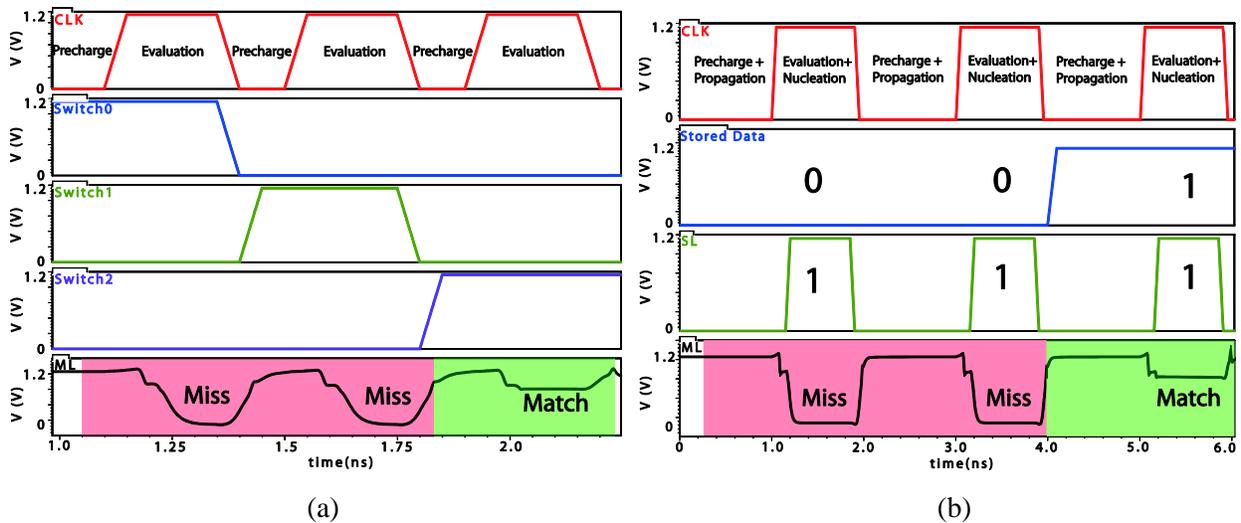


Figure R.24 Simulations transitoires de CAM proposée : (a) sans la propagation de parois de domaine ; (b) avec la propagation de parois de domaine.

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 conséquent, il s'agit d'une technologie prometteuse pour améliorer la densité des MRAMs et mettre en œuvre le calcul neuromorphique.

Du point de vue 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.

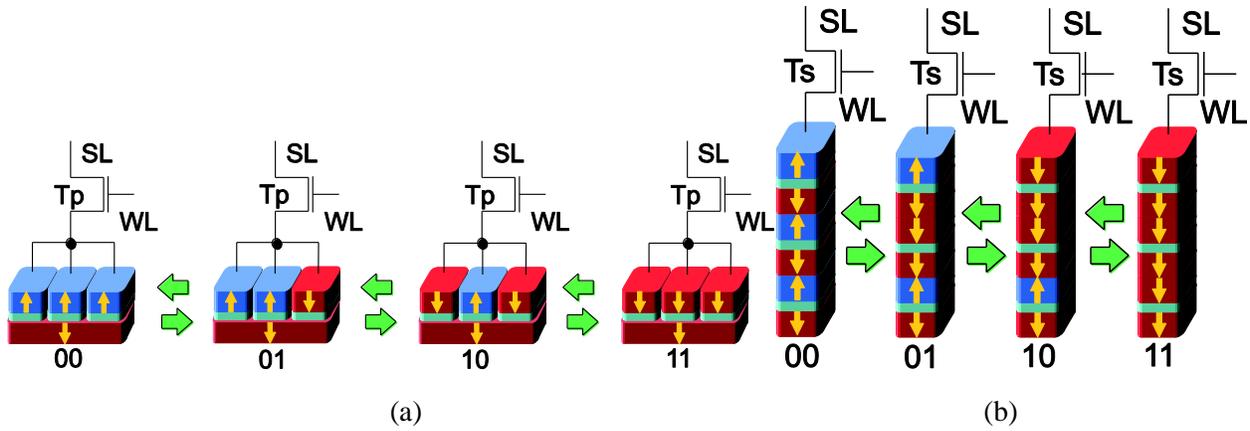


Figure R.25 Structures de 2-bit MLC : (a) série ; (b) parallèle.

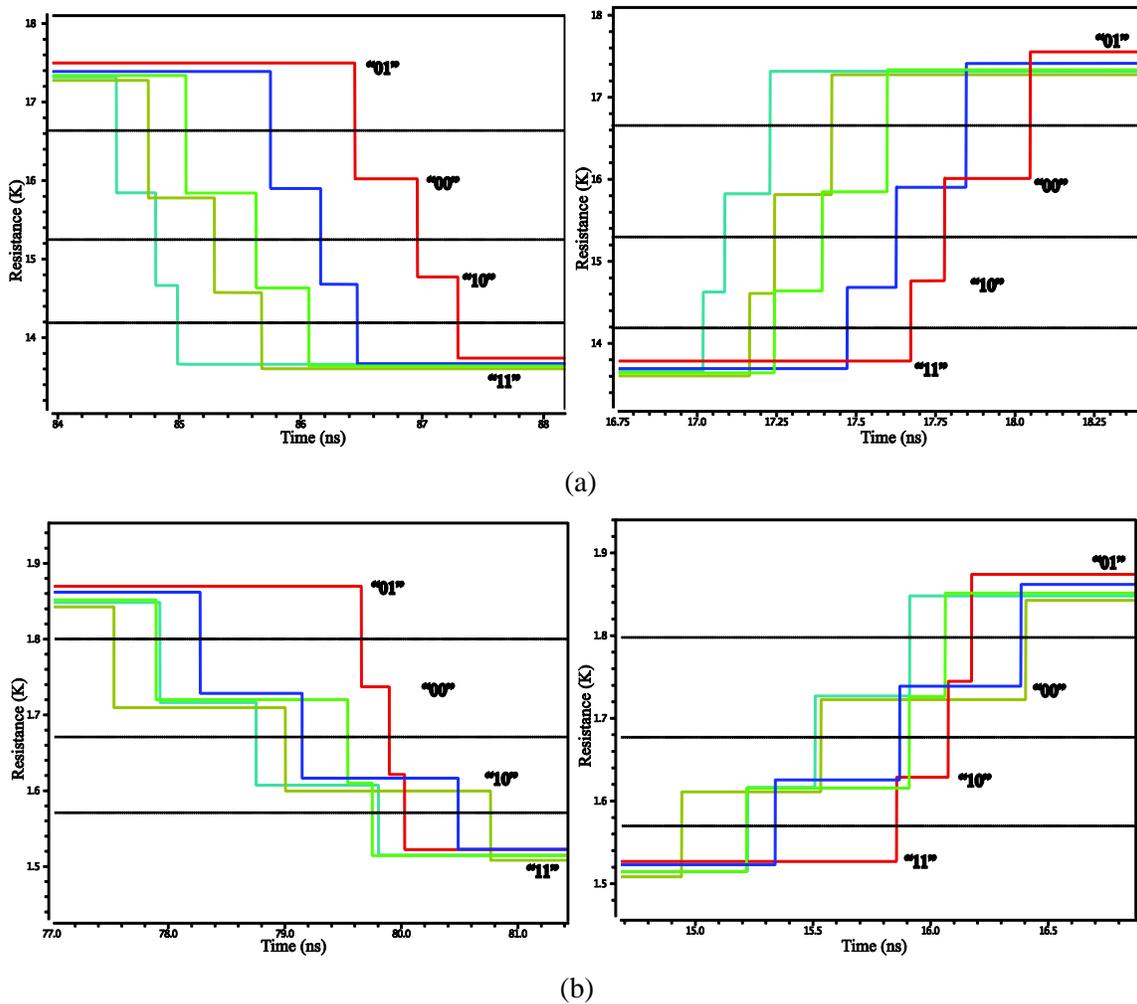


Figure R.26 Simulations Monte-Carlo de 2-bit MLC : (a) série; (b) parallèle.

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 avons 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.

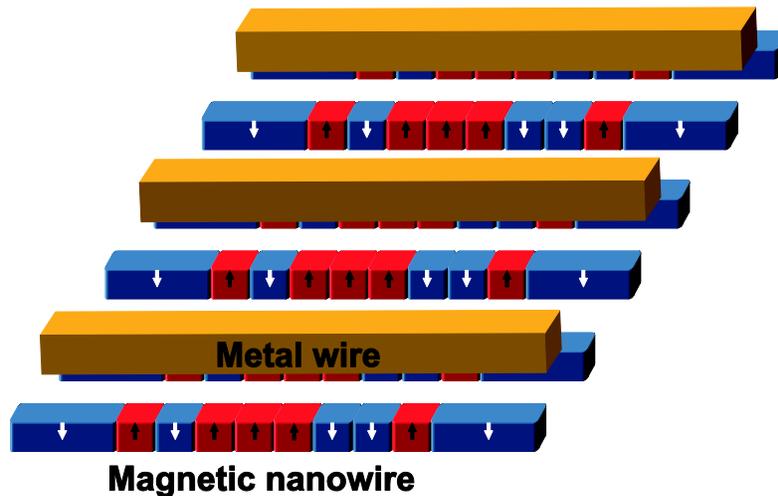


Figure R.27 Structure de l'AMP mémoire racetrack assisté par champ magnétique.

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.

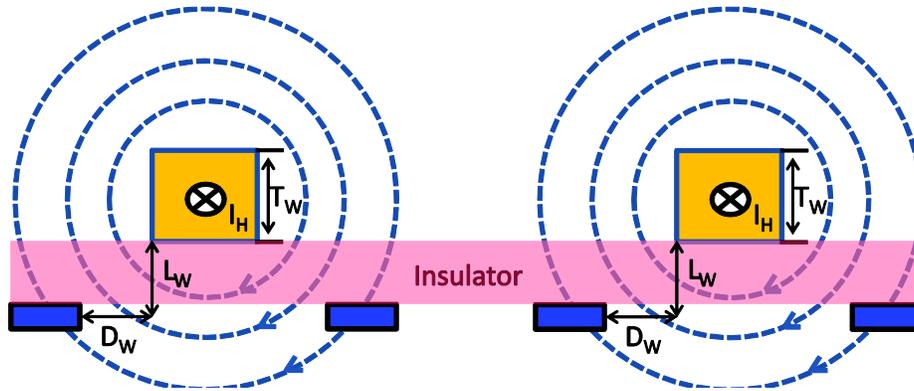


Figure R.28 Vue en coupe de la structure de l'AMP mémoire racetrack à l'aide de champ magnétique.

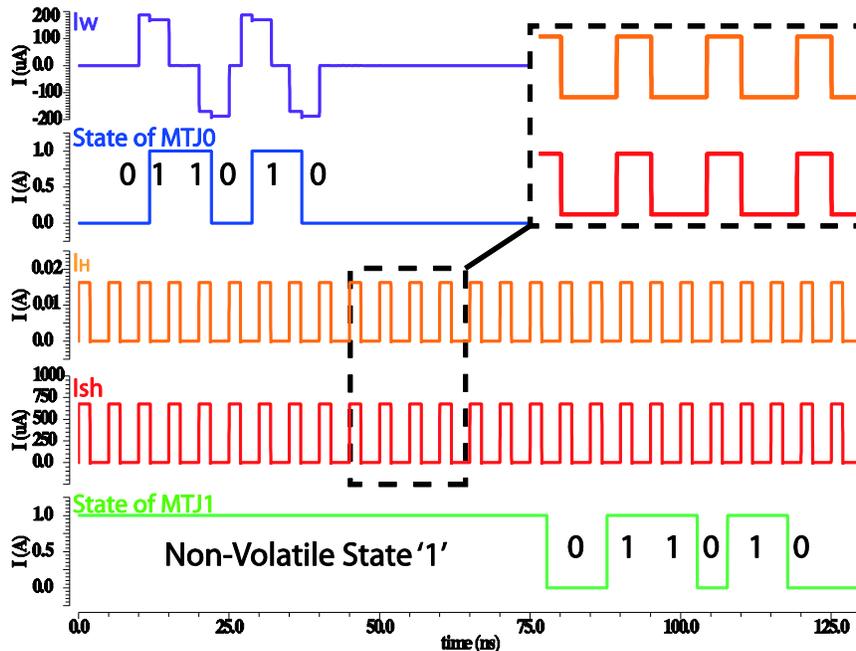


Figure R.29 Simulation transitoire de mémoire racetrack 16-bit à AMP à l'aide de champ magnétique.

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 $3.1 \times 10^{11} \text{ A/m}^2$ par rapport au courant critique intrinsèque de $4.7 \times 10^{11} \text{ A/m}^2$. 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.

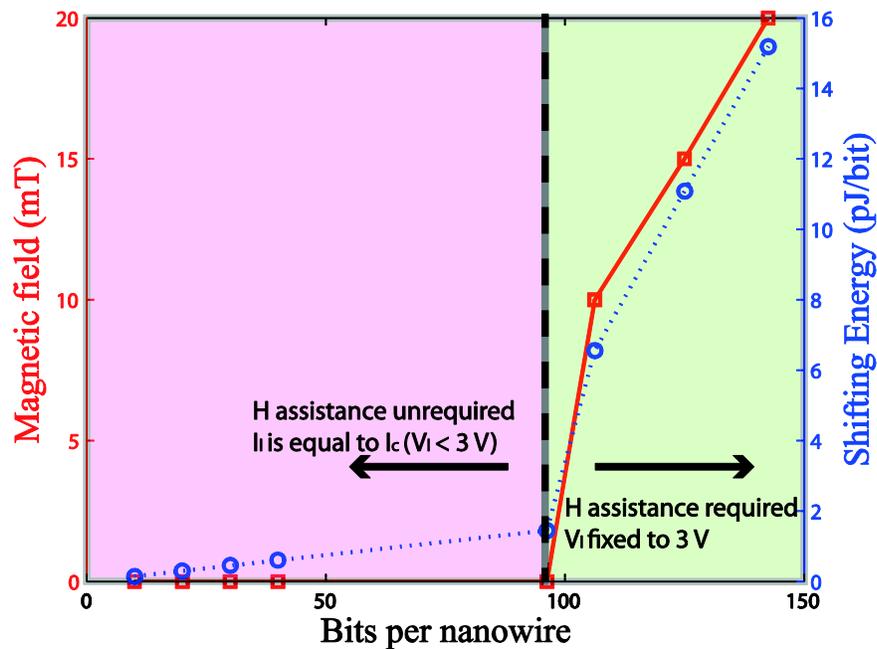


Figure R.30 Dependence of magnetic field and shifting energy versus number of bits per nanowire in the racetrack memory.

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.