Primitives de sécurité à base de mémoires magnétiques

Frédéric Martial Ouattara

THESE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITE DE MONTPELLIER

En SyAM - Systèmes Automatiques et Micro-électroniques

École doctorale I2S - Information, Structures et Systèmes

Unité de recherche LIRMM

Primitives de sécurité à base de mémoires magnétiques

Présentée par Frederic Ouattara

Le 06/11/2020

Sous la direction de Lionel TORRES

Devant le jury composé de

Lilian BOSSUET	Professeur à l'Université Jean-Monnet	Rapporteur
Jean-Michel PORTAL	Professeur à l'Université Aix-Marseille	Rapporteur
Ian O'CONNOR	Professeur à l'Ecole Centrale de Lyon	Examinateur
Bertrand CAMBOU	Professeur à Northern University of Arizona	Examinateur
Pascal NOUET	Professeur à l'Université de Montpellier	Président
Ken MACKAY	Ingénieur à CROCUS Technology	Invité
Lionel TORRES	Professeur à l'Université de Montpellier	Directeur de thèse

UNIVERSITÉ DE MONTPELLIER

Primitives de sécurité à base de mémoires magnétiques

REMERCIEMENTS

Ces travaux de thèse n'auraient pu être réalisés sans le soutien et l'aide de certaines personnes. Je tiens par la présente à remercier toutes les personnes qui ont participé, de façon plus ou moins directe, à cette étape si importante pour moi.

Tout d'abord, j'adresse un immense remerciement à mon directeur de thèse Lionel TORRES. Merci pour la confiance accordée, son dynamisme et sa bonne humeur. Je n'aurai jamais pu aller aussi loin dans ce travail sans sa disponibilité, sa patience, ses conseils et son savoir. Je le remercie encore pour la rigueur scientifique qu'il a su m'insuffler.

Je souhaite ensuite remercier Lilian BOSSUET et Jean-Michel PORTAL qui m'ont fait l'honneur de consacrer du temps pour relire cette thèse et en être les rapporteurs. Je tiens également à remercier Ian O'CONNOR, Bertrand CAMBOU, Pascal NOUET et Ken MACKAY d'avoir accepté d'assister à la soutenance de cette thèse, je les remercie pour leurs commentaires constructifs et leur participation active et fondamentale lors de ma soutenance.

Mes années au LIRMM ont été une formidable expérience grâce à mes collègues et au personnel administratif que je tiens à remercier. Je remercie particulièrement Caroline Lebrun, secrétaire du département microélectronique pour sa disponibilité et son aide si précieuse. Merci également à mes collègues doctorant du LIRMM pour leur sympathie et leur accueil chaleureux. Ils ont participé à ce que l'ambiance de travail soit très agréable à travers nos pauses café et les différentes sorties organisées.

J'exprime ma gratitude à Arash, pour sa disponibilité, son implication dans mes travaux de thèse, à travers nos moments de collaboration.

Mes remerciements vont aussi à ma famille notamment mes parents, mes frères, mes cousins. Je les remercie pour leur soutien, leur encouragement qui ont été sans faille durant ces années de thèse.

Enfin je remercie ma tendre Sahadatou pour sa patience et son soutien inconditionnel. Elle est pour beaucoup dans l'achèvement de ce travail de thèse.

Table des matières

Table des matières 5
Table des figures 9
Liste des tableaux 13
Glossaire 15
Chapitre I Introduction Générale 17
1. Contexte 17
2. Objectif de la thèse 19
Chapitre II Les Mémoires magnétiques à accès aléatoire (MRAM) 21
1. Introduction 21
2. La spintronique 22
2.1. La Magnétorésistance à effet tunnel (TMR) 25
2.2. La Jonction tunnel magnétique (JTM) 27
3. Les technologies MRAM 29
3.1. Mécanisme d'écriture FIMS: Field Induced Magnetic Switching 29
3.2. Mécanisme d'écriture Toggle 30
3.3. Mécanisme d'écriture TAS: Thermally Assisted Switching 32
3.4. Mécanisme d'écriture STT : Spin Transfert Torque 38
3.5. Mécanisme d'écriture VCMA : Voltage-Controlled Magnetic Anisotropy 40
3.6. Mécanisme d'écriture SOT : Spin Orbit Torque 42
4. Conclusion 43
Chapitre III Générateur de nombres véritablement aléatoires à base de STT-MRAM 45
1. Introduction 45
2. Les générateurs de nombres pseudo aléatoires (PRNG) 46
2.1. Principe et fonctionnement des PRNG 47
2.2. Conception de PRNG 47
3. Les générateurs de nombres véritablement aléatoires (TRNG) 51
3.1. Les sources d'aléa et les différentes techniques d'extraction d'aléa dans les circuits électroniques 52
3.2. Méthodes de post-traitement des TRNG 56
3.3. Evaluation des générateurs de nombre aléatoires 58
4. TRNG à base de STT-MRAM 61
4.1. La revue des travaux dans la littérature 62

4.2. Environnement de conception hybride CMOS/STT-MRAM 70
4.3. Les amplificateurs de lecture 74
4.4. Comparaison des différentes cellules de lecture 79
4.5. Le circuit d'écriture 81
4.6. Système complet basé sur le spin dice 83
4.7. Flot de conception du TRNG 85
4.8. Caractérisation et test du démonstrateur 86
5. Conclusion 89
Chapitre IV Générateur de nombre véritablement aléatoire à base de TAS-MRAM 91
1. Introduction 91
2. Principe de fonctionnement 91
3. Outils et méthodes d'extraction d'aléa dans les TAS-MTJ 92
4. Instrumentation et expérimentation 94
4.1. Caractérisation de l'écriture des jonctions TAS-MTJ 94
4.2. Exploration de TRNG sur une jonction TAS 96
4.3. Utilisation des procédures de post-traitement pour améliorer la probabilité de sortie 97
4.4. Amélioration du biais à travers une boucle de contre réaction 99
5. Evaluation statistique des nombres aléatoires générées 101
6. Conclusion 102
Chapitre V Fonctions Non Clonables Physiquement à base de MRAM 105
1. Introduction 105
2. Propriétés et paramètres des PUF 106
2.1. Intra-distance ou la reproductibilité 107
2.2. Inter-distance ou l'unicité 108
3. La Classification des PUF 109
3.1. Les PUF à base de dispositif silicium et non-silicium 109
3.2. PUF fort et PUF faible 116
4. Les méthodes d'évaluation des PUF 116
4.1. Méthode d'évaluation basé sur le calcul de la distance ou du poids de Hamming 116
4.2. Méthode d'évaluation statistiques de la qualité des PUF 118
5. Revues sur les PUF à base de MRAM 119
5.1. Extraction basée sur l'écriture des JTM 119

5.2. Extraction basée sur les variations aléatoires de la résistance des JTM 121
5.3. Synthèse 121
6. Fonction non clonable physiquement basée sur la TAS-MRAM 122
6.1. Principe d'extraction 122
6.2. Instrumentation et Expérimentation 122
7. Conclusion 127
Chapitre VI Conclusion générale et perspectives 129
1. Conclusion générale 129
2. Perspectives 130
Publications et communications 132
Publications 132
Posters 132
Références 133

Table des figures

Figure I-1 : Projection de l'évolution du marché des MRAM [1]. 17

Figure II-1 : Orientation des moments magnétiques selon le type de matériau [6] 22

Figure II-2 : Effet de la magnétorésistance anisotrope (AMR) [7] 23

Figure II-3 : Première observation de la GMR sur des multicouches Fe/Cr à 4,2 K [11] 24

Figure II-4 : Effet de la magnétorésistance géante (GMR) [7] 25

Figure II-5 : Jonction tunnel magnétique [6] 28

Figure II-6 : (a) Distribution des valeurs de résistances min et max d'une mémoire MRAM de 8Kbits (b) Dépendance de la TMR à la tension de lecture pour des JTM à base de MgO et AlOx 29

Figure II-7 : (a) Structure d'une JTM en technologie FIMS (b) Principe d'écriture FIMS 30

Figure II-8 : (a) Structure d'un JTM en technologie Toggle (b) Principe de lecture/écriture Toggle 31

Figure II-9 : Séquence d'écriture Toggle 32

Figure II-10 : Structure d'une JTM en technologie TAS 33

Figure II-11 : Séquence d'écriture d'une JTM TAS : Passage d'un état parallèle à un état antiparallèle 34

Figure II-12 : Structure d'une jonction SR ou MLU 35

Figure II-13 : Unité logique magnétique (a) cellule TAS-MTJ (b) cellule SR-MTJ (c) porte logique virtuel XOR 37

Figure II-14 : Principe du « Match in place » (a) code non valide, (b) code valide 37

Figure II-15 : (a) structure de la STT-JTM (b) Passage d'un état AP => P (c) Passage d'un état P=> AP 40

Figure II-16 : Commutation induite par tension (MeRAM) [38] 41

Figure II-17 : Spin Orbit Torque- MRAM (a) Jonction STT (b) Jonction SOT 43

Figure III-1 : Classification des générateurs de nombres aléatoires 46

Figure III-2 : Combinaison de LFSR par une fonction booléenne 48

Figure III-3 : PRNG filtré à base de LFSR 49

Figure III-4 : Générateur par rétrécissement 50

Figure III-5 : Générateur par auto-rétrécissement 51

Figure III-6 : Structure générique d'un TRNG 51

Figure III-7 : Schéma de principe du TRNG développé par Intel et discuté dans [68] 53

Figure III-8 : Echantillonnage de la gigue d'un signal périodique par un signal d'horloge 54

Figure III-9 : TRNG basé sur l'accumulation de la gigue proposée dans [72] 54

Figure III-10 : TRNG basé sur la métastabilité présenté dans [77] 55

Figure III-11 : Précession de l'aimantation sous l'influence d'un courant de spin : dépendance temporelle de (a) Mz et (b) Mx, (c) processus d'inversion du moment magnétique. θ et ϕ représente l'état initial du moment magnétique de la couche libre [91]. 61

Figure III-12 : Probabilité de commutation d'une JTM-STT en fonction de l'amplitude du courant d'écriture et de la durée de l'impulsion. (Commutation AP -> P) [95] 63

Figure III-13 : La méthode « spin dice ». Les impulsions de réinitialisation et d'excitation sont appliquées de manière séquentielle [93]. 64

Figure III-14 : Méthodes de génération des Rbs (à gauche) la méthode avec réinitialisation de la JTM et (à droite) la méthode sans réinitialisation de la JTM proposé dans [97]. 65

Figure III-15 : TRNG basé sur une jonction STT avec écriture d'excitation conditionnelle, suivi d'un contrôle de la probabilité de sortie en temps réel. 66

Figure III-16 : TRNG basé sur (a) plusieurs JTM en parallèle et (b) sur une paire de JTM en série[104]. 68

Figure III-17 : Circuit mémoire STT-MRAM utilisé comme TRNG à travers un échantillonnage du temps de commutation de la JTM lors de son écriture à courant faible [105]. 69

Figure III-18 : Simulation d'une séquence d'écriture d'une jonction STT 71

Figure III-19 : Probabilité de commutation de la JTM en fonction du courant d'écriture 73

Figure III-20 : Vue layout de la JTM ainsi que les layers magnétiques du DKIT 74

Figure III-21 : Structure de Black & Das avec deux JTM STT et le circuit d'écriture 75

Figure III-22 : Structure de lecture Hass adaptée au circuit d'écriture STT 77

Figure III-23 : Structure de lecture PCSA 78

Figure III-24 : Structure de lecture SRAM Rémanent (RSRAM) (a) avec transistor d'accès PMOS et (b) avec transistor d'accès NMOS..... 79

Figure III-25 : Cellule RSRAMn adaptée à un système d'écriture non complémentaire (référence composée de 4JTM avec le circuit d'initialisation) 81

Figure III-26 : Miroir de courant programmable comme source de courant pour l'écriture de la JTM 82

Figure III-27 : Simulation comportementale de la source de courant programmable 83

Figure III-28 : Circuit d'écriture et de lecture d'une JTM avec la possibilité de faire varier le courant d'écriture a « 1 » de la JTM. 84

Figure III-29 : Algorithme de contrôle et de gestion de la probabilité de sortie à « 1 » de la JTM. Les trois états en rectangle correspondent aux probabilités de sortie calculées après chaque reset. L'état au milieu indique l'état inconnu de la JTM après chaque écriture aléatoire et les 3 états en gris indiquent l'état de la JTM ainsi que la probabilité de sortie actuelle après chaque lecture. 85

Figure III-30 : Layout du circuit et son packaging dans un boitier de type DIL40 86

Figure III-31 : Mesure de la résistance parallèle et de la TMR sur deux wafers contenant notre circuit (résultats de SPINTEC) 87

Figure III-32 : Test d'un circuit validant le fonctionnement de la partie digital du TRNG..... 88

Figure III-33 : Résultats de mesures d'un circuit présentant un phénomène qui semble aléatoire..... 88

Figure IV-1 : (a) 1kbits TAS-MRAM dans un Package QFN44 (b) image microscopique de l'intérieur du QFN44 93

Figure IV-2 : Architecture de la puce composée de 1024 bits TAS-MRAM..... 93

Figure IV-3 : Banc de test comportant une carte FPGA pour la programmation des composants utilisés pour l'écriture de la JTM TAS, une carte d'acquisition pour la lecture et le PCB contenant une puce 1kbit de TAS-MRAM 94

Figure IV-4 : Probabilité de commutation des JTM TAS en fonction de l'amplitude de la tension de chauffage 95

Figure IV-5 : Histogramme de la probabilité de commutation obtenu à partir de chaque 10^4 Rbs sur un total de 10^7 Rbs générés par les jonctions MTJ 1 (rouge) et MTJ 2 (bleu). 97

Figure IV-6 : La probabilité de commutation après un post-traitement de Von Neumann 98

Figure IV-7 : Histogramme de la probabilité de commutation après un post-traitement XOR sur deux JTMs (MTJ1 et MTJ2) 99

Figure IV-8 : Architecture de la boucle de contre réaction 100

Figure IV-9 : Histogramme de la probabilité de commutation après un post-traitement basé sur la RTOPT 100

Figure V-1 : Evaluation de la reproductibilité de la réponse d'une PUF sous différentes températures de fonctionnement 107

Figure V-2 : Evaluation de l'unicité de la réponse des PUF..... 108

Figure V-3 : Principe de base des PUF à base de délai à travers la comparaison de deux temps de propagation de deux chemins identiques..... 110

Figure V-4 : Principe de base de la PUF d'arbitrage..... 111

Figure V-5 : Une PUF d'arbitrage typique avec deux chemins de propagation nominalement identiques suivis par un arbitre pour générer un bit de réponse en fonction du chemin le plus rapide 111

Figure V-6 : Principe de base d'une PUF à base d'oscillateur en anneau 112

Figure V-7 : PUF à base d'oscillateur en anneau présenté dans [133] 113

Figure V-8 : (a) Cellule SRAM typique à 6 transistors et ses caractéristiques en tension des inverseurs rebouclés 114

Figure V-9 : Butterfly-PUF basé sur le rebouclage de deux verrous [135] 115

Figure V-10 : Paysage énergétique dans (a) une JTM à géométrie parfaite, (b) et (c) une JTM avec des variations dans la geometrie de la couche de stockage 120

Figure V-11 : Lecture d'une matrice 32X32 de JTM TAS (a) après écriture à "1" de toutes les JTM et (b) après écriture de toutes les JTM à "0"..... 123

Figure V-12 : Variations inter-PUF correspondant à chaque puce pour différentes tensions de chauffage..... 124

Figure V-13 : Variation inter-PUF entre toutes les 26 puces pour différentes tensions de chauffage..... 125

Figure V-14 : Variation intra-PUF correspondant à chaque puce pour différentes tensions de chauffage 125

Figure V-15 : Variations intra-PUF moyenne entre toutes les puces pour différentes tensions de chauffage 126

Figure V-16 : Uniformité moyenne entre toutes les matrices pour différentes valeurs de tensions de chauffage 126

Liste des tableaux

Tableau II-1 : Les Technologies MRAM 44

Tableau III-1 : Comparaison des TRNG à base de jonctions magnétiques dans la littérature . 70

Tableau III-2 : Paramètres électriques de la jonction STT utilisé dans le projet GREAT 72

Tableau III-3 : Valeurs des paramètres électriques de la JTM sous les 3 principaux corners .. 72

Tableau III-4 : Caractéristiques des différentes cellules de lecture à base de SRAM dans [114] 80

Tableau IV-1 : Résultat des tests de NIST sur les nombres aléatoires générés sans correction (Raw) ainsi qu'après les corrections XOR^3, Von Neumann, RTOPT et une combinaison de RTOPT et XOR^2. Les tests sont validés si P-value ≥ 0,0001 et Prop ≥ 96/100 102

Tableau V-1 : Comparaison des differents PUF à base de MRAM dans la littérature 122

Glossaire

AMR	Anisotropic Magnetoresistance
CMOS	Complementary Metal-Oxyde Semiconductor
DKit	Design Kit
FIMS	Field Induced Magnetic Switching
FSR	Feedback Shift Register
GMR	Giant Magnetoresistance
JTM	Jonction Tunnel Magnétique
LFSR	Linear Feedback Shift Register
MLU	Magnetic Logic Unit
MRAM	Magnetic Random Access Memory
MTJ	Magnetic tunnel Junction
NIST	National Institue of Standards and Technology
PRNG	Pseudo Random Number Generator
PUF	Physically Unclonable Function
PVT	Process Voltage Temperature
RA	Resistance Area products
Rb	Random bit
RTOPT	Real Time Output Probability Tracking
SAF	Synthetic Antiferromagnet
SoC	System on Chip

SOT	Spin Orbit Torque
SRAM	Static Random Access Memory
SR-MRAM	Self Referenced MRAM
STT	Spin Transfert Torque
TAS	Thermally Assisted Switching
TMR	Tunneling Magnetoresistance
TRNG	True Random Number Generator

Chapitre I Introduction Générale

1. Contexte

Le développement rapide des systèmes sur puce (*SoC, System on chip)* à faible consommation d'énergie, haute densité et haute performance a poussé les mémoires embarquées à leurs limites et a ouvert le champ au développement des technologies de mémoires émergentes. La mémoire magnétique à accès aléatoire (*MRAM, magnetic random access memory*) est apparue comme un choix prometteur pour les mémoires intégrées en raison de sa latence de lecture/écriture réduite et de sa grande capacité d'intégration CMOS. En effet cette technologie émergente qui a été rapidement commercialisée, intègre plusieurs propriétés, notamment la non-volatilité, une haute densité, une endurance illimitée, une robustesse thermique et une immunité aux radiations, lui permettant de se développer sur diverses plateformes d'application.

Une projection de l'évolution du marché des MRAM présentée sur la Figure I-1 montre l'intérêt suscité par cette mémoire émergente [1].

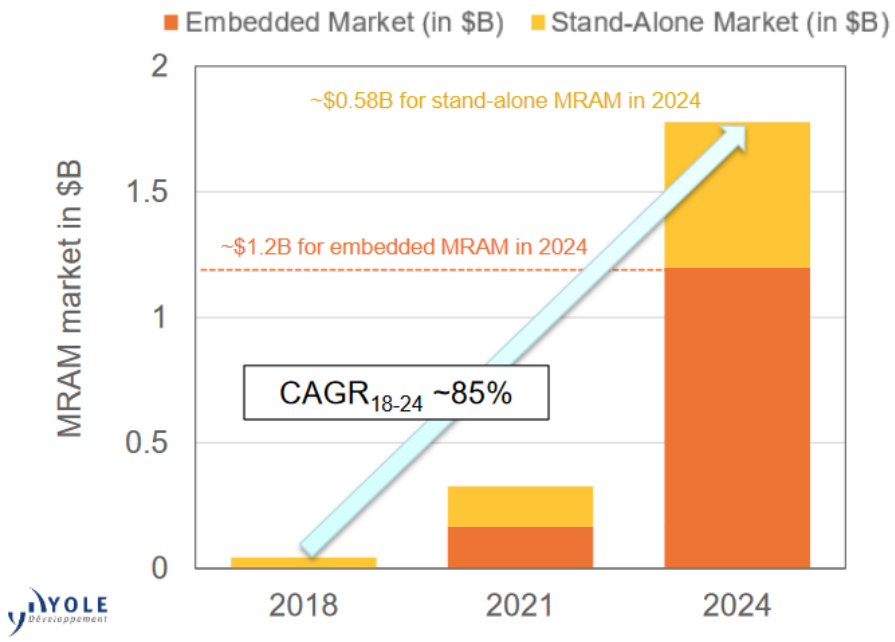


Figure I-1 : Projection de l'évolution du marché des MRAM [1].

En outre, cette forte croissance du marché des MRAM peut être mise en relation avec l'utilisation croissante des appareils électroniques dans tous les aspects de la vie quotidienne et pour un large éventail d'application augmentant dans le même le temps le besoin de

sécurité de l'information circulant à travers ces appareils. En effet l'industrie électronique ayant toujours été confronté aux problèmes de sécurité, consacre un budget assez conséquent dans la sécurisation de ses produits.

La cryptographie est la technique traditionnelle utilisée pour remédier aux problèmes de sécurité. En fonction du système à sécuriser et de la nature des informations sécrètes, une ou toutes les mesures de sécurité suivantes peuvent être appliquées : l'authentification, l'intégrité, la confidentialité et la non-répudiation. Toutefois leur niveau de sécurité dépend fortement de la clé utilisée dans le cas du cryptage, et de l'identifiant dans le cas d'une authentification et donc de la qualité des primitives de cryptographie utilisées.

Les primitives de cryptographie classiques basées sur des algorithmes ou des fonctions mathématiques sont difficile d'implémentation et le plus souvent sont très onéreux. Comme alternative à ces primitives, de nouvelles primitives de sécurité matérielle sont utilisées aujourd'hui, telles que les générateurs de nombres véritablement aléatoires (*TRNG, true random number generator*) et les fonctions physiques non-clonable (*PUF, physically unclonable function*). Les TRNG et les PUF sont des primitives émergentes utilisées pour la génération des clés secrètes sécurisées et l'authentification ou l'identification des dispositifs à faibles coûts.

Les TRNG sont utilisés pour générer des nombres aléatoires à partir d'un processus physique, plutôt que d'un programme informatique. Ils sont généralement mis en œuvre en tirant parti du bruit thermique ou d'autres phénomènes physiques imprévisibles et sont censés générer des bits aléatoires avec une entropie très élevée et une corrélation nulle. Une conception de TRNG sur puce devrait occuper une petite surface, avoir un débit binaire élevé et avoir une faible consommation d'énergie, tout en assurant des flux binaires non biaisés avec une entropie élevée par bit et une faible, voire aucune, corrélation.

Les PUF exploitent la variabilité intrinsèque de fabrication introduite dans des dispositifs au cours de leur processus de fabrication pour générer une signature, unique à chaque dispositif. Afin de garantir sa sécurité, la signature générée doit être unique d'un dispositif à l'autre (non clonable), et, pour un même dispositif, elle doit être robuste par rapport au vieillissement et aux variations environnementales (reproductible).

En plus de leurs fonctionnalités traditionnelles de stockage d'information, les propriétés internes des MRAM les rendent adaptées pour la mise en œuvre de ces deux primitives de cryptographie. L'utilisation astucieuse de ces propriétés internes des MRAM pourrait contribuer à réduire le coût lié à la sécurisation des dispositifs électroniques.

2. Objectif de la thèse

Les recherches intensives autour des MRAM ces dernières années ont placé cette technologie émergente comme un candidat prometteur des mémoires non volatiles embarquées. Parallèlement à leur fonction principale de stockage d'information, les MRAM sont des dispositifs offrant des possibilités dans plusieurs autres domaines d'applications telles que les capteurs, les récepteurs RF, la sécurité matérielle, etc.

Dans cette thèse, nous nous sommes intéressés principalement à l'utilisation des MRAM dans la conception des primitives de cryptographie de base. Plus concrètement une exploration au niveau de la conception de TRNG basée sur des jonctions de type STT-MRAM a été étudiée puis des expérimentations pratiques sur des MRAM de type TAS dans le but de générer des nombres aléatoires ainsi que des identifiants pour en faire des PUF ont été effectuées.

Ce document est organisé comme suit :

- Le Chapitre II présente un état de l'art sur les MRAM à travers une brève historique des principales découvertes ayant abouti à sa conception. Les différents types de MRAM y sont présentés ainsi que leurs méthodes d'écriture et de lecture.
- Dans le Chapitre III, nous abordons la conception de TRNG à base de STT-MRAM. Nous faisons d'abord un état de l'art des méthodes de génération de nombres aléatoires utilisées dans la littérature puis une proposition de conception est présentée. Elle est basée sur l'extraction d'aléa lors de l'écriture d'une jonction STT à travers un calibrage des impulsions d'écriture.
- Le Chapitre IV aborde également la génération de nombres aléatoires à travers cette fois ci des jonctions de type TAS. Des expérimentations pratiques sont menées dans ce sens sur des circuits de test fabriqués par la société CROCUS Technologie. Les résultats de ces expérimentations y sont présentés.
- Dans le Chapitre V, nous présentons une méthode pour l'extraction de signatures unique de ces circuits de test. Des expériences permettant d'analyser l'efficacité d'une PUF conçue avec ces circuits sont menées. Le comportement de commutation stochastique des jonctions TAS comme source d'aléa est mis en évidence.
- Enfin dans le Chapitre VI, une conclusion générale est donnée avec quelques perspectives.

Chapitre II Les Mémoires magnétiques à accès aléatoire (MRAM)

Ce chapitre présente les bases de la spintronique en présentant les principales découvertes relatives aux interactions des spins des électrons avec les propriétés magnétiques des matériaux. De plus, un aperçu des applications de la spintronique est donné avec un accent particulier sur les dispositifs mémoires.

1. Introduction

En électronique, la charge électrique de l'électron a longtemps été la seule propriété fondamentale utilisée dans le traitement de l'information. Dans les années 1980, des découvertes sur les phénomènes de transport d'électrons dépendant du spin ont donné naissance à ce qu'on appelle aujourd'hui la spintronique [2]. La spintronique est un nouveau paradigme pour le stockage d'informations et le traitement logique utilisant une autre propriété fondamentale de l'électron : le spin. L'information est portée à travers les spins de l'électron en lieu et place de sa charge électrique. L'information est contenue dans l'orientation magnétique relative des spins (« up » ou « down ») par rapport à une référence (par exemple l'orientation magnétique d'une couche ferromagnétique). La détection de l'orientation magnétique relative des spins se fait en utilisant les propriétés de transport d'électrons à travers ces matériaux semi-conducteurs ferromagnétiques. Comparés aux dispositifs à semi-conducteurs classiques, les dispositifs de la spintronique sont potentiellement plus rapides, plus économes en énergie et plus denses en plus de leur non volatilité [3].

La MRAM est le fruit de recherches menées dans le domaine de la spintronique qui ont déjà abouti à de grandes avancées dans le domaine du stockage d'information. Une des plus importantes avancées est l'invention de la vanne de spin [4] utilisée dans les disques durs et qui a permis d'augmenter la densité et d'accroitre significativement la quantité d'information pouvant y être stockée.

La suite du chapitre est organisée comme suit : La section II2 présente les principaux phénomènes liés aux interactions de spin dans les dispositifs à semi-conducteurs. Ensuite, l'élément de base dans la conception des mémoires magnétiques est introduit. La section II.3 passes en revue les technologies MRAM existantes. La section II.4 conclut ce chapitre.

2. La spintronique

Le terme « spintronique » résulte de la contraction de deux mots « spin » et « électronique » faisant référence à l'électronique de spin. Le spin est une propriété quantique intrinsèque associée à une particule, qui est caractéristique de la nature de la particule au même titre que sa masse et sa charge électrique. Par analogie le spin peut être vu comme étant l'équivalent quantique du moment angulaire intrinsèque d'une particule élémentaire en physique classique [5].

Les électrons ont une charge et un spin mais, pendant longtemps, charge et spin ont été utilisés séparément. L'électronique classique ignore le spin et déplace les électrons en agissant seulement sur leur charge tandis que la spintronique ou l'électronique de spin, est une technique qui associe contrôle de courants de spins et de charges dans des nouveaux dispositifs intégrables aux circuits de la microélectronique. Le spin d'un électron peut avoir deux états : spin « up » ou spin « down ». Il est communément décrit comme étant la « rotation » de l'électron sur lui-même : la rotation dans un sens correspond au spin up et dans l'autre sens au spin down.

Dans un matériau, c'est le spin des électrons qui est responsable de ses propriétés magnétiques. Ainsi, dans un matériau conducteur classique tel que le cuivre ou l'aluminium, les moments magnétiques sont désordonnés on dit que c'est un matériau paramagnétique (Figure II-1 a) [6]. Par contre dans un matériau comme le fer ou le nickel, les moments magnétiques sont tous alignés dans la même direction et orientés dans le même sens, ce qui se traduit par une aimantation non nulle à l'échelle macroscopique au sein du matériau. Ce type de matériau est dit ferromagnétique (Figure II-1 b). Enfin dans un matériau antiferromagnétique (Figure II-1 c) tel que le chrome, les moments magnétiques sont parallèles entre eux mais avec des orientations opposées au voisinage. Ce qui se traduit par une aimantation nulle à l'échelle macroscopique.

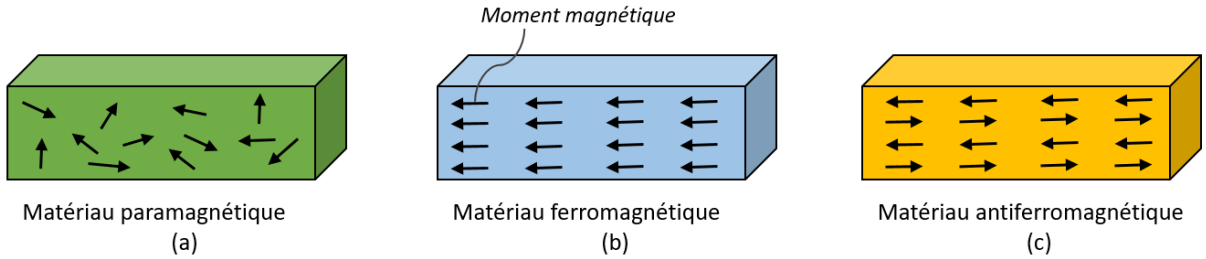


Figure II-1 : Orientation des moments magnétiques selon le type de matériau [6]

Lorsqu'un courant électrique traverse une couche ferromagnétique, les électrons à travers leurs spins sont filtrés en fonction du sens de l'aimantation du matériau. En effet, les

électrons qui ont leurs spins orientés dans la même direction que ceux du matériau, traversent plus facilement la couche ferromagnétique que ceux qui ont leurs spins orientés dans la direction opposée. Ce phénomène a été mis en évidence par William Thomson en 1856 à travers la magnétorésistance anisotrope (*AMR*, *anisotropic magnetoresistance*) grâce à des expériences sur le fer et le nickel. La Figure II-2 [7] présente l'effet AMR résultant de l'interaction entre l'aimantation du matériau (flèches vertes sur la figure) et le spin de l'électron. Les électrons dont le spin est dans le sens inverse de l'aimantation (sphères blanches sur la figure) sont plus dispersés que les électrons dont le spin est parallèle à l'aimantation (sphères rouges sur la figure). Dans les matériaux ferromagnétiques, l'AMR affecte la résistance de l'ordre de quelques pourcents. Cependant, à la fin des années 1970, cela suffisait à développer avec succès des capteurs AMR pour remplacer les capteurs inductifs comme tête de lecture dans les disques durs.

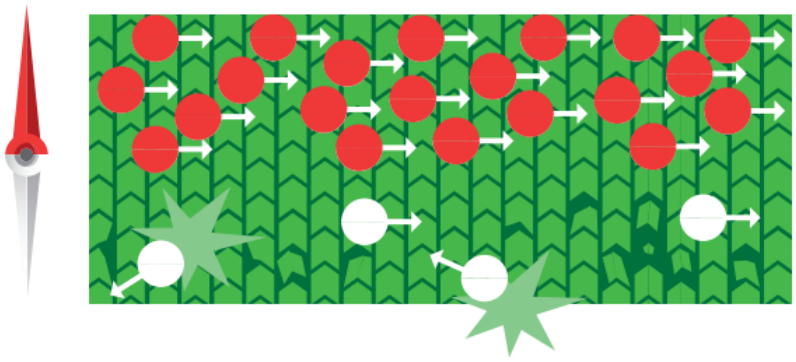


Figure II-2 : Effet de la magnétorésistance anisotrope (AMR) [7]

Le concept général de la spintronique est de placer des matériaux ferromagnétiques sur le trajet des électrons et d'utiliser l'influence du spin sur la mobilité des électrons dans ces matériaux. Cette influence, d'abord suggérée par Mott [8] en 1936, a été ensuite démontrée expérimentalement et décrite théoriquement à la fin des années 60 [9] [10].

Le développement de la spintronique a permis la découverte de la magnétorésistance géante (GMR, *giant magnetoresistance*) en 1988 par deux groupes de chercheurs dirigés respectivement par Albert Fert et Peter Grunberg, de façon indépendante, et qui leur a valu le prix Nobel de physique en 2007 [11] [12]. Le phénomène de la GMR intervient dans des structures qui alternent des couches ferromagnétiques (FM) et des couches non magnétiques (NM).

Le groupe de Albert Fert a étudié la magnétorésistance de trente à soixante structures composées de fer et de chrome (Fe / Cr) empilés et a observé un facteur de presque 2 entre

les résistivités des structures sous champ magnétique externe nul et les résistivités des structures sous champ magnétique externe saturé (Figure II-3).

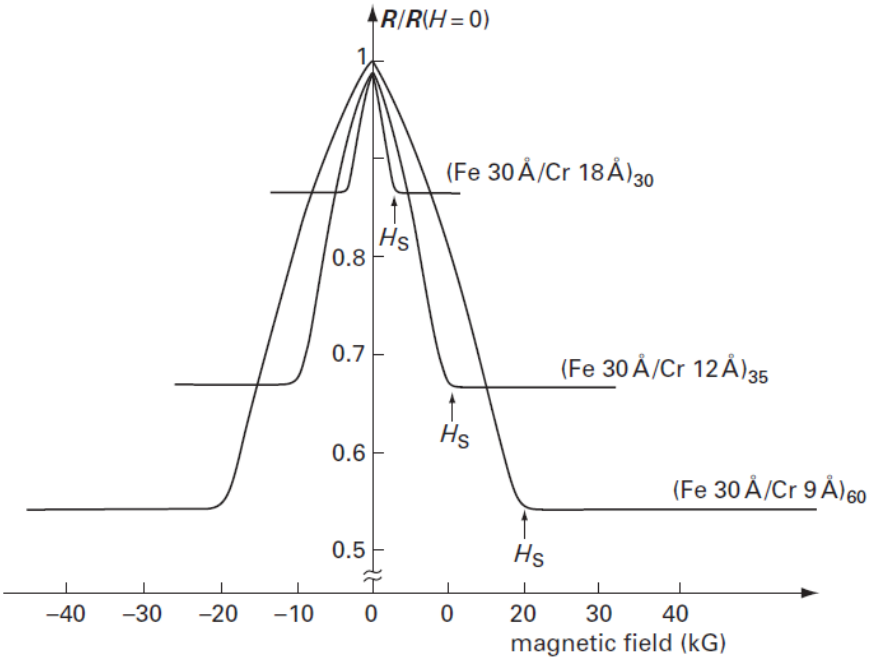


Figure II-3 : Première observation de la GMR sur des multicouches Fe/Cr à 4,2 K [11]

D'autre part, le groupe de P. Grünberg a étudié la magnétorésistance d'une structure composée de Fe / Cr / Fe et a constaté qu'un alignement antiparallèle de l'aimantation des deux couches Fe augmente la résistivité électrique, bien plus que l'effet AMR. La Figure II-4 présente bien ce phénomène. Lorsque le spin des électrons est parallèle à la direction d'aimantation de la couche FM (Fe), les électrons sont faiblement dispersés (traversent facilement les couches FM) et la couche FM présente une faible résistance. D'autre part, lorsque le spin des électrons est antiparallèle à la direction d'aimantation de la couche FM, les électrons ont une forte diffusion et la couche FM présente une grande résistance.

En absence de champ magnétique externe, la structure FM / NM / FM présente un comportement antiferromagnétique c'est-à-dire que les aimantations des deux couches FM s'opposent. Dans une telle structure, les électrons de spin up et les électrons de spin down auront une forte diffusion car ils devront nécessairement traverser une couche avec une aimantation antiparallèle au spin. Alors la résistivité électrique de la structure est grande. Par contre lorsqu'un champ magnétique externe est appliqué de sorte à ce que les deux couches FM soient alignées dans la même direction, les électrons dont le spin est parallèle à l'aimantation des couches FM seront faiblement dispersés. Dans ce cas, la résistivité électrique est faible par rapport à la configuration antiparallèle de la structure.

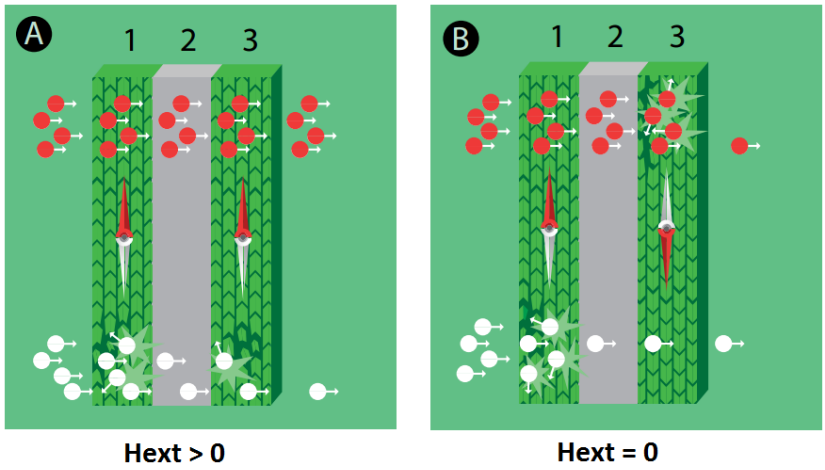


Figure II-4 : Effet de la magnétorésistance géante (GMR) [7]

Une conséquence de la découverte de la GMR a été l'invention de la vanne de spin. Dans ce cas, le dispositif est constitué de plusieurs couches minces. Le plus simple est 3 couches minces : deux couches ferromagnétiques séparées par une couche non ferromagnétique conductrice. Une des deux couches ferromagnétiques a une aimantation fixe. C'est la couche de référence. L'autre couche ferromagnétique a une aimantation libre, c'est la couche de stockage, on peut donc l'orienter dans les deux directions, parallèle ou antiparallèle à celle de la couche de référence, en appliquant un champ magnétique externe. On change donc l'orientation de la couche de stockage en fonction de la donnée à écrire. Pour lire l'information, il suffit de faire passer un courant à travers l'empilement et déterminer si l'on a une faible ou une forte résistance. Cette invention a permis d'accroitre fortement la densité des disques durs actuels et est la principale application industrielle de la spintronique. Ce pendant la différence de résistance que l'on peut obtenir avec un tel dispositif ne dépasse pas 20%. L'augmentation de cette différence de résistance a été possible avec la découverte de la magnétorésistance à effet tunnel (TMR, *tunneling magnetoresistance*).

2.1. La Magnétorésistance à effet tunnel (TMR)

En 1960, I. Giaever démontre pour la première fois que les électrons sont capables de traverser un isolant pour passer d'une couche conductrice à une autre couche conductrice par effet tunnel [13]. Il reçoit le prix Nobel de la Physique en 1973 pour ces travaux [13] [14] [15]. En 1971, P.M. Tedrow et R. Meservey observent l'effet de conservation de spin à travers un tel tunnel composée d'une couche supraconductrice d'aluminium et une couche ferromagnétique de nickel séparées par un isolant [16]. Et en 1975, l'effet tunnel magnétique (TMR) est démontré par M. Jullière qui étudie la conductance de deux couches

ferromagnétiques séparées par un mince isolant (Fe/Ge/Co) [17]. Il observe que la conductance d'une telle jonction est liée à l'orientation magnétique relative des 2 couches ferromagnétiques (Fe et Co). Cet empilement de couche (« ferromagnétique / isolant / ferromagnétique ») appelé jonction tunnel magnétique (JTM ou *MTJ pour Magnetic Tunnel Junction*) est le composant principale des MRAM.

Contrairement au GMR qui est lié à la diffusion d'électrons, la TMR repose sur la polarisation des spins des électrons. En effet, dans une configuration parallèle des orientations magnétiques des couches ferromagnétiques, les électrons dont les spins sont parallèles à la direction de l'aimantation des couches vont « creuser » facilement un tunnel à travers la barrière isolante, alors que les électrons dont les spins sont antiparallèles à la direction de l'aimantation des couches seront majoritairement filtrés ou réfléchis. Dans une configuration antiparallèle de la structure, les flux d'électrons à la fois de spin up et de spin down seront réduits, ce qui entraîne une grande résistance. La variation de résistance dans les jonctions tunnel magnétiques est exprimée en pourcentage par la TMR selon l'équation ($II-1$), où Rp est la résistance de la jonction quand l'aimantation des couches ferromagnétiques est parallèle, Rap quand l'aimantation est antiparallèle, P1 et P2 sont les polarisations de spin des deux couches ferromagnétiques :

$$\mathrm{TMR} = \frac{\triangle \mathrm{R}}{\mathrm{Rp}} = \frac{Rap - Rp}{Rp} = \frac{2.P1.P2}{1 - P1.P2} \qquad (II-1)$$

La valeur de la TMR est étroitement liée au type de matériau ferromagnétique et surtout de l'isolant utilisé. Les premières expériences, faites à température ambiante sur des jonctions à base d'alumine (Al_2O_3), ont montré des TMR de l'ordre des dizaines de pourcent (de 11.8% obtenu par J. Moodera en 1995 [18] jusqu'à 70% obtenu en 2004 par D. Wang [19]).

Cependant, l'évolution concernant la TMR viendra avec des jonctions à base de l'oxyde de magnésium (MgO) comme isolant. En 2001, Butler et al. et Mathon et Umersk ont prédit théoriquement qu'une TMR supérieur à 1000% pourrait être obtenu dans des structures de type Fe (001) / MgO (001) / Fe (001). Ainsi, beaucoup de travaux ont été réalisés, notamment au niveau des procédés de fabrication des couches magnétiques et en 2004, des jonctions à base de MgO du type Fe/MgO/Fe, atteignant une TMR de 200% à température ambiante [20] ont été fabriqués. Puis en 2008, des TMR de 600% ont été observées, toujours à température ambiante sur des structures à base de MgO, mais avec des couches magnétiques à base de Fer, Colbalt et Bore (CoFeB/MgO/CoFeB) [21].

Grace aux différentes recherches et découvertes de ces dernières années, le niveau de la TMR n'a cessé d'augmenter, ce qui est une satisfaction du point de vue de la conception microélectronique. En effet, dans le but d'intégrer des composants magnétiques du type

jonctions tunnel magnétiques aux procédés CMOS, toujours de plus en plus à la pointe de la technologie, et toujours de plus en plus performants, il est important d'avoir une TMR la plus élevée possible. Cela permet de coder et décoder facilement et de façon stable un niveau logique « 0 » ou « 1 ».

2.2. La Jonction tunnel magnétique (JTM)

Une JTM est composée de trois couches, deux couches ferromagnétiques séparées par une mince barrière isolante appelée barrière tunnel. La Figure II-5 montre la structure de base d'une telle jonction, soit dans une configuration où les 2 aimantations sont dans le sens parallèle (a), soit dans le cas où les aimantations sont dans le sens antiparallèle (b). Le fonctionnement est similaire à celui de la vanne de spin. Une des couches ferromagnétiques, appelée couche de référence, a une aimantation fixe tandis que l'autre couche, la couche libre ou couche de stockage, peut voir son aimantation orientée dans les deux sens, parallèle ou antiparallèle à celle de la couche de référence, avec un comportement hystérétique. Ainsi, lorsque les aimantations des deux couches sont dans une configuration parallèle, la résistance de la JTM est faible tandis que dans une configuration antiparallèle, la résistance est grande. Le courant passe à travers la jonction par effet tunnel. L'effet physique qui intervient dans le cas de la JTM est la magnétorésistance à effet tunnel (TMR) comme décrit précédemment. Il est nécessaire d'appliquer un champ magnétique suffisamment important pour passer de l'état parallèle à l'état antiparallèle. Ce champ est appelé champ coercitif (Hc) et doit être orienté dans un sens ou dans l'autre en fonction de la donnée à écrire. Lorsque que le champ est en dessous de cette valeur (en valeur absolue), l'état de la jonction reste stable, la donnée est donc bien stockée de façon permanente. En effet, lorsque le champ magnétique extérieur est nul, l'aimantation d'un matériau ferromagnétique s'aligne spontanément parallèlement à une direction. Il existe deux types d'anisotropie : l'anisotropie magnéto-cristalline et l'anisotropie de forme. Dans le cas de l'anisotropie magnéto-cristalline, le moment magnétique a tendance à s'aligner sur un axe, l'axe dit facile, dont la direction dépend de la structure cristalline du matériau. L'anisotropie de forme est due à la forme de la jonction tunnel. A champ nul, l'aimantation du matériau est parallèle à l'axe facile qui est la position dans laquelle l'énergie d'anisotropie est minimale. C'est pourquoi, dans le cas d'une JTM en forme d'ellipse par exemple, le moment magnétique s'aligne dans la direction du plus grand axe et reste dans le plan des couches. Lorsqu'un champ magnétique extérieur est appliqué, l'anisotropie s'oppose au changement d'aimantation car elle tend à l'orienter dans la direction de l'axe facile. L'anisotropie permet donc une bonne stabilité de l'aimantation de la couche de stockage et permet en particulier d'éviter les changements spontanés d'aimantation dus aux fluctuations thermiques. Cependant, il est nécessaire d'appliquer un champ magnétique

extérieur suffisamment élevé pour contrer l'anisotropie, ce qui augmente l'énergie nécessaire pour l'écriture de la JTM. Il y a donc un compromis à trouver entre stabilité de la donnée et énergie d'écriture de la jonction. En pratique, les mémoires non-volatiles ont une stabilité ou une durée de rétention minimale de 10 ans [22].

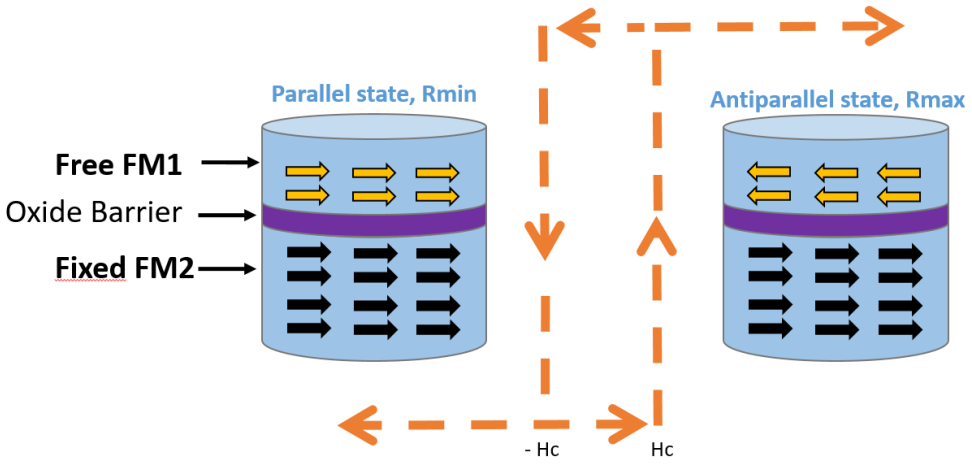


Figure II-5 : Jonction tunnel magnétique [6]

L'opération de lecture de la JTM s'effectue à travers la mesure de la résistance effective de la jonction, qui est fonction de l'état d'aimantation de la couche de stockage. Ceci peut être réalisé en appliquant une tension de polarisation à travers la jonction et en mesurant le courant résultant (approche de détection de courant), ou en envoyant un courant dans la jonction puis en mesurant la tension correspondante (approche de détection de tension). Dans les deux cas, le paramètre détecté est comparé à une valeur de référence pour déterminer l'état de la couche de stockage et ainsi connaitre l'information binaire mémorisée dans la JTM (0 pour l'état parallèle et 1 pour l'état antiparallèle). Un amplificateur de détection à base de CMOS (Sense Amplifier) est généralement utilisé pour obtenir cette information. La lecture fiable de la donnée écrite nécessite que les niveaux de résistance parallèle et antiparallèle soient bien définis et surtout suffisamment éloignés. La Figure II-6 (a) montre les dispersions de résistance mesurées avec une tension de polarisation de 350mV sur une assemblé de 8 milles JTM à base de MgO obtenus sur un circuit mémoire présenté dans [23].

Néanmoins, la valeur de résistance de la jonction ne dépend pas uniquement de l'orientation relative des couches ferromagnétiques, mais aussi de la tension de polarisation appliquée aux bornes de la JTM pendant la lecture [23]. Comme le montre la Figure II-6 (b) pour des jonctions utilisant des barrières tunnel en AlOx et en MgO, la TMR est fortement affectée par la tension de lecture, par conséquent, la tension de lecture doit être aussi faible que possible.

Le mécanisme de lecture est suffisamment mature en comparaison au mécanisme d'écriture. Plusieurs générations différentes de JTM ont été élaborées. Toutes se distinguent par leur mécanisme d'écriture ainsi que par l'empilement des couches qui les composent. Ces différents mécanismes d'écriture seront présentés dans les sections suivantes.

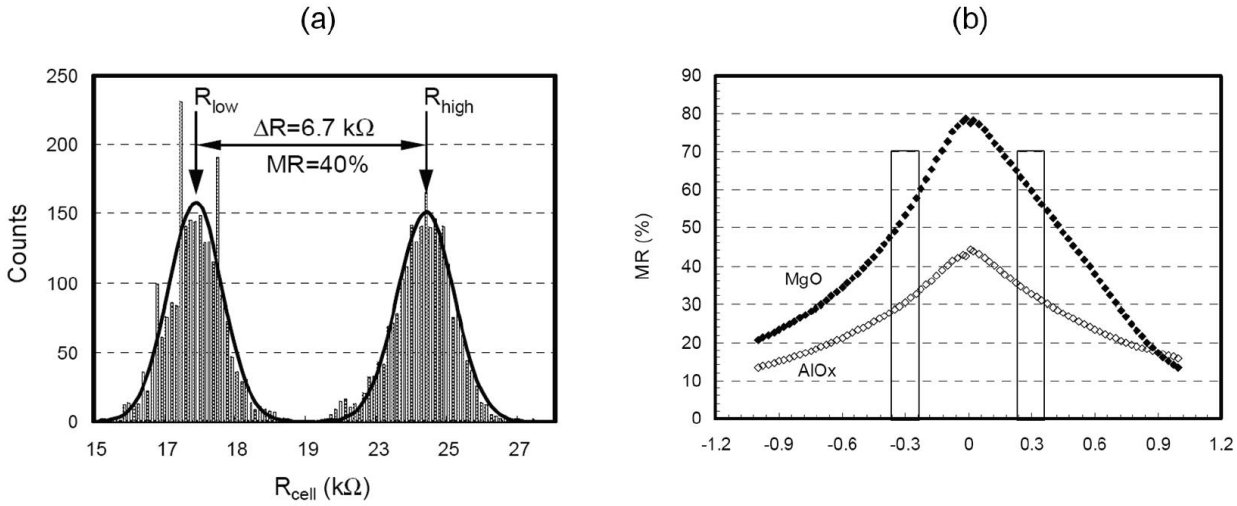


Figure II-6 : (a) Distribution des valeurs de résistances min et max d'une mémoire MRAM de 8Kbits (b) Dépendance de la TMR à la tension de lecture pour des JTM à base de MgO et AlOx.

3. Les technologies MRAM

Les MRAM font partie de la catégorie des mémoires ferromagnétiques. L'élément principal des MRAM est la jonction tunnel magnétique (JTM), dont les différentes variantes seront décrites ci-après dans ce manuscrit selon leur mode d'écriture. Les mémoires MRAM sont aujourd'hui considérées comme de très bonnes candidates parmi toutes les mémoires non volatiles émergentes, grâce à leurs caractéristiques et propriétés intéressantes pour une majeure partie des applications du monde actuel : consommation, vitesse, endurance quasi infinie, immunité aux radiations et miniaturisation.

3.1. Mécanisme d'écriture FIMS: Field Induced Magnetic Switching

La technologie FIMS (*Field Induced Magnetic Switching*) est la première génération de JTM à être intégrée aux procédés de fabrication CMOS. Pour cette techno, le champ nécessaire au changement de l'aimantation de la couche de stockage est si élevé que le courant nécessaire pour générer ce champ est bien trop grand (plusieurs dizaines de mA). Alors deux lignes de champ magnétique (lignes d'écritures) perpendiculaires sont utilisées pour générer un champ magnétique suffisant au retournement de la couche de stockage. La JTM est située à

l'intersection de ces deux lignes de champ disposées en angle droit au-dessus et en dessous de la cellule comme présenté dans la Figure II-7 (b). Au cours de l'écriture, deux courants électriques sont envoyés dans les lignes de champ, générant un champ magnétique au point de croisement des lignes d'écriture, la composition des deux champs magnétiques génère un champ plus élevé que le champ coercitif de la couche de stockage impliquant le changement de son orientation magnétique.

Deux problèmes se sont posés avec cette méthode. Premièrement, un courant important est nécessaire pour générer un champ magnétique suffisant pour inverser la magnétisation de la couche de stockage. Deuxièmement, cette approche présente un problème de sélectivité dans des matrices de JTM dense. En effet, dans un réseau de mémoire, pour obtenir une densité convenable, la taille des JTM doit être réduite ainsi que les distances qui les séparent. Cependant, la réduction des JTM n'impacte pas son champ de seuil de commutation, par conséquent, la probabilité d'écrire une jonction à moitié sélectionnée augmente très fortement lorsque la densité mémoire augmente, rendant l'approche FIMS difficilement scalable.

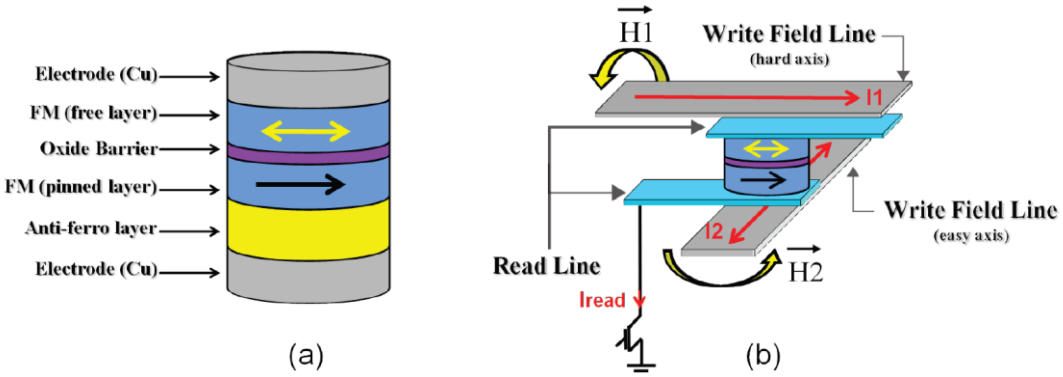


Figure II-7 : (a) Structure d'une JTM en technologie FIMS (b) Principe d'écriture FIMS

3.2. Mécanisme d'écriture Toggle

La méthode d'écriture toggle a été proposée par Savtchenko [24] dans le but de résoudre le problème de sélectivité observé dans la technologie FIMS. La structure de la cellule mémoire toggle se diffère de celle de la mémoire FIMS par l'empilement plus complexe de couches constituant la JTM, la position de la jonction par rapport aux lignes de champ et la séquence d'écriture. La Figure II-8 (a) présente une cellule de type toggle qui est constituée d'un empilement d'une dizaine de couches ultra minces. En remplaçant la couche de référence et la couche de stockage par des systèmes antiferromagnétiques synthétiques (SAF, *Synthetic antiferromagnet*), on obtient un empilement dont l'aimantation totale est nulle, ce qui permet

d'augmenter la stabilité du système. L'écriture d'une telle jonction se fait en commutant l'orientation magnétique des deux couches ferromagnétiques constituant le « Free SAF » et la résistance totale de la jonction est déterminée en fonction de l'orientation magnétique de la couche de référence et celle de la couche de stockage au contact de la barrière tunnel (oxide barrier).

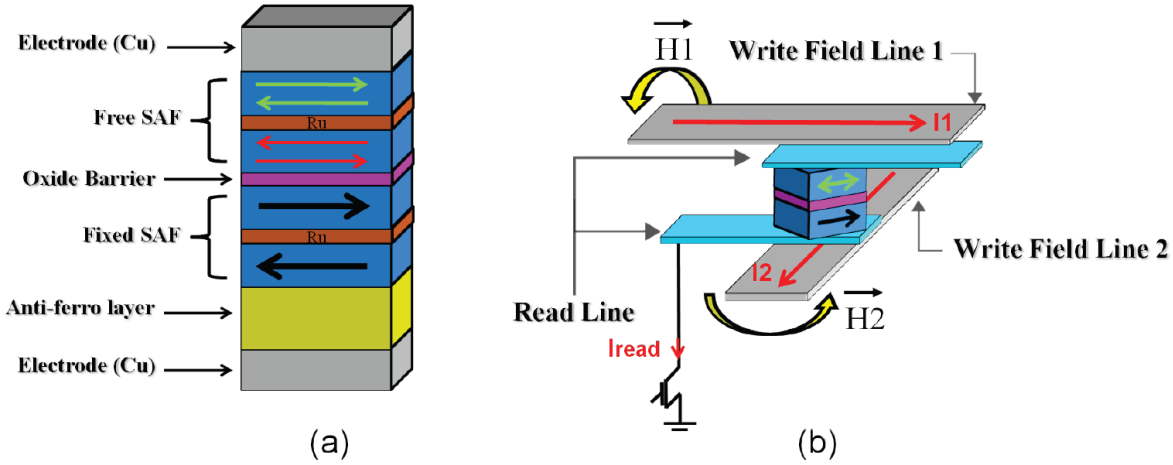


Figure II-8 : (a) Structure d'un JTM en technologie Toggle (b) Principe de lecture/écriture Toggle

Les jonctions sont aussi orientées à 45 degrés par rapport aux lignes de champ comme indiqué sur la Figure II-8 (b). La commutation du « Free SAF » constituant la couche de stockage se fait en appliquant une séquence spécifique de champs magnétiques permettant un retournement uniforme de la structure (Figure II-9). Cette technique permet de faire basculer l'orientation magnétique de la couche de stockage dans un sens comme dans l'autre. Avant d'écrire, il faut donc d'abord lire la donnée stockée pour ensuite la comparer à la donnée à écrire. Si les deux valeurs sont différentes, alors on applique la séquence de courant pour faire basculer l'orientation magnétique de la « Free SAF », sinon il n'est pas nécessaire d'appliquer la séquence d'écriture.

Le fait d'appliquer une séquence d'impulsions de courant pour générer un champ permet de s'affranchir des problèmes de sélectivité observés dans la techno FIMS. Cependant, le courant d'écriture reste encore très élevé ce qui rend la miniaturisation de cette technologie difficile en dessous du nœud technologique 90 nm.

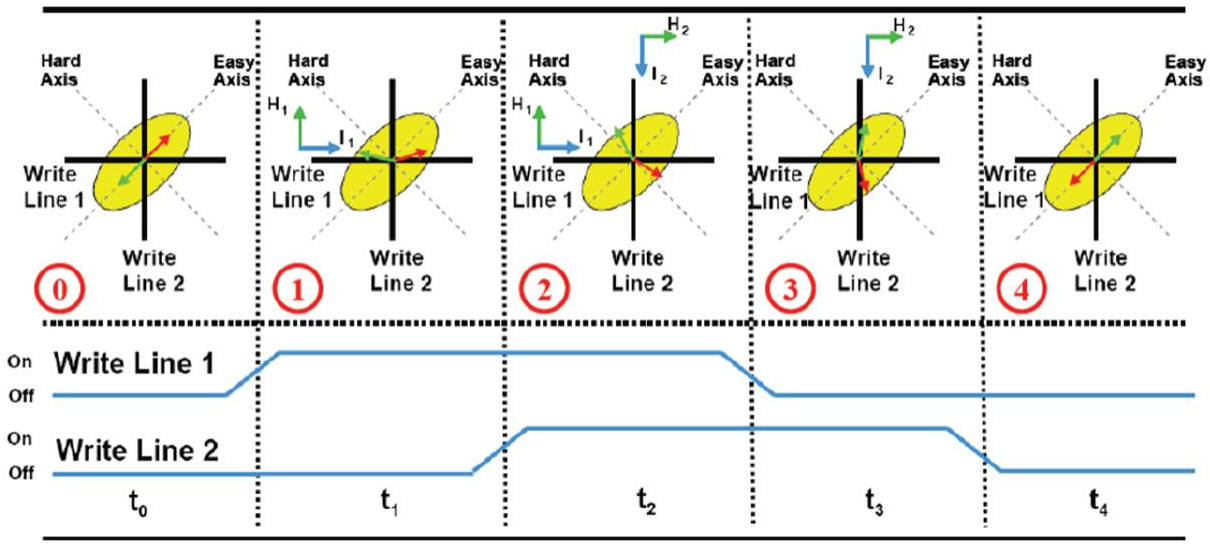


Figure II-9 : Séquence d'écriture Toggle

3.3. Mécanisme d'écriture TAS: Thermally Assisted Switching

Pour résoudre à la fois le problème de sélectivité et de miniaturisation des JTM, une autre méthode d'écriture a été développée et brevetée par le laboratoire SPINTEC, CEA/CNRS en 2001 appelée TAS (Thermally Assisted Switching). La structure de la jonction est présentée sur la Figure II-10. Deux couches antiferromagnétiques (AFM1 et AFM2) avec des températures de blocage différentes sont ajoutées de part et d'autre de la jonction. Ces couches antiferromagnétiques permettent de maintenir l'aimantation des couches FM de stockage et de référence stable à température ambiante. La couche AFM2 placée au-dessus de la couche de stockage, a une température de blocage faible (environ 150°C). L'orientation magnétique de la couche de stockage peut donc être changée lors de l'écriture en chauffant la JTM à une température supérieure à celle de AF2 (T°>150°C). Afin d'éviter toute commutation de la couche de référence, la couche AFM1 placée en dessous de la couche de référence, a une température de blocage beaucoup plus élevée (autour de 300°C). La direction de l'orientation magnétique de la couche de référence est donc choisie et fixée lors de la fabrication.

La particularité de ce type d'empilement est qu'il est très stable du point de vue de l'aimantation à température ambiante. Afin de pouvoir changer l'aimantation de la couche de stockage il est nécessaire de chauffer la couche antiferromagnétique placée au-dessus de la couche de stockage, au-delà de sa température de blocage. Lorsque la jonction est suffisamment chauffée, les spins de la couche de AFM2 sont complètement désordonnés et un faible champ magnétique suffit pour imposer une aimantation à la couche de stockage, pour laquelle les spins vont être dans une direction dépendante du champ. La température ne

doit pas non plus être excessive pour ne pas risquer de dé-piéger l'aimantation de la couche de référence.

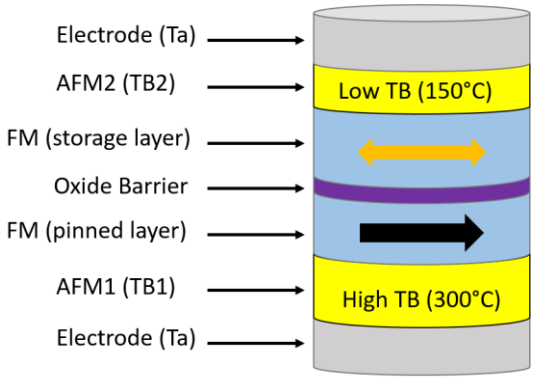


Figure II-10 : Structure d'une JTM en technologie TAS

Le principe d'écriture TAS est décrit par la Figure II-11. Tout comme avec la méthode d'écriture toggle, la méthode TAS est séquentielle et se fait en 3 étapes. On part d'un état initial de la JTM en configuration parallèle. La première étape consiste à chauffer la jonction à l'aide d'un courant à travers le transistor en série, c'est la phase dite « Heating ». Lorsque la température de la JTM est supérieure à la température de la couche AFM2, on passe à l'étape 2 en faisant circuler un courant dans la ligne de champs induisant un champ magnétique tout en continuant à chauffer la jonction. La couche de stockage étant dé-piégée, un faible champ magnétique suffit à retourner son aimantation c'est la phase dite « Switching ». Enfin, pour piéger l'aimantation de la couche de stockage dans son nouvel état, on passe à la troisième étape en stoppant le courant de chauffage tout en maintenant le champ magnétique actif à travers le courant de ligne d'écriture. C'est la phase de refroidissement sous champ magnétique dite « Cooling ». Le refroidissement sous champ permet de stabiliser le couplage d'échange entre la couche AFM2 et la couche de stockage. La jonction est alors, à l'état final dans une configuration antiparallèle. Chaque étape a une durée d'une dizaine de nanosecondes pour une durée totale d'une trentaine de nanosecondes pour un cycle d'écriture.

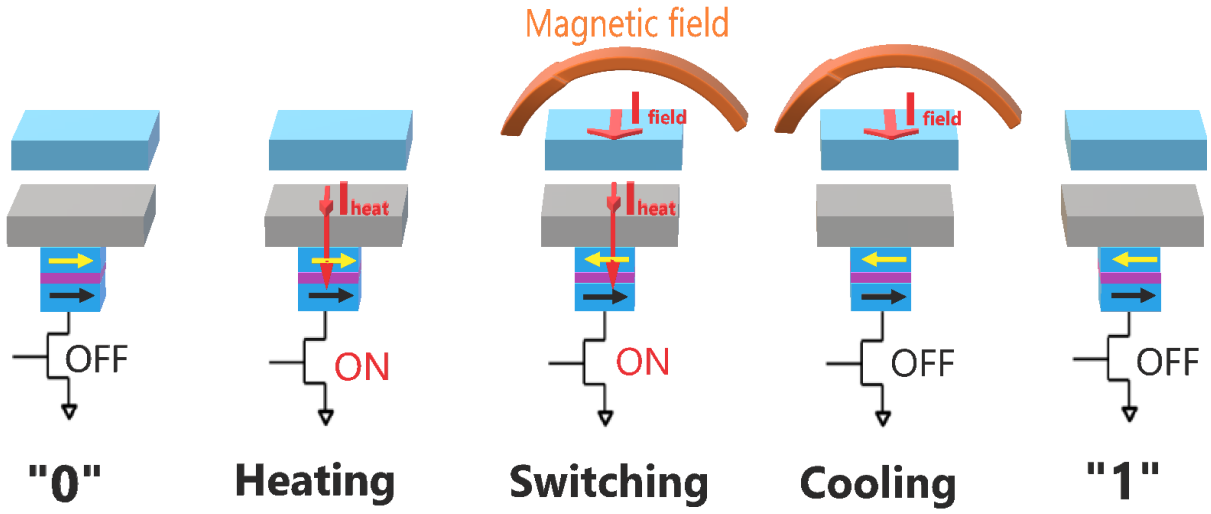


Figure II-11 : Séquence d'écriture d'une JTM TAS : Passage d'un état parallèle à un état antiparallèle

Ce principe d'écriture permet d'économiser de l'énergie par rapport aux deux précédentes méthodes. En effet le courant pour chauffer la cellule dépend fortement de la résistance de la jonction définie par le paramètre RA (*Resistance-Area products*) et la surface de la JTM, comme le montre l'équation $(II-2)$. Aussi, le paramètre PA (mW/um^2) correspondant à la puissance nécessaire pour atteindre la température de blocage peut être réduit en intégrant des barrières thermiques afin de confiner la chaleur. De plus, les lignes de champ d'écriture ainsi que les générateurs de courant nécessaires à la génération de ce champ peuvent être mutualisés entre plusieurs jonctions permettant ainsi d'écrire plusieurs jonctions simultanément dans une configuration matricielle de mémoire TAS-MRAM. Ce qui réduit considérablement l'énergie globale d'écriture de la mémoire. Cette méthode TAS permet également de palier au problème de sélectivité, car en effet il est facile de maitriser les jonctions qui seront chauffées et celles qui ne le seront pas.

$$I_{heat}(mA) = \sqrt{\frac{Pheat}{Rp}} = \sqrt{\frac{PA * S}{\frac{RA}{S}}} = \sqrt{\frac{PA * S^2}{RA}} \qquad (II-2)$$

La Société Crocus Technology, issue du laboratoire Spintec en 2004, détient les brevets de la TAS-MRAM et exploite actuellement cette technologie. Son portefeuille de produits s'étend des puces de mémoire aux capteurs magnétiques ainsi qu'aux microcontrôleurs à sécurité renforcée. En 2010, la société crocus a breveté une autre implémentation de la TAS appelée « Self referenced » MRAM (SR-MRAM) ou Magnetic logic unit (MLU) [25]. La MLU diffère des structures MRAM habituelles en ce qu'elle ne comporte aucune couche de référence dure dont la direction d'aimantation reste fixe pendant le processus de fabrication, mais plutôt une

couche de référence souple à aimantation commutable, appelée couche de détection (SRL). Dérivée de la TAS-MRAM, sa couche de stockage est toujours piégée à température ambiante par une couche antiferromagnétique pour assurer la conservation des données. La Figure II-12 présente la structure simplifiée de la MLU.

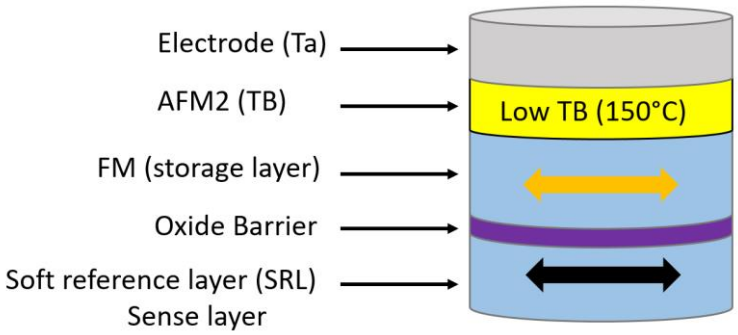


Figure II-12 : Structure d'une jonction SR ou MLU

La couche de référence SRL n'étant pas soumis à l'effet d'une couche antiferromagnétique de blocage, un faible champ magnétique externe suffit à faire commuter son orientation magnétique. De plus, en absence d'un champ magnétique externe, les interactions magnétostatiques entre les deux couches ferromagnétiques feront que la couche SRL passera spontanément à une configuration antiparallèle par rapport à la couche de stockage. La seule valeur de résistance mesurable sera donc la valeur maximale (Rmax) en l'absence de champs. La MLU ne peut donc pas être lue avec un schéma de mesure de résistance directe comme dans les précédentes JTM.

Bien que le principe d'écriture reste le même pour la TAS que pour le MLU, l'opération de lecture est assez différente. Dans ce cas, la résistance de la jonction est mesurée deux fois pour une lecture. Une première fois avec la magnétisation de la couche SRL commutée dans un sens et une seconde fois avec la magnétisation de la SRL commutée dans le sens opposé. La variation de résistance entre les deux mesures fournit des informations sur la magnétisation de la couche de stockage. Cette nouvelle approche augmente le temps de lecture, mais la tolérance aux variations de procédés de fabrication est nettement améliorée puisque chaque cellule est auto-référencée. De plus, cette approche réduit considérablement les erreurs de lecture observées dans les autres technologies et qui sont principalement liées aux problèmes de chevauchement des résistances Rp et Rap des JTMs quand la densité de la mémoire augmente. Pour une cellule MLU, l'utilisation de la lecture différentielle de résistance n'est pas sensible à ce problème de fabrication. En outre, le matériau antiferromagnétique à haute température de blocage utilisé dans les couches de référence

des TAS-MRAM n'est plus nécessaire dans les cellules SR-MRAM, ce qui entraîne une augmentation importante de la plage de températures de fonctionnement. En effet, dans la TAS-MRAM l'utilisation de deux couches antiferromagnétiques avec des températures de blocage différentes limite la plage de température de programmation (par exemple une cellule avec 150 ° C pour la couche de stockage contre 300 ° C pour la couche de référence, permet des opérations dans une plage inférieure à 130 ° C). Grâce à l'utilisation d'une seule couche antiferromagnétique, l'une de ces limites est supprimée et la plage de programmation peut s'étendre à des températures beaucoup plus élevées car il n'y a plus de risque de dé-piéger la couche de référence lors de l'écriture. Des couches antiferromagnétiques avec une température de blocage plus élevée peuvent également être utilisées pour piéger la magnétisation de la couche de stockage pour les applications à haute température.

En outre, le fonctionnement de la SR-MRAM conduit également à une autre caractéristique intéressante : la cellule peut également servir de porte logique OU exclusif (XOR). En supposant que les deux couches magnétiques soient les entrées et que la résistance de la JTM soit la sortie, la table de vérité d'une porte logique XOR peut être construite (Figure II-13) [26]. Cela fait de la SR-MRAM un composant particulièrement utile dans des applications de sécurité. Un exemple d'application introduite par la technologie Crocus est le match in place [27], illustré à la Figure II-14. Si la direction actuelle appliquée sur la ligne de champ pendant une opération de lecture est considérée comme une entrée d'information, le dispositif peut utiliser sa capacité logique XOR pour comparer in situ les données stockées en mémoire avec l'entrée (et contrôler s'il existe une concordance ou non, en comparant la résistance de sortie avec une référence). D'autres applications possibles de SR-MRAM sont les suivantes: mémoire à contenu adressable, NOR-MRAM, NANDMRAM [26].

Bien que la structure de la TAS-MRAM lui confère une meilleure évolutivité que celle des technologies précédentes, elle a besoin d'un délai non négligeable pour terminer son opération d'écriture en raison des processus de chauffage / refroidissement. De plus, puisqu'un champ magnétique externe est utilisé pour commuter la JTM, le courant reste élevé, même s'il est inférieur à celui de la technologie Toggle. Un moyen possible de réduire l'énergie d'écriture consiste à combiner la méthode TAS avec l'effet STT (que nous verrons juste après). La TAS-MRAM pourrait être extensible jusqu'à 45 nm [26]. Cependant, l'association de la TAS et la STT a pour objectif d'améliorer encore la miniaturisation de la jonction. Le principal avantage de la TAS-MRAM est sa grande stabilité thermique grâce à sa structure. Cette dernière permet une très bonne conservation des données et une bonne fiabilité contre les perturbations de champ magnétique.

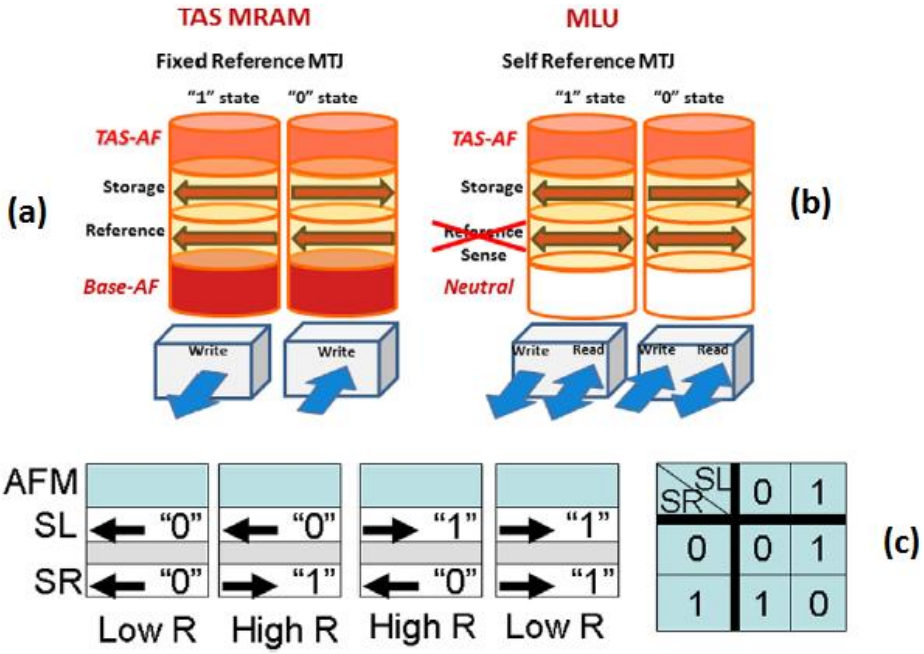


Figure II-13 : Unité logique magnétique (a) cellule TAS-MTJ (b) cellule SR-MTJ (c) porte logique virtuel XOR

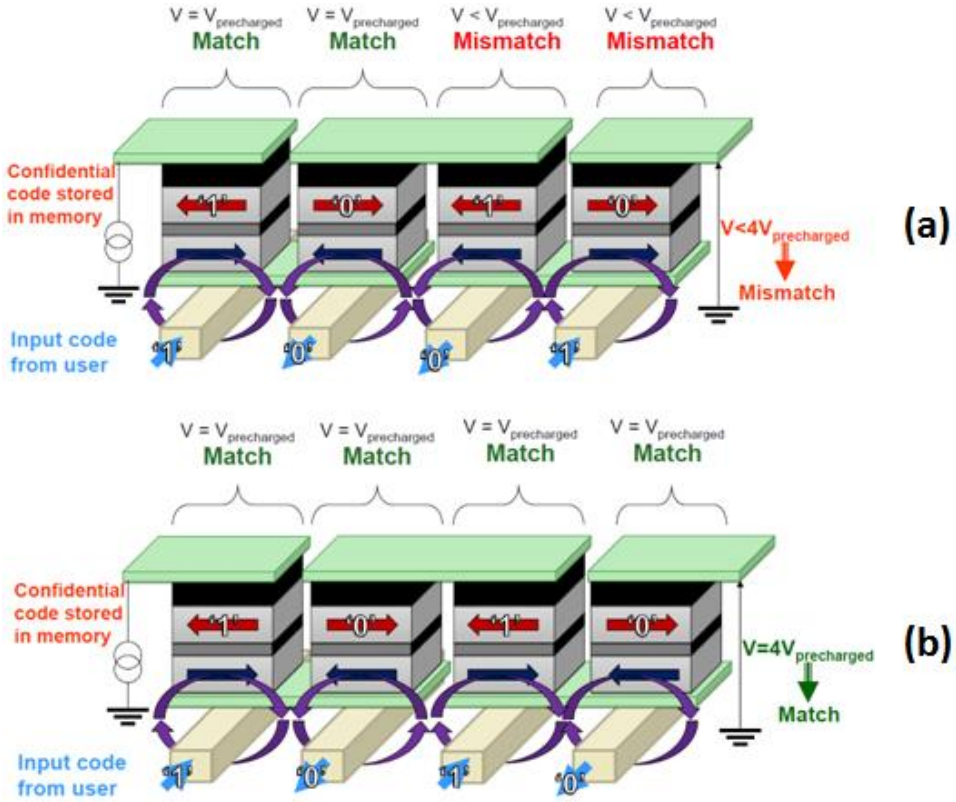


Figure II-14 : Principe du « Match in place » (a) code non valide, (b) code valide

3.4. Mécanisme d'écriture STT : Spin Transfert Torque

Les paragraphes précédents portaient sur l'écriture par champ magnétique externe. Cependant, il est possible d'utiliser un courant polarisé en spin pour changer l'orientation de la couche de stockage de la jonction. En effet un courant électrique est généralement non polarisé (composé de 50% d'électrons de spin-up et 50% d'électrons de spin-down). En faisant circuler un courant à travers une épaisse couche ferromagnétique (appelée polariseur), on peut produire un courant polarisé en spin de forte densité composé majoritairement des électrons dont le spin est parallèle à l'orientation magnétique du polariseur. Ainsi, en redirigeant ce courant à travers une mince couche magnétique, on peut transférer le spin des électrons à la couche et changer son orientation. L'effet physique lié à ce phénomène est appelé couple de transfert de spin (*STT, spin transfert torque*).

Le déclenchement des travaux de recherche sur la méthode d'écriture STT s'est fait suite aux 2 publications de Slonczewski [28] et de Berger [29] qui ont prédit indépendamment que la circulation d'un courant suffisamment fort et perpendiculaire au plan dans un empilement métallique peut réorienter l'aimantation d'une couche de cet empilement. Par la suite, plusieurs équipes de recherche l'ont démontré expérimentalement à travers le monde.

Le phénomène de transfert de spin survient donc lorsqu'un courant polarisé en spin traverse un matériau magnétique. La structure de la jonction utilisée pour ce type d'écriture est présentée dans la Figure II-15 (a). La couche de référence ayant une orientation magnétique fixe est utilisée comme un polariseur pour créer un courant polarisé en spin, soit en transmission ou en réflexion en fonction de la direction du courant dans la jonction. La direction du courant à travers la JTM détermine donc la valeur de la donnée à écrire, à savoir '0' ou '1'.

L'opération d'écriture d'une structure STT est décrite dans la figure Figure II-15 (b) et la Figure II-15 (c). Lorsqu'un courant est envoyé à travers la jonction de l'électrode supérieure vers l'électrode inférieure comme décrit sur la Figure II-15 (b), cela provoque l'injection d'une majorité d'électrons polarisés en spin en direction de la couche piégée vers la couche de stockage. Les électrons polarisés en spin vont affecter l'aimantation de la couche de stockage grâce au couple de transfert de spin. Lorsque la densité de courant dans la jonction est supérieure à la densité de courant de seuil d'écriture ou courant critique (J_{c-}), le couple de transfert de spin renverse l'aimantation de la couche de stockage et l'aligne avec celle de la couche piégée. On passe alors d'un état antiparallèle à un état parallèle de la JTM. Dans la Figure II-15(c), un courant est envoyé à travers la jonction, de l'électrode inférieure vers l'électrode supérieure, les électrons dont le spin est opposé à la couche de référence sont réfléchis. Le spin des électrons réfléchis exerce un couple sur l'aimantation de la couche de

stockage. Lorsque la densité de courant dans la jonction est supérieure à la densité de courant de seuil d'écriture (J_{c+}), le couple de transfert de spin renverse l'aimantation de la couche de stockage et la rend opposée à celle de la couche piégée. On passe ainsi d'un état parallèle à un état antiparallèle.

Il existe de nos jours deux types de magnétisation des couches ferromagnétiques : la magnétisation dans le plan (*In-plane magetization*) et la magnétisation perpendiculaire (*perpendicular magnetization ou PMA*). Dans la magnétisation dans le plan, qui est celle utilisée dans les générations antérieures de MRAM, l'orientation magnétique est parallèle au plan de la JTM, tandis que dans la magnétisation perpendiculaire, l'orientation magnétique est perpendiculaire au plan de la JTM. La STT-MRAM perpendiculaire (PMA STT) a été introduite pour réduire davantage le courant de commutation de la JTM et augmenter sa miniaturisation.

L'opération de lecture de la JTM repose sur la mesure de la résistance effective de la cellule. Pour réaliser cette action, on applique une tension de polarisation aux bornes de la jonction et on en mesure le courant résultant. La lecture diffère d'une écriture par l'utilisation d'un courant plus faible. L'état de l'art a montré que le temps d'accès en lecture des STT-MRAM est similaire et parfois meilleur que son équivalent SRAM [30][31] [32].

En ce qui concerne les opérations d'écriture, malgré le fait que cette nouvelle approche réduit considérablement l'énergie d'écriture par rapport aux méthodes précédentes, certaines limitations ont été observées.

Tout d'abord, les opérations de lecture et d'écriture utilisent le même chemin, ce qui peut entraîner des écritures inattendues lorsque la lecture est en cours, en particulier avec des nœuds technologiques avancés. Pour pallier ce problème, une solution a été proposée au niveau dispositif, consistant à concevoir une structure JTM à trois terminaux avec deux chemins séparés spatialement et électriquement, pour l'écriture et la lecture [33].

Deuxièmement, le courant nécessaire pour commuter la couche de stockage d'un état parallèle à un état antiparallèle (et inversement) n'est pas symétrique [34]. Passer d'un état parallèle à un état antiparallèle nécessite plus de courant que l'inverse. En effet, le passage d'un état antiparallèle à un état parallèle est effectué par des électrons qui sont polarisés en spin (électrons majoritaires) et qui traversent la jonction facilement, tandis que le passage d'un état parallèle à un état antiparallèle est réalisé par des électrons réfléchis (électrons minoritaires) par la couche de référence ou polariseur. Une solution a également été proposée pour éliminer ce problème en ajoutant un polariseur complémentaire [35]. Dans le dispositif proposé, la JTM a deux couches piégées au lieu d'une, avec des orientations magnétiques

opposées. En fonction des informations à écrire, le courant d'écriture traverse la couche piégée correspondante.

Troisièmement, les STT-MRAM sont confrontées à un problème de miniaturisation. Lorsque l'on réduit la taille de la jonction, le facteur de stabilité thermique décroît de façon linéaire avec la surface de la jonction. Cela peut entraîner des problèmes de rétention [36]. De plus, bien que son énergie d'écriture soit faible comparée à celles des autres méthodes d'écriture vu précédemment, la jonction STT a souvent besoin de transistors d'accès de taille supérieures à la taille minimale du nœud technologique [37], ce qui limite donc la densité de la mémoire. Ceci est également un problème pour les applications hautes performances qui nécessitent une vitesse d'écriture élevée, car le courant d'écriture de la STT augmente lorsque la durée d'impulsion d'écriture diminue.

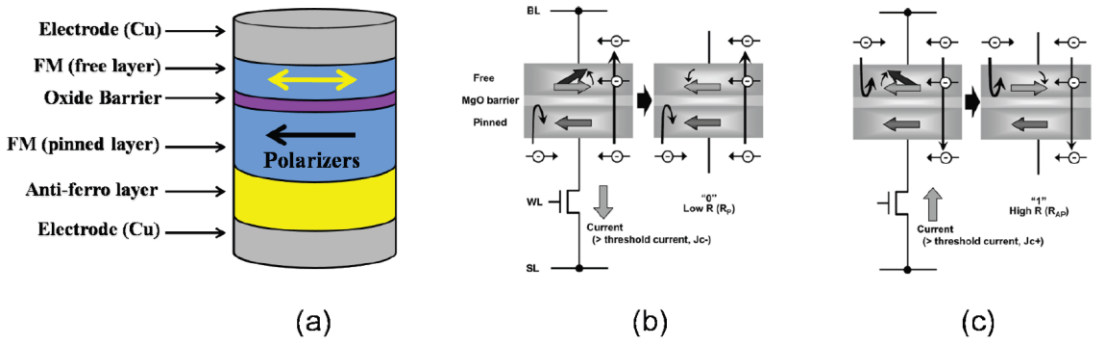


Figure II-15 : (a) structure de la STT-JTM (b) Passage d'un état AP => P (c) Passage d'un état P=> AP

3.5. Mécanisme d'écriture VCMA : Voltage-Controlled Magnetic Anisotropy

Afin de poursuivre la miniaturisation des JTM et de réduire davantage l'énergie de commutation observée avec les STT-MRAM, une JTM commandée en tension ou champ électrique a été proposée [38][39][40], également connu sous le nom de mémoire magnétoélectrique à accès aléatoire (MeRAM).

Dans les matériaux ferromagnétiques tels que ceux généralement utilisés dans la conception des JTM, les champs électriques sont filtrés en fonction de la conductivité du matériau et ne pénètrent donc que de quelques angströms dans la surface du film [41]. Par conséquent, la concentration du champ électrique près de la surface est en principe une limitation du contrôle par le champ électrique des propriétés magnétiques. Cependant, en utilisant des matériaux ferromagnétiques ultra-minces (<2 nm), les propriétés magnétiques peuvent être sensibles aux effets d'interface, voire dominées par ceux-ci, fournissant ainsi un

mécanisme permettant de coupler le champ électrique appliqué à l'anisotropie magnétique du matériau. Ainsi, la tension au niveau de l'interface des matériaux ferromagnétiques peut être utilisée pour réaliser des dispositifs nano magnétiques contrôlés par un champ électrique. Comme le montre la Figure II-16, cette approche utilise la tension plutôt que le courant pour inverser l'aimantation de la couche de stockage grâce à l'effet d'anisotropie magnétique contrôlé par tension (VCMA). La couche de stockage a une anisotropie magnétique perpendiculaire qui peut être modifiée par la tension. Par conséquent, une commutation induite par la tension de l'aimantation peut être effectuée en modifiant l'anisotropie magnétique de la jonction. La structure de la JTM contrôlée en tension (VMTJ) est semblable à celles des technologies MRAM antérieures, ce qui permet de conserver ainsi les mêmes procédés de fabrication [38]. La VMTJ a un comportement contrôlé par tension unipolaire, c'est-à-dire que la commutation est effectuée par des tensions de réglage / réinitialisation d'amplitudes différentes mais de même polarité, alors que la STT-MRAM utilise des polarités de courant opposés pour commuter la JTM.

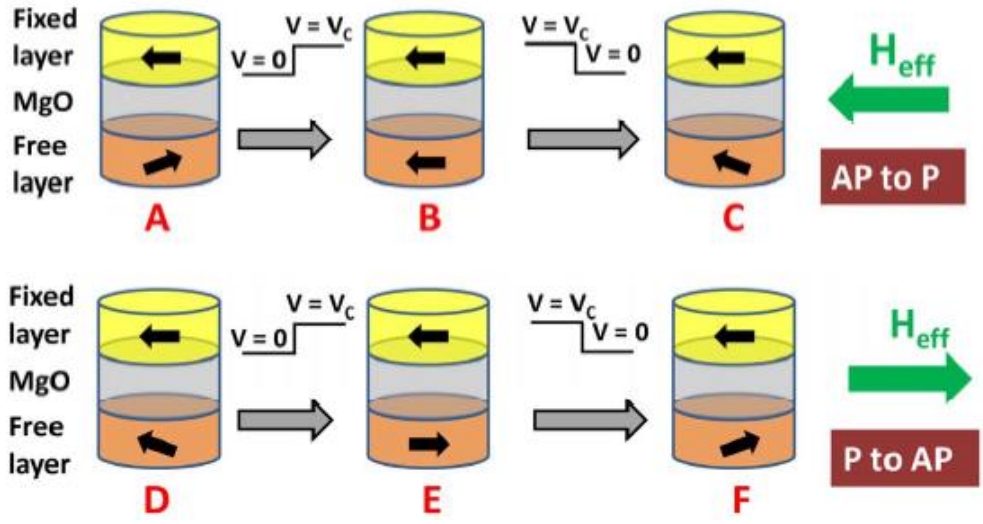


Figure II-16 : Commutation induite par tension (MeRAM) [38]

Etant donné que la commutation de la VMTJ est effectuée via une tension, l'épaisseur de la barrière peut être augmentée afin de réduire la conductance parasite et donc l'effet des couples induits par le courant (c'est-à-dire l'effet STT). De plus, une TMR élevée (supérieure à 100%) est possible permettant une lecture facile de l'aimantation de la couche de stockage.

Un des avantages clés des JTM contrôlées en tension par rapport à celles contrôlées en courant (STT), ou en champs magnétique, est le potentiel de réduction drastique de la dissipation de puissance. En effet, en éliminant le besoin de courants pour l'écriture les JTM,

la dissipation ohmique qui dans la plupart des cas est le mécanisme de perte prédominant dans la mémoire magnétique est considérablement réduite, ce qui entraîne une très faible dissipation d'énergie dynamique. Bien que les mémoires non volatiles telles que la MRAM sont connues pour réduire ou éliminer la puissance statique, le gain supplémentaire de la dissipation de puissance dynamique est également d'une importance considérable, en particulier pour les applications où une réécriture fréquente des bits de mémoire a lieu pendant le fonctionnement, par exemple, dans des architectures de mémoires cache ou dans la logique en mémoire [42].

Bien qu'il en soit encore au stade expérimental et qu'il faille encore améliorer sa conception, la MeRAM devrait améliorer l'évolutivité des JTM en termes de miniaturisation, notamment en diminuant d'une manière sensible le courant d'écriture.

3.6. Mécanisme d'écriture SOT : Spin Orbit Torque

La méthode d'écriture SOT (Spin Orbit Torque) [43] est l'approche la plus récente parmi les méthodes d'écriture des MRAMs. Elle a été développée pour réduire les problèmes observés dans les STT-MRAM. Contrairement à l'approche STT, cette nouvelle technique utilise une structure à trois terminaux pour séparer les chemins de lecture et d'écriture, comme illustré à la Figure II-17. L'effet physique responsable du renversement de la magnétisation de la couche de stockage n'est pas encore complètement compris. Selon certains auteurs, l'effet Rashba [44] ou l'effet de spin de Hall [45] pourrait expliquer le changement de magnétisation de la couche de stockage.

Contrairement à la STT-MRAM, la SOT-MRAM sépare intrinsèquement les chemins de lecture et d'écriture et permet un courant d'écriture symétrique entre les deux états de la JTM. Par conséquent, la stabilité de la lecture est améliorée, ce qui réduit fortement la possibilité d'écrire un bit au cours d'une opération de lecture. En outre, les concepteurs peuvent optimiser la lecture et l'écriture séparément. Par ailleurs, la taille de cellule des SOT-MRAM est supérieure à celle de STT-MRAM en raison de sa structure à trois terminaux. La SOT-MRAM étant une technologie jeune comparée à d'autres technologies MRAM, des recherches supplémentaires sont nécessaires pour en optimiser les jonctions. Tout comme la MeRAM, cette technologie devrait permettre d'atteindre les mêmes performances que les SRAM [46].

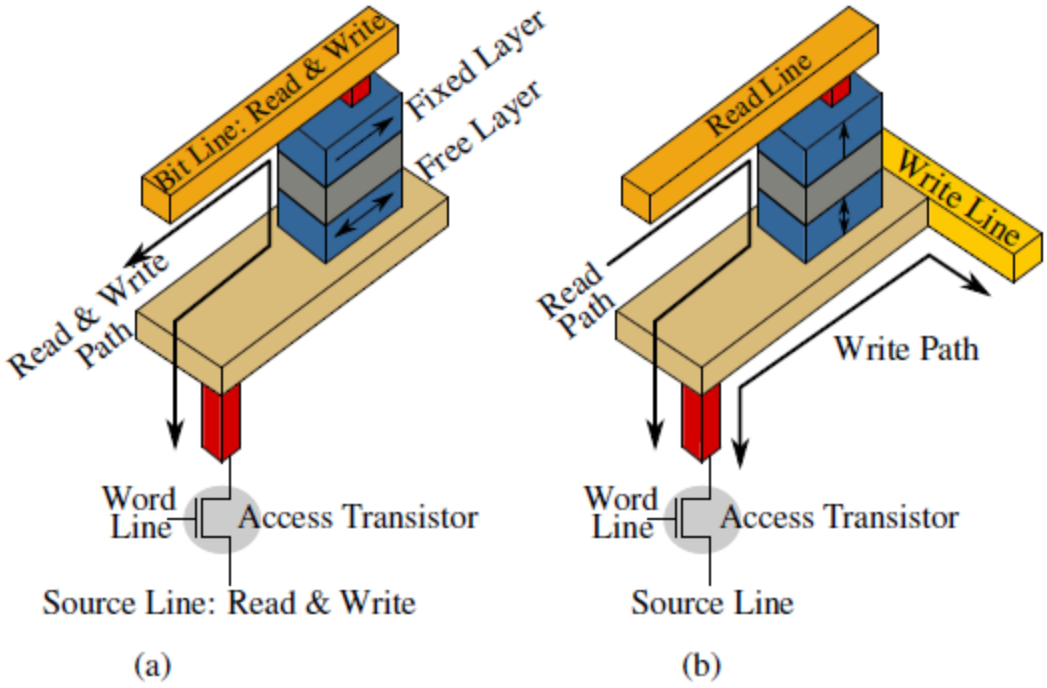


Figure II-17 : Spin Orbit Torque- MRAM (a) Jonction STT (b) Jonction SOT

4. Conclusion

Ce chapitre présente les principaux phénomènes liés à la magnétorésistance (effet MR), qui ont conduit au développement de la technologie MRAM. Comme décrit ci-dessus, un matériau constitué de deux couches ferromagnétiques (FM) séparées par une couche non ferromagnétique (NM) présente une faible résistance électrique lorsque les moments magnétiques des couches FM sont parallèles et une grande résistance électrique quand ils sont antiparallèles. Ainsi, il a été observé que le transport des électrons dépend du spin en raison des interactions entre le spin des électrons et les propriétés magnétiques du matériau. Cette découverte a marqué le début de la spintronique. Les avancées sur l'effet MR ont permis d'atteindre un ratio de plus de 100% entre la résistance la plus faible et la résistance la plus élevée des JTM. En conséquence, les technologies MRAM ont émergé et des investigations intensives sont en cours pour améliorer leurs performances. En raison de sa méthode d'écriture commandée en tension, la jonction MeRAM nécessite un courant d'écriture très faible par rapport aux autres technologies MRAM. Par conséquent, une très grande évolutivité est attendue. La STT-MRAM et la SOT-MRAM affichent presque les mêmes performances globales et sont de très bons candidats pour faire partie de la hiérarchie mémoire des systèmes sur puce. Quant à la technologie TAS-MRAM, elle est la plus fiable grâce à la structure des jonctions qui permet une excellente stabilité thermique, puis une très bonne

conservation des données. Le Tableau II-1 présente un résumé des différentes technologies MRAM et leur niveau de maturité.

Bien que la MeRAM et la SOT-MRAM affichent des performances très prometteuses, ils se situent toujours au niveau expérimental et doivent encore être développés. Au contraire, TAS-MRAM et STT-MRAM sont bien matures puisque des puces existent déjà dans le commerce.

Les chapitres suivants porteront principalement sur deux technologies MRAM utilisés durant cette thèse. A savoir la TAS-MRAM et la STT-MRAM. L'utilisation de ces MRAM pour la conception de blocs de sécurité primitifs sera abordée. Le Chapitre III et le Chapitre IV porteront sur la conception et l'expérimentation des générateurs de nombres véritablement aléatoires (*TRNG pour True random number generator*) à base de STT-MRAM et de TAS-MRAM et le Chapitre V sur les fonctions physiques non clonable (*PUF pour physically uncloble functions*) à base de TAS-MRAM.

Tableau II-1 : Les Technologies MRAM

Technologie	Taille de la cellule (F^2)	Courant d'écriture	Endurance	Maturité	Avantages/Inconvénients
Toggle [47][48][49]	50	> 30 mA	10^{15}	Commercialisé	*Maturité* *Consommation*
TAS - MLU [26][50][51]	<50	1-10 mA	10^{15}	Commercialisé	*Fiabilité* *Temps d'accès*
STT [52][34]	10-20	> 50 uA	> 10^{16}	Commercialisé	*Consommation* *Fiabilité*
MeRAM [38][53][54]	<10	1-10 uA	> 10^{16}	Prototype	*Consommation* *Maturité*
SOT [55][56][57]	>50	< 100 uA	> 10^{16}	Test chip	*Consommation* *Maturité*

Chapitre III Générateur de nombres véritablement aléatoires à base de STT-MRAM

1. Introduction

Un générateur de nombres aléatoires idéal est une construction mathématique qui génère des nombres aléatoires indépendants et uniformément répartis [58]. En pratique, les générateurs de nombres aléatoires sont divisés en deux catégories avec possibilité d'hybridation entre les deux : les générateurs de nombres véritablement aléatoires (True Random Number Generators - TRNG) et les générateurs de nombres pseudo-aléatoires (Pseudo Random Numbers Generators - PRNG). La Figure III 1 présente une classification des RNG.

Les PRNG utilisent des algorithmes déterministes pour générer des nombres qui, en apparence, semblent aléatoires mais qui en réalité ne le sont pas. Ils sont facilement intégrables dans des circuits numériques et peuvent fonctionner à des débits élevés. Cependant, théoriquement, les PRNG ne garantissent pas l'imprévisibilité des nombres générées. De ce fait, ils utilisent souvent une graine d'initialisation issue d'un TRNG pour augmenter cette imprévisibilité.

Les TRNG quant à eux ne sont pas de nature algorithmique : il s'agit de constructions électriques ou mécaniques qui extraient l'aléa à partir d'un phénomène le plus souvent physique et possédant des propriétés aléatoires (bruit thermique, phénomènes quantiques, etc.). Leur débit est en général faible et dépend fortement du phénomène dont ils tirent l'aléa (fréquence à laquelle une main peut jeter un dé, spectre fréquentiel du bruit électronique...) et par les contraintes liées à la technique d'extraction ainsi qu'à son implémentation. La qualité statistique des suites générées dépend à la fois de la qualité de la source d'aléa et de la technique avec laquelle on l'extrait. Cependant, en pratique des défauts statistiques liés à l'implémentation peuvent être observés. Néanmoins, lorsqu'ils sont correctement conçus, les TRNG permettent de générer des suites imprévisibles.

Une source importante de bruit dans les circuits intégrés est l'imprécision de leur processus de fabrication, appelée variation de processus [59]. Cette source peut être exploitée pour concevoir des TRNG car elle fait varier les propriétés physiques et, par conséquent, les caractéristiques électriques ou magnétiques d'un dispositif dans chacune de ses instances fabriquées. Par exemple, une caractéristique importante des jonctions tunnel magnétique (JTM) des MRAM est le courant de seuil de commutation, qui est le courant minimal requis

pour changer l'état d'une JTM [60]. En raison de la variation des procédés de fabrication, les attributs physiques de la JTM (tels que l'épaisseur et la section transversale de l'oxyde de tunnel) et par conséquent le courant de seuil dans chaque JTM d'une MRAM fabriquée est légèrement différent. Cette variabilité entraîne un comportement de commutation stochastique dans les JTM tout en appliquant un courant proche du seuil.

Ce chapitre présente une nouvelle méthode de conception de TRNG basée sur les STT-MRAM. Dans un premier temps, les notions de base sur les TRNG sont introduites (section 2), ensuite la section 3 définit les TRNG à travers un état de l'art sur les TRNG dans les circuits numériques. La section 4 aborde la conception de TRNG basé sur les STT-MRAM, et enfin la section 5 conclut ce chapitre.

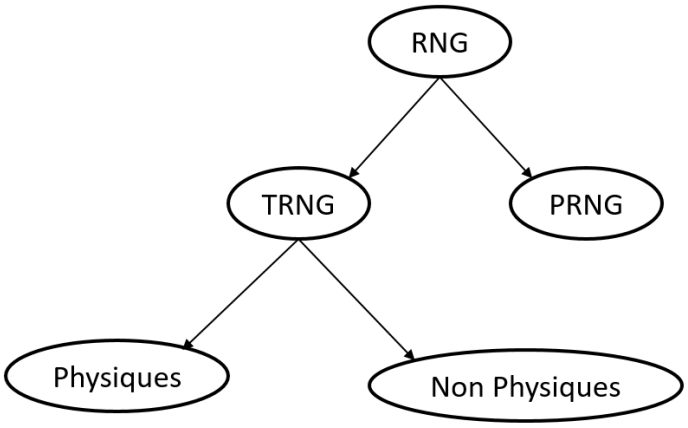


Figure III-1 : Classification des générateurs de nombres aléatoires

2. Les générateurs de nombres pseudo aléatoires (PRNG)

Un générateur de nombres pseudo-aléatoires (PRNG) génère une séquence binaire présentant certaines propriétés du hasard telle qu'aurait une séquence réellement aléatoire, mais qui, théoriquement est prévisible. La qualité des bits générés est évaluée à l'aide de tests statistiques, et son niveau de sécurité grâce à la cryptanalyse de l'algorithme utilisé. De tels générateurs sont faciles à concevoir et permettent d'obtenir un débit assez élevé tout en produisant des suites qui ont de bonnes propriétés statistiques. Par conséquent, ils sont très adaptés pour les applications qui ne nécessitent pas l'imprévisibilité des suites (la simulation numérique par exemple).

2.1. Principe et fonctionnement des PRNG

Les séquences générées par un PRNG ne sont pas tout à fait aléatoires dans le sens où elles sont déterminées par un ensemble de paramètres initiaux. L'algorithme utilisé peut donc être complètement caractérisé par la succession déterministe de ses états internes. La période d'un PRNG est définie par le nombre de tous ses états, généralement exprimé en bits. Il est assez facile de construire des algorithmes avec des périodes suffisamment longues pour la plupart des applications. D'autres part, ces algorithmes sont construits de manière à fournir des nombres qui ont de bonnes répartitions statistiques. La facilité d'implémentation, la rapidité et la multitude d'algorithmes efficaces connus font que ce type de générateur est très utilisé dans les applications telles que la simulation numérique, les automates de jeux, etc.

Les PRNG peuvent également être utilisés dans les application cryptographiques nécessitant des suites imprévisibles, lorsque leur graine d'initialisation est renouvelée fréquemment à l'aide d'un TRNG. Ceci permet de profiter à la fois de la rapidité des PRNG et de l'imprévisibilité des TRNG. Il est toutefois important de noter que si un PRNG est toujours initialisé au même état, il produira toujours la même séquence de sortie. Il est donc essentiel que la graine d'initialisation soit renouvelée avec une période inférieure à celle de l'algorithme.

2.2. Conception de PRNG

Généralement les PRNG sont conçus en utilisant des Registres à Décalage à Rétroaction (*FSR, Feedback Shift Register*). Les FSR ont été largement étudiés dans la littérature [61][62]. Un FSR de longueur n est constitué d'une paire de fonction (F, f) où f est une fonction booléenne définie sur n variables et F une fonction de $n-uplet$ également définie sur n variables telle que :

$$F(x_1, x_2, \dots, x_n) = (x_2, x_3, \dots, x_n, f(x_1, x_2, \dots, x_n))$$

En effet, F décrit la façon d'obtenir le nouvel état en fonction de l'état précédent tandis que f est utilisé pour définir le $n-$ème terme du nouvel état. On appelle donc F la fonction de l'état suivant et f la fonction de rétroaction.

A partir d'un état initial $(s_1, s_2, \dots, s_n)$, on obtient une succession d'état suivant :

$$F(s_1, s_2, \dots, s_n) = (s_2, s_3, \dots s_n, s_{n+1})$$
$$F(s_2, s_3, \dots, s_{n+1}) = (s_3, s_4, \dots s_{n+1}, s_{n+2})$$
$$F(s_3, s_4, \dots, s_{n+2}) = (s_4, s_5, \dots s_{n+2}, s_{n+3}) \qquad (III-1)$$

La séquence binaire produite par un FSR est alors $(s_1, s_2, \dots s_n, s_{n+1}, \dots)$.

Si la fonction de rétroaction f est linéaire on parle alors de Registre à Décalage à Rétroaction Linéaire (LFSR *pour Linear Feedback Shift Register*), sinon on parle de Registre à Décalage à Rétroaction Non Linéaire (NLFSR *pour Non Linear Feedback Shift Register).*

Les LFSR ont été largement utilisés dans le passé en raison de leur bonne propriété statistique, leur grande période de séquence binaire et de la simplicité de leur implémentation. Toutefois ils ne sont plus conseillés dans leur configuration initiale pour des applications cryptographiques. En effet il a été démontré à travers l'algorithme de *Berlekamp-Massey* [63] qu'en connaissant $2n$ bits successifs d'une séquence on peut facilement retrouver le LFSR qui les a engendrés. Cependant, il existe des solutions pour contourner l'effet des propriétés de linéarité des LFSR et ainsi renforcer la sécurité des PRNG à base de LFSR. Plusieurs conceptions basées sur les trois méthodologies suivantes sont fréquemment employées :

1. Associer une fonction non linéaire aux sorties de plusieurs LFSR ;
2. Utiliser une fonction de filtrage non linéaire basé sur le contenu d'un seul LFSR ;
3. Utiliser plusieurs LFSR en parallèle ou une horloge externe qui peut provenir d'un autre LFSR.

2.2.1. PRNG à combinaison non linéaire de plusieurs LFSR

Une technique générale pour supprimer la linéarité inhérente des LFSR est d'utiliser plusieurs LFSR en parallèles. La sortie est générée comme une fonction non linéaire f des sorties des composantes LFSR ; cette construction est illustrée par la Figure III-2, et la fonction f est appelée fonction de combinaison.

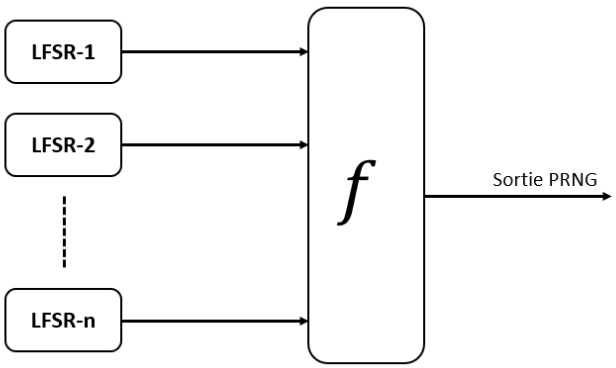


Figure III-2 : Combinaison de LFSR par une fonction booléenne

2.2.2. PRNG à base de filtrage non linéaire d'un LFSR

Ce type de PRNG masque les linéarités du LFSR en utilisant une fonction de sortie non linéaire sur quelques bits. A chaque étape, la fonction de sortie notée f prend en entrée t $bits$ de l'état interne du LFSR ; cette technique est illustrée par la Figure III-3. Les t bits ne sont généralement ni consécutifs, ni uniformément espacés dans le registre. La fonction de sortie produit un seul bit, et représente la sortie du LFSR. La principale difficulté lors de l'utilisation de ce générateur est de choisir une fonction f adéquate. Habituellement f est décrite comme une table de valeurs ou comme un polynôme, ce type de fonction est appelé fonction booléenne.

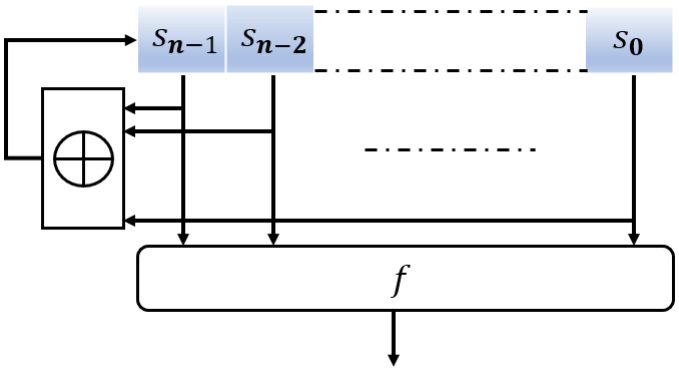


Figure III-3 : PRNG filtré à base de LFSR

2.2.3. PRNG par rétrécissement à base de LFSR

Dans cette catégorie de générateur, on trouve deux variantes, le générateur par rétrécissement (*shrinking generator*) et le générateur par auto-rétrécissement (*self shrinking generator*) :

Le générateur par rétrécissement a été introduit en 1994 par Coppersmith et al. [64]. Il est basé sur la combinaison de deux LFSR. Le premier, le registre A, génère les bits de sortie. L'autre, le registre S, contrôle les sorties. A et S sont tous les deux synchronisés sur l'horloge du système. Si le bit du registre S est à 1 alors le bit du registre A est envoyé en sortie du générateur. Si le bit de S est à zéro, aucun bit n'est envoyé en sortie et le générateur avance d'un coup d'horloge. Le gros désavantage de cette approche est une diminution du taux de génération ou débit qui devient irrégulier par la même occasion. La Figure III-4 présente le principe de cette technique de générateur.

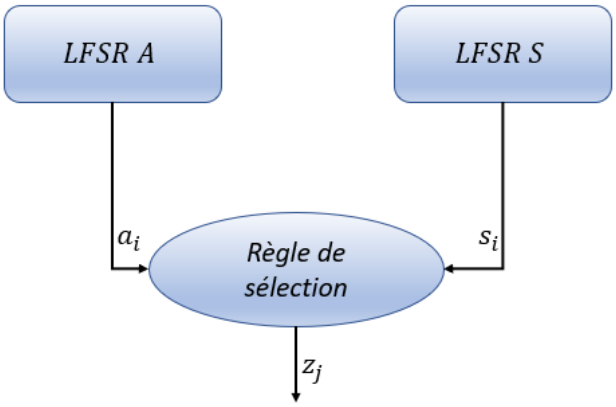


Figure III-4 : Générateur par rétrécissement

Malgré sa simplicité, il n'y a actuellement aucune attaque connue meilleure que la recherche exhaustive (attaque par force brute) lorsque les polynômes de rétroaction sont secrets. Cependant, si les polynômes de rétroaction sont connus, l'attaque la plus connue nécessite moins de $A \cdot S$ bits de sortie [65].

Le générateur par auto-rétrécissement : est une version modifiée du générateur par rétrécissement et il est encore plus simple, il a été présenté par W. Meier et O. Staffelbach dans [66]. A la différence du générateur par rétrécissement qui utilise un deuxième LFSR pour contrôler la sortie du premier, le générateur d'auto-rétrécissement utilise les bits de sortie alternés d'un seul LFSR pour contrôler sa sortie finale. La Figure III-5 illustre le fonctionnement de cette technique. La procédure de synchronisation de ce type de générateur est la suivante :

1. Echantillonner la sortie du LFSR par paire de bits ;
2. Si la paire est « 10 », la sortie finale est un « 0 » ;
3. Si la paire est « 11 », la sortie finale est un « 1 » ;
4. Sinon, la sortie n'affiche rien ;
5. Retournez à l'étape 1.

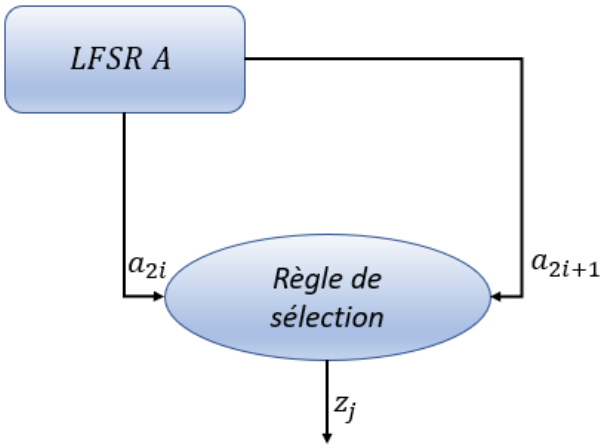


Figure III-5 : Générateur par auto-rétrécissement

En dépit de cette similitude, le générateur par auto-rétrécissement a montré une résistance à la cryptanalyse supérieure au générateur par rétrécissement [67].

3. Les générateurs de nombres véritablement aléatoires (TRNG)

Les générateurs de nombres véritablement aléatoires (TRNG) produisent des nombres qui ne suivent aucun ordre ou séquence en particulier. S'ils sont bien conçus, même une connaissance absolue de leur architecture et de leur fonctionnement interne ne devrait permettre de prédire leurs bits de sortie. Les TRNG sont donc basés sur des phénomènes physiques purement aléatoire (bruit thermique, métastabilité, etc.). Le principe générique d'un TRNG est présenté sur la Figure III-6. Il est composé d'une source d'aléa, d'un extracteur d'entropie et éventuellement d'un post-traitement. La source d'aléa correspond au processus produisant l'aléa. L'extraction d'entropie est le processus qui permet d'extraire les propriétés aléatoires de la source pour générer des nombres aléatoires. Le post-traitement consiste à corriger les défauts statistiques que l'on observe en général à la sortie de l'extracteur d'entropie. Cela est due au fait que la source d'aléa et l'extracteur d'entropie ne sont jamais parfaits.

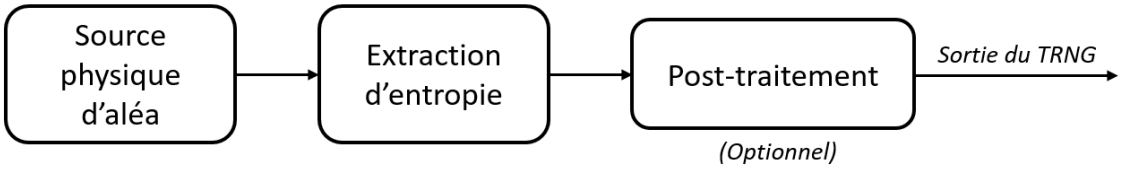


Figure III-6 : Structure générique d'un TRNG

Un TRNG doit générer des séquences aléatoires sous un format défini : entier, binaire, caractère, etc. Dans notre contexte, nous allons parler plus précisément de séquences binaires. Chaque séquence peut être représentée par le lancer d'une pièce de monnaie non biaisé en l'air, le résultat de chaque lancer a une probabilité de 50% d'être pile et une probabilité de 50% d'être face. Et chaque lancer doit être indépendant des autres.

3.1. Les sources d'aléa et les différentes techniques d'extraction d'aléa dans les circuits électroniques

Les sources d'aléa constituent la partie analogique d'un TRNG. Chaque TRNG utilise une source aléatoire spécifique qui est basée sur un phénomène physique imprévisible. Dans les technologies sur silicium, les sources d'aléa physiques sont limitées et semblent toutes liées au bruit. Ce dernier est omniprésent dans les circuits électroniques. Il agit sur les signaux manipulés de différentes manières (fluctuations de la tension du signal analogique, gigue, etc.). Nous allons dans cette section donner quelques exemples de techniques habituellement utilisée dans la littérature pour la conception des TRNG.

3.1.1. Amplification directe du bruit

Le bruit des composants analogiques (résistances, diode Zener, etc.) peut être utilisé comme une source d'aléa. Le principe du TRNG utilisant ce type de bruit consiste à amplifier les variations de la tension de sortie du signal analogique issu de la source de bruit et à les capter grâce à d'autres composants analogiques (amplificateur opérationnels, comparateurs, etc.).

Un exemple de cette technique développée par *Intel* et dont l'évaluation a été effectué par la société *Cryptography Research* est présenté dans [68]. La Figure III-7 présente le diagramme de blocs de ce TRNG. La source d'aléa ici est le bruit thermique ou bruit de Johnson aux bornes d'une résistance. Pour s'affranchir des sources de bruit externes présentant souvent des corrélations, les concepteurs de ce TRNG ont réalisé une source différentielle en soustrayant le bruit obtenu aux bornes de deux résistances identiques. Ainsi, les perturbations induites par le bruit d'alimentation ou les sources électromagnétiques environnantes ou encore la dépendance à la température sont fortement réduites. Le bruit différentiel ainsi obtenu est appliqué à un VCO (oscillateur contrôlé en tension) de basse fréquence, qui est lui-même échantillonné par un oscillateur haute fréquence à intervalles réguliers pour générer des nombres aléatoires. Un correcteur de Von Neumann (cf. section 3.2.1 de ce chapitre) est associé pour corriger les biais à la sortie du mixeur (Super Latch).

Parmi les principes de TRNG publiés basés sur l'amplification du bruit de composants analogiques on peut également citer [69][70][71]. Le principal défaut de ces techniques est

qu'elles font appel à des composants analogiques qui nécessitent un soin particulier d'implémentation dans les circuits intégrés et qui ont des débits assez faible (de l'ordre de la centaine de Ko/s [70][71]).

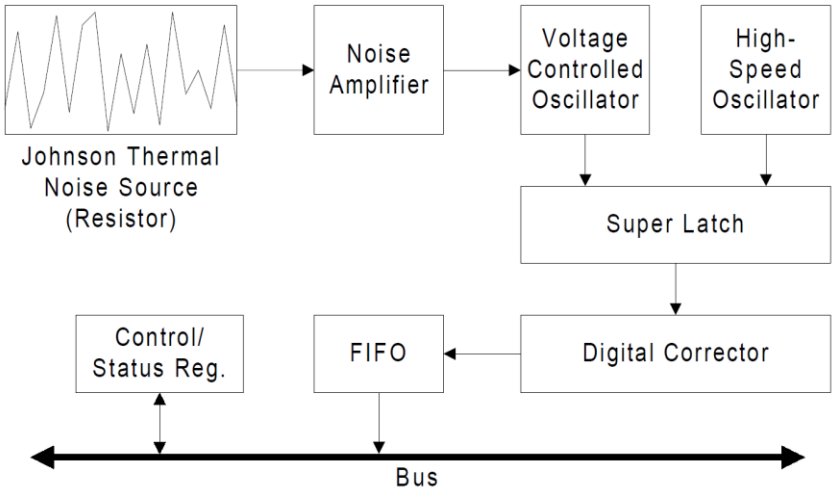


Figure III-7 : Schéma de principe du TRNG développé par Intel et discuté dans [68]

3.1.2. L'Extraction de la gigue (Jitter)

La gigue (*Jitter en anglais*) est un phénomène de fluctuations temporelles (fluctuations au niveau des transitions) d'un signal périodique autour de sa valeur nominale. Elle est due au bruit dans les composants électroniques. Elle peut être mesurée sur un signal numérique dans le domaine fréquentiel ou temporel et peut être extraite par échantillonnage. Les signaux manipulés peuvent être générés à l'intérieur du circuit, et l'extraction se base uniquement sur des composants numériques (bascules D, latchs, compteurs, etc.). L'amplitude de la gigue est égale à la valeur du plus grand retard moins la valeur de la plus grande avance des transitions (toujours en référence à la valeur nominale du signal périodique). Contrairement aux applications de communications où l'on cherche à réduire l'effet de la gigue, l'objectif ici est d'augmenter la gigue afin d'avoir une plage aléatoire plus grande.

Traditionnellement il existe deux méthodes pour générer des nombres aléatoires à partir de la gigue d'un signal périodique. Soit échantillonner le signal au plus près de sa transition comme présenté sur la Figure III-8 ; Soit procéder à une accumulation de la gigue en comptant le nombre de transition du signal sur un intervalle de temps donné (les variations temporelles dues à la gigue s'additionnent dans le temps et influencent sur le nombre de transitions pour un intervalle de temps suffisamment long).

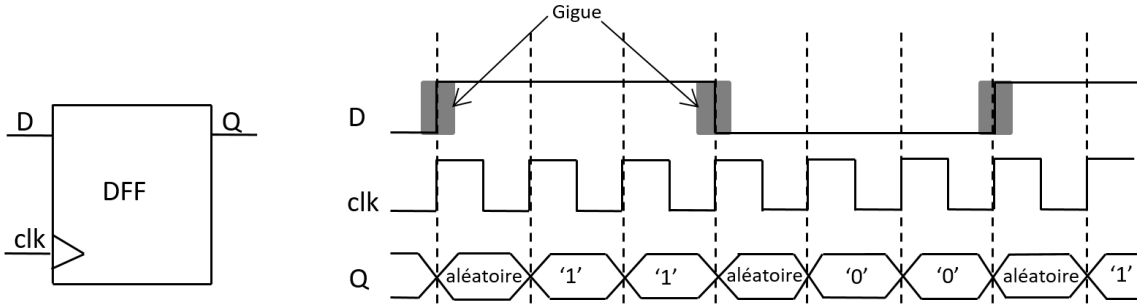


Figure III-8 : Echantillonnage de la gigue d'un signal périodique par un signal d'horloge

La méthode d'accumulation de la gigue est la plus utilisée dans la littérature. Elle a d'abord été proposé par C. Fairfield et al. [72] puis repris largement et améliorée par la communauté scientifique. Le principe de fonctionnement est décrit dans la Figure III-9. Un signal haute fréquence (Fd) de 8MHz, généré grâce à un oscillateur embarqué, est échantillonné par un signal basse fréquence (Fc) ajustable. La valeur échantillonnée dépend de l'accumulation des variations temporelles des deux signaux Fd et Fc.

Les principaux inconvénients de cette technique sont liés à la faiblesse des débits ainsi qu'au fait que les variations dues aux bruits déterministes sont plus importantes que celles dues aux bruits aléatoires [73] ce qui peut induire des défauts statistiques sur les bits générés.

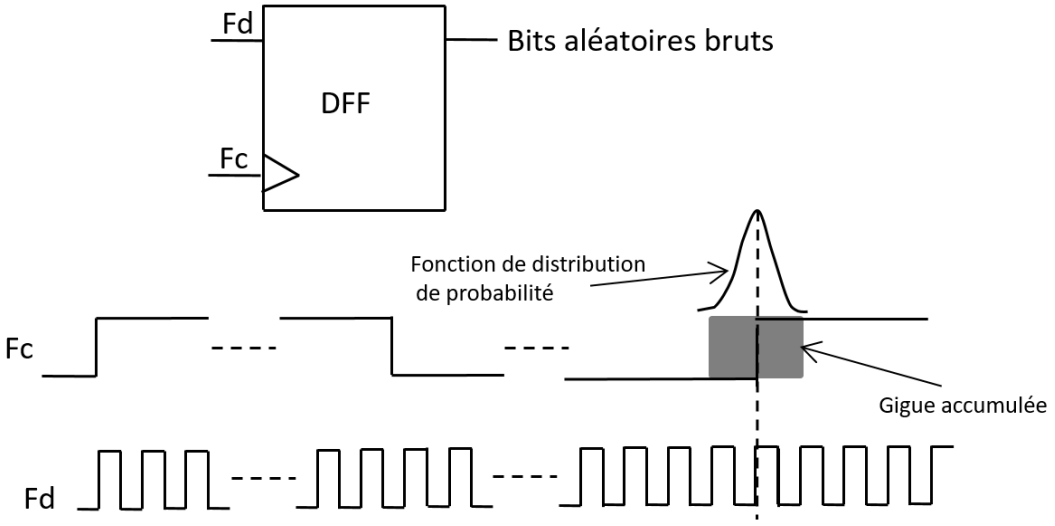
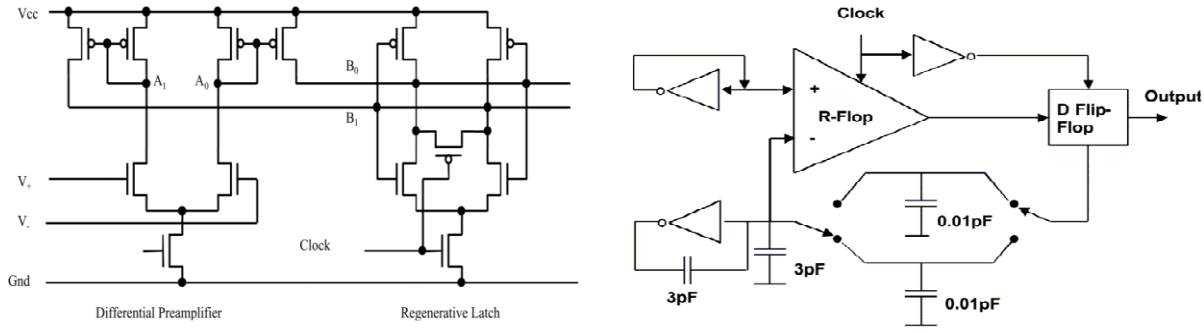


Figure III-9 : TRNG basé sur l'accumulation de la gigue proposée dans [72]

3.1.3. Résolution d'états métastables

La métastabilité désigne la capacité d'un système électronique à rester pendant un temps indéterminé dans un état d'équilibre instable [74]. Par exemple, si on reboucle la sortie d'une porte logique inverseuse à son entré, la tension de sortie peut stagner sur une valeur intermédiaire entre le niveau logique haut et le niveau logique bas [75][76]. Les techniques exploitant la métastabilité pour générer des nombres aléatoires ont pour principe commun de placer un système bistable dans une situation intermédiaire de telle sorte que son basculement à un état stable soit grandement influencé par le bruit.

En 2002, les auteurs dans [77] ont présenté le premier TRNG basé sur le principe de la métastabilité. Ce TRNG utilise une circuit analogique personnalisée qui est contrôlée afin de forcer l'état métastable. Le circuit de contrôle de la métastabilité est appelé *R-flop.* La Figure III-10 présente le circuit personnalisé complet *R-flop* ainsi que la méthode d'extraction des nombres aléatoires. Le circuit *R-flop* est composé d'un amplificateur différentiel et d'un latch. L'étage d'entrée analogique de l'amplificateur différentiel pilote la cellule bistable (latch) avec une petite différence de courant. Ainsi, la sortie prend une valeur logique « 1 » ou « 0 », lorsque le latch est cadencé avec des niveaux de tension très bas. Dans le circuit présenté, si $(\Delta V = V_+ - V_-)$ à l'entrée de l'amplificateur est inférieur à $0{,}1\ mV$, alors ΔV à l'entrée de la cellule bistable est inférieur à l'incertitude de tension du bruit thermique. Le circuit *R-flop* génère des nombres aléatoires avec des biais provoqués par des variations de température qui doivent être ajustées à l'aide d'un condensateur à commutation de circuits suivi d'une bascule D. Le bloc extracteur d'entropie est illustré à la Figure III-10 (b) [77]. Ensuite, un circuit de post-traitement génère la séquence de sortie finale de bits aléatoires.



(a) ciruit R-flop (b) TRNG basé sur la métastabilité du circuit R- flop

Figure III-10 : TRNG basé sur la métastabilité présenté dans [77]

Parmi les techniques exploitant la résolution de situation métastables on peut également citer [78] et [79]. La difficulté principale avec ces techniques réside dans la quantification de l'influence du bruit sur la résolution des états métastables.

3.2. Méthodes de post-traitement des TRNG

La sortie brute du TRNG peut présenter de mauvaises propriétés statistiques telles que de longues séquences de « 1 » ou de « 0 » ou un motif répétitif de bits. Ce biais peut être causé par un bruit déterministe provenant de sources externes telles que des variations de la tension d'alimentation ou des effets de diaphonie (crosstalk) provenant des circuits logiques environnantes. Dans ce cas, un algorithme de post-traitement peut être utilisé afin d'augmenter l'entropie. Le bloc de post-traitement n'est pas obligatoire si la sortie brute du TRNG présente de bonnes propriétés statistiques. Les techniques de post-traitement peuvent être arithmétiques ou cryptographiques [80]. Les méthodes de post-traitement arithmétique typiques utilisent une fonction de réduction simple du biais, c'est le cas du correcteur de Von Neumann [81][82] et des filtres de parité (XOR).

3.2.1. La méthode de post-traitement Von Neumann

C'est la méthode de correction la plus utilisée car elle nécessite peu d'implémentation matériel. Il s'agit d'une simple procédure de post-traitement qui donne des résultats parfaitement non biaisés [81]. Elle consiste à regrouper consécutivement les bits d'une séquence binaire en sous-séquences de paires de bits non chevauchantes et à générer des sorties comme suit :

- Si la paire de bits est « 00 » ou « 11 », elle est rejetée (non prise en compte).
- Si la paire de bits est « 01 » ou « 10 », la sortie est le premier bit de la paire.

Supposons que les bits d'une séquence donnée aient un biais ε ; cela signifie que pour chaque bit « x » de la séquence, la probabilité P est :

$$P(x = 0) = \frac{1}{2} + \varepsilon \; ; \; P(x = 1) = \frac{1}{2} - \varepsilon$$

En appliquant la méthode de correction de Von-Neumann à cette séquence on a une sortie « y » avec la probabilité d'avoir « 0 » comme suit :

$$P\,(y\; =\; 0) = \frac{P(\text{"01"})}{P(\text{"01" ou "10"})} = \frac{\left(\frac{1}{2}+\varepsilon\right).\left(\frac{1}{2}-\varepsilon\right)}{\left(\frac{1}{2}+\varepsilon\right).\left(\frac{1}{2}-\varepsilon\right)+\left(\frac{1}{2}-\varepsilon\right).\left(\frac{1}{2}+\varepsilon\right)} = \frac{1}{2} \qquad (III-2)$$

L'inconvénient de cette méthode est le temps d'attente inconnu jusqu'à ce que les bits aléatoires post-traités requis soient disponibles [81][83]. En effet, le débit de la sortie post-

traitée dépend du caractère aléatoire de la sortie brute. Il varie donc en fonction de l'entropie de la sortie brute.

3.2.2. Post-traitement basé sur les filtres de parité ou le OU-Exclusif (XOR)

Les filtres de parité sont couramment utilisés pour réduire le biais statistique d'une suite aléatoire. Le principe d'un filtre de parité consiste à effectuer l'addition en base 2 de n bits successifs pour générer un bit de sortie, ce qui correspond à une opération de OU exclusif (XOR). Le XOR est utilisé dans la plupart des algorithmes cryptographiques tels que le système de cryptage avancé (AES) et le système de cryptage des données (DES), il ajoute de la confusion et de l'aléa dans le processus de cryptage tout en étant réversible dans le processus de décryptage. Le biais de probabilité diminue lorsqu'une opération XOR est effectuée entre deux bits indépendants.

Supposons que deux bits indépendants « x_1» et « x_2 » aient des biais de probabilité « ε_1 » et « ε_2 », respectivement. Cela signifie que :

$$P(x_1 = 0) = \frac{1}{2} + \varepsilon_1 \; ; \; P(x_1 = 1) = \frac{1}{2} - \varepsilon_1$$

$$P(x_2 = 0) = \frac{1}{2} + \varepsilon_2 \; ; \; P(x_2 = 1) = \frac{1}{2} - \varepsilon_2$$

En appliquant un XOR entre les bits « x_1» et « x_2 » ($x_1 \otimes x_2$), la probabilité d'avoir un « 0 » en sortie est :

$$P(x_1 \otimes x_2 = 0) = P(x_1 = x_2 = 0) \cup P(x_1 = x_2 = 1)$$

$$P(x_1 \otimes x_2 = 0) = \{P(x_1 = 0)\, P(x_2 = 0)\} \cup \{\, P(x_1 = 1\, P(x_2 = 1)\}$$

$$P(x_1 \otimes x_2 = 0) = \left(\frac{1}{2} + \varepsilon_1\right).\left(\frac{1}{2} + \varepsilon_2\right) + \left(\frac{1}{2} - \varepsilon_1\right).\left(\frac{1}{2} - \varepsilon_2\right)$$

$$P(x_1 \otimes x_2 = 0) = \frac{1}{2} + 2\varepsilon_1\varepsilon_2$$

Comme « ε_1 » et « ε_2 » sont inférieurs à $\frac{1}{2}$, alors $2\varepsilon_1\varepsilon_2$ est inférieur à « ε_1 » et « ε_2 ».

Une généralisation de la formule est donnée comme suit :

$$P(x_1 \otimes x_2 \ldots \otimes x_n = 0) = \frac{1}{2} + 2^{n-1} \prod_{i=1}^{n} \varepsilon_i \qquad (III - 3)$$

On constate donc que le biais diminue en fonction du nombre de XOR utilisé. Cependant, le débit peut également être divisé par le nombre de XOR en fonction de la méthode d'implémentation.

D'autres méthodes de post-traitement arithmétique sont également largement utilisées, comme les LFSR vue dans la section 2.2. Les post-traitements généralement cryptographiques sont des fonctions de hachage ou un algorithme de chiffrement de données [80]. Les algorithmes de chiffrement par blocs, tels que AES, sont également couramment utilisés pour le post-traitement. Cela se trouve, par exemple dans le processeur Ivy bridge d'Intel [84].

3.3. Evaluation des générateurs de nombre aléatoires

Prouver qu'un RNG est aléatoire est généralement difficile. En effet, malgré le fait d'avoir une source d'aléa, le processus d'extraction de l'aléa et la conversion des signaux en données binaires peuvent provoquer un biais sur la distribution souhaitée et rendre l'exploitation des résultats difficile. L'utilisation d'un correcteur, permet de compenser cela, mais ceci en réduisant le débit à la sortie du générateur. Dans la pratique, le degré d'aléa se vérifie à travers des tests statistiques. En fonction de ces résultats, il est possible d'affirmer avec un degré de confiance que le générateur fournit des valeurs aléatoires non prédictives.

Par définition, un test fournit un moyen de réaliser des décisions quantitatives, basées sur une hypothèse de départ ou hypothèse nulle (H_0), laquelle peut être validée ou pas. Dans le contexte des générateurs aléatoires, l'hypothèse H_0 est définie comme étant *"La suite analysée n'est pas aléatoire".*

Une fois l'hypothèse initiale définie, il faut maintenant quantifier les résultats des tests en utilisant comme indicateur les valeurs de probabilité aléatoire (la $p - value$) [85]. L'idée générale est de prouver que l'hypothèse nulle n'est pas vérifiée car dans le cas où elle le serait le résultat observé serait fortement improbable. Il s'agit à ce titre d'une extension du principe de la preuve par l'absurde.

Ainsi l'idée de l'analyse quantitative utilisée implique d'assurer une hypothèse initiale, qui ne peut jamais être acceptée directement, mais qui peut être rejetée seulement par le test statistique. La $p - value$ est considérée comme la mesure à partir de laquelle les données plaident contre l'hypothèse initiale. Les seuils suivants sont généralement pris pour référence :

- $p \leq 0{,}01$: très forte présomption contre l'hypothèse nulle
- $0{,}01 \leq p \leq 0{,}05$: forte présomption contre l'hypothèse nulle
- $0{,}05 \leq p \leq 0{,}1$: faible présomption contre l'hypothèse nulle
- $p \geq 0{,}01$: Pas de présomption contre l'hypothèse nulle

Il existe différentes batteries de tests qui ont été développées pour évaluer la qualité des générateurs de nombres aléatoires. Ces tests standardisés sont basés sur l'analyse de

certaines propriétés statistiques, telles que l'uniformité de la distribution des nombres générés et la non répétabilité des séquences, permettant ainsi de quantifier le caractère non prédictif du générateur. Le principe de ces tests consiste à comparer des propriétés statistiques des suites testées avec celle d'une suite (théorique) qui aurait été générée par un générateur idéal. Les plus utilisés dans le domaine de la cryptographie sont le NIST SP 800-22 [86] proposés par l'organisme de standards technologiques américain NIST (*National Institute of Standards and Technology*), ainsi que DIEHARD, proposé par des universitaires.

3.3.1. Les Tests DIEHARD

Développés et publiés en 1995 par George Marsaglia, ces tests ont été écrits en langage C et Fortran puis inclus dans un CD-ROM [87], ainsi qu'une collection de fichiers contenant des bits aléatoires d'origine différente (construits en combinant les sorties de TRNG et PRNG).

Une décennie plus tard, Robert Brown [88] a produit une version étendue de la suite de tests de Diehard, appelée Dieharder. Le code a été réécrit et publié sous licence GNU, intégré à la bibliothèque statistique GNU et paramétrable (on peut faire varier facilement la taille de l'échantillon). De nouveaux tests ont été ajoutés et d'autres améliorées. Le package résultant est pris en charge par les distributions Linux traditionnelles.

3.3.2. Les Tests de NIST SP 800-22

Développée par l'Institut National de Standards et Technologie (*NIST pour National Institute of Standards and Technology*), la batterie de tests statistiques *NIST SP 800-22* comptait à l'origine 16 tests statistiques, ce nombre a été réduit à 15 dans une révision datant de 2001. La dernière révision *NIST SP 800 rev-1a* a été publiée en Avril 2010 [86]. Ces tests sont spécifiquement conçus pour des applications cryptographiques. Ils sont basés sur une synthèse d'algorithmes existants auxquels s'ajoutent d'autres tests basés sur de nouveaux algorithmes créés par l'équipe de sécurité informatique et d'ingénierie statistique du NIST. Cette suite de tests est devenue un standard bien reconnu dans le monde des tests de l'aléatoire et de la cryptographie.

Bien que ce soit à ce jour la suite de tests la plus complète et la plus évoluée, les auteurs préviennent qu'aucun test statistique ne permet de certifier qu'un générateur est approprié pour une application cryptographique donnée : ces tests ne doivent pas se substituer à la cryptanalyse, ils permettent néanmoins d'éliminer les candidats qui ne sont pas adaptés à une application donnée.

Le principe de ces tests consiste à examiner deux hypothèses : l'hypothèse nulle H_0, qui est que la suite examinée est aléatoire, et l'hypothèse alternative H_1, qui est que la suite

examinée n'est pas aléatoire. Les seuils de confiance sont définis en fonction des deux types d'erreurs rencontrées dans une telle évaluation : l'erreur de premier ordre α, la probabilité qu'une suite réellement aléatoire ne soit pas reconnue comme tel par le test statistique, et l'erreur de second ordre β, la probabilité qu'une suite non-aléatoire soit reconnue comme aléatoire par le test statistique. Chaque test est basé sur le calcul d'une quantité statistique qui est fonction des données et du type de test. Les résultats du test sont donnés sous forme de $p - value$, qui résume la force de l'évidence contre l'hypothèse nulle. Pour chaque test, chaque $p - value$ est la probabilité qu'un RNG parfait ait produit une séquence moins aléatoire que la séquence testée au vu du paramètre statistique évalué par le test. Si ($p - value \geq \alpha$), alors l'hypothèse nulle est acceptée (la séquence paraît aléatoire). α désigne alors le seuil de confiance du test, sa valeur est typiquement comprise entre 0.001 et 0.01 dans le cadre des applications cryptographiques. Une valeur de α égale à 0.001 indique qu'une séquence aléatoire parmi 1000 testées ne devrait pas avoir l'air aléatoire alors qu'elle l'est réellement. Si ($p - value \geq 0{,}001$) , la séquence est considérée aléatoire à 99.9% de chances. Si ($p - value \leq 0{,}001$), la séquence est considérée non aléatoire à 99.9% de chances également.

La liste des 15 tests est la suivante :

1. Le test de fréquence globale (monobit)
2. Le test de fréquence par blocs
3. Le test de répétition globale
4. Le test de répétition par blocs
5. Le test de rang de la matrice binaire
6. Le test spectral
7. Le test de non chevauchement de séquences
8. Le test de chevauchement de séquences
9. Le test universel de Maurer
10. Le test de la complexité linéaire
11. Le test Série
12. Le test de l'entropie approximée
13. Les sommes cumulatives
14. L'excursion aléatoire

15. Le variant sur l'excursion aléatoire.

Chaque test requiert plus ou moins de données pour son application, les plus exigeants nécessitant $10^9\ bits$ (comme le test universel de Maurer).

4. TRNG à base de STT-MRAM

La commutation des jonctions STT dépend de la direction, de l'amplitude et de la durée du courant appliqué aux JTM. Elle est intrinsèquement stochastique due aux fluctuations thermiques internes de la JTM [28][89]. En effet, plusieurs recherches théoriques et expérimentales ont prouvé qu'un courant polarisé en spin passant à travers une mince couche ferromagnétique induit son moment angulaire de spin au conducteur ferromagnétique. Par conséquent, cela entraine une précession des moments magnétiques de la couche pour aboutir à une commutation de l'alimentation de la couche lorsque le courant de spin est suffisamment important [90]. La précession de la magnétisation sous l'influence d'un courant de spin est illustrée par la Figure III-11 [91]. Les fluctuations thermiques de la magnétisation de la couche dues au passage du courant polarisé en spin font que l'état initial de la couche (déterminé par θ) diffère à chaque mesure. Ce qui entraîne une commutation stochastique des directions de magnétisation de la couche.

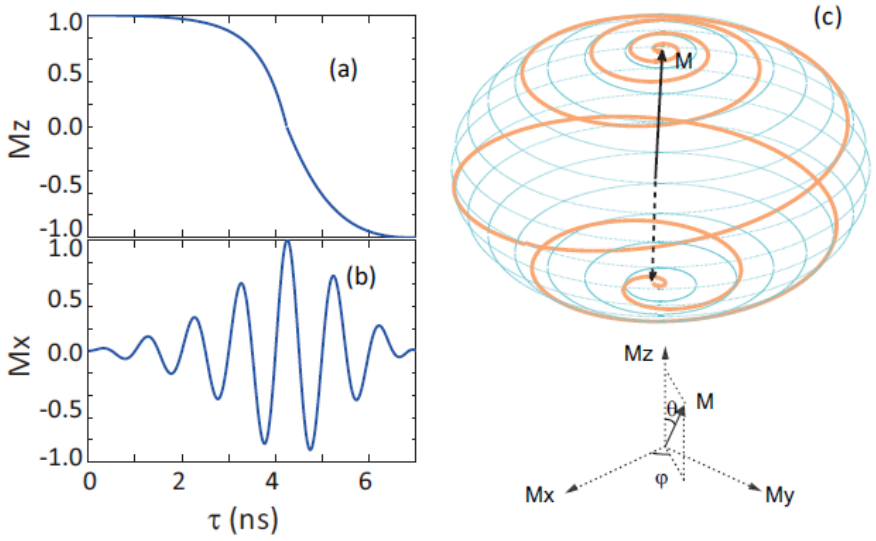


Figure III-11 : Précession de l'aimantation sous l'influence d'un courant de spin : dépendance temporelle de (a) Mz et (b) Mx, (c) processus d'inversion du moment magnétique. θ et ϕ représente l'état initial du moment magnétique de la couche libre [91].

Ainsi la durée de transition ou de commutation des jonctions STT n'est pas une valeur constante mais suit une distribution statistique. L'écriture d'une telle jonction est sujet à un

phénomène aléatoire. Des erreurs d'écriture peuvent se produire avec un courant d'écriture insuffisant ou une impulsion d'écriture courte. Tandis qu'une commutation inattendue peut se produire lors de la lecture [92].

Néanmoins, cette propriété peut être utilisée de manière appropriée comme source d'alea pour la génération de nombres aléatoires.

Plusieurs méthodes de génération de nombres aléatoires utilisant les jonctions STT ont déjà été proposées. La plupart d'entre elles est basée sur de la simulation.

Dans le cadre d'un projet européen (*GREAT*) dans lequel nous avions la possibilité de faire fabriquer des circuits de test basé sur les STT-MRAM, nous avons proposé de concevoir un petit circuit pour vérifier la faisabilité d'un TRNG basé sur les STT-MRAM.

Dans la suite de cette partie, nous présenterons une revue de quelques TRNG à base de MRAM, ensuite nous présenterons l'environnement de développement et le circuit conçu ainsi que les résultats après fabrication du circuit.

4.1. La revue des travaux dans la littérature

De nombreuses études ont été menées sur l'utilisation des MRAM dans la conception des générateurs de nombres aléatoire. La plupart de ces circuits sont basés sur des jonctions STT et des méthodes de générations de nombres utilisant l'amplitude et la durée de l'impulsion du courant d'écriture de la JTM.

La relation entre la probabilité de commutation (P_{sw}) et l'impulsion de courant appliquée aux JTM-STT est définie comme suit [93]:

$$P_{sw}(I_{sw}, t) = 1 - exp\left\{-\frac{t}{\tau_0}\, exp\left[-\Delta\left(1 - \frac{I_{sw}}{I_{c0}}\right)^2\right]\right\} \qquad (III-4)$$

Où t est la durée de l'impulsion de courant, τ_0 le temps d'attente ou le temps de non-réaction [94], Δ est le paramètre de stabilité thermique de la JTM (il définit la durée maximale pendant laquelle la JTM peut conserver son état magnétique initial après écriture contre l'agitation thermique), I_{sw} est le courant d'écriture et I_{c0} le courant critique de commutation.

La Figure III-12 par exemple présente la probabilité de commutation (AP -> P) d'une JTM STT en fonction de l'amplitude du courant d'écriture et de la durée de l'impulsion [95].

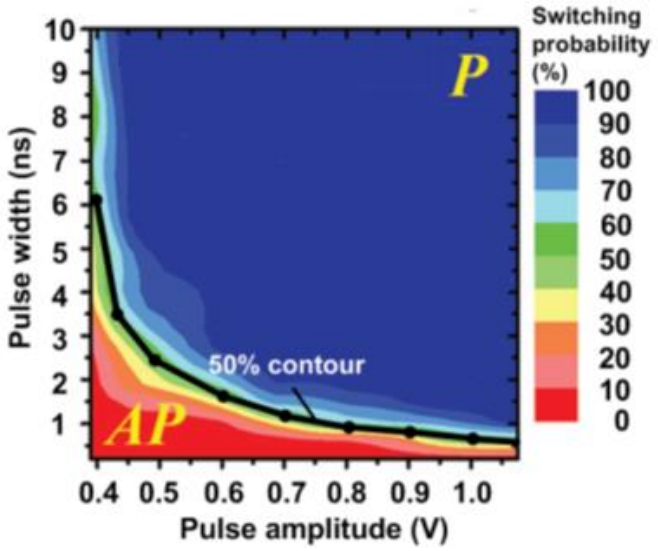


Figure III-12 : Probabilité de commutation d'une JTM-STT en fonction de l'amplitude du courant d'écriture et de la durée de l'impulsion. (Commutation AP -> P) [95]

Il est donc possible de déterminer la valeur de courant ainsi que de la durée d'impulsion pouvant donner une probabilité de commutation de 50% de la JTM. Une utilisation astucieuse de cette caractéristique peut potentiellement permettre de générer des nombres aléatoires. Il existe dans la littérature quelques travaux dans ce sens que nous allons présenter dans les lignes suivantes.

4.1.1. Méthode de génération « Spin dice »

Les auteurs dans [93] ont été les premiers à proposer un TRNG à base de jonctions STT appelé « spin dice », dans lequel les nombres aléatoires (RN) ou bits aléatoires (RB) sont générés en utilisant la nature stochastique de la commutation des jonctions STT. Le mécanisme de base pour la génération d'un RB est le même que celui pour l'écriture d'un bit de donnée dans la mémoire magnétique (STT-MRAM). La principale différence est que la probabilité de commutation est fixée à 50% pour la génération d'un RB, tandis qu'elle est de 100% dans le cas d'une utilisation mémoire.

Le principe de la méthode « spin dice » est présenté à la Figure III-13. Tout d'abord, une étape d'initialisation de la JTM à l'état parallèle (P) est effectuée à travers une impulsion de réinitialisation. L'amplitude et la durée de l'impulsion de réinitialisation sont suffisamment grandes pour initialiser la JTM correctement ($P_{sw} = 100\%$). Après cette étape, la JTM étant à l'état P, une deuxième impulsion est appliquée à la jonction dans le sens opposé, appelé impulsion d'excitation.

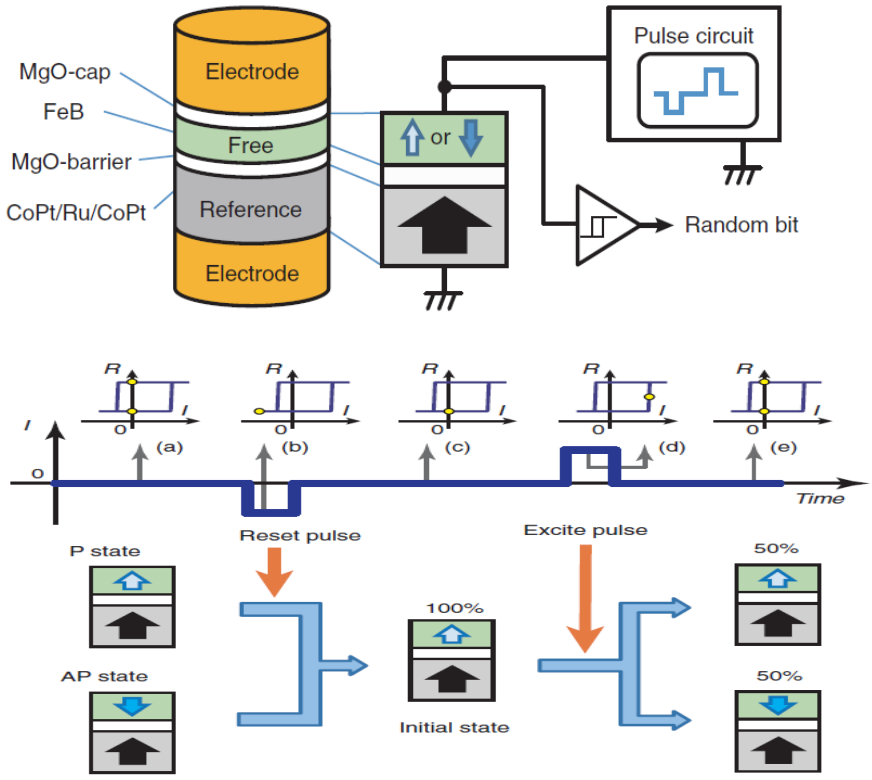


Figure III-13 : La méthode « spin dice ». Les impulsions de réinitialisation et d'excitation sont appliquées de manière séquentielle [93].

L'amplitude et la durée de l'impulsion d'excitation sont réglées pour satisfaire la probabilité de commutation de 50% ($P_{sw} = 50\%$) correspondant au point de bifurcation [96]. A ce point de bifurcation, l'agitation thermique induit une petite déviation aléatoire des moments magnétiques entrainant le basculement de la magnétisation totale de la couche de stockage à l'état AP ou P avec la même probabilité de 50%. Enfin, l'état final de la JTM est obtenu en mesurant sa résistance, donnant ainsi un bit aléatoire.

Dans la pratique, les Rbs générés par cette méthode, bien que présentant de bonnes propriétés statistiques nécessitent au moins un post-traitement arithmétique pour passer les tests d'évaluation. En effet, la probabilité fluctue autour de 50% dues aux variations thermiques et de courants. Après utilisation d'un post traitement arithmétique simple comme le XOR sur huit « Spin dice », les RB générés ont pu passer tous les tests de NIST-SP 800.

4.1.2. TRNG basé sur la STT-MTJ utilisant une boucle à verrouillage de probabilité à commande numérique

Pour remédier aux problèmes de fluctuation de la probabilité de sortie des Rbs, les auteurs dans [97] ont proposé une méthode de correction. En effet pour maintenir une entropie élevée, il faut une équiprobabilité entre les Rbs. Et pour garantir cette équiprobabilité il est primordial de contrôler régulièrement la distribution des Rbs générés.

Dans le même temps, Ils proposent également de réduire le temps nécessaire pour générer un Rb en supprimant l'étape de réinitialisation de la JTM. La méthode proposée consiste à écrire la JTM dans ses 2 états (AP -> P et P-> AP) avec une probabilité de 50%. Pour cela il faut d'abord lire la JTM et en fonction de la valeur lue, une impulsion d'excitation dans le sens inverse est alors appliquée à la JTM. Une comparaison des deux méthodes de génération est présentée sur la Figure III-14.

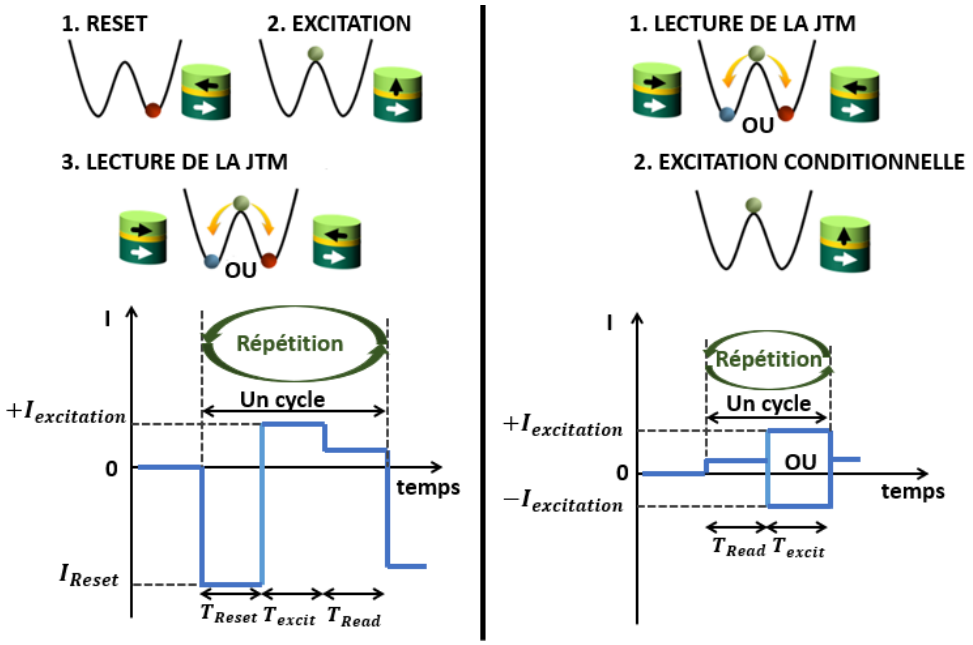


Figure III-14 : Méthodes de génération des Rbs (à gauche) la méthode avec réinitialisation de la JTM et (à droite) la méthode sans réinitialisation de la JTM proposé dans [97].

La méthode de correction de la probabilité est composée de deux générateurs de pulse haute résolution, une carte d'acquisition, un combineur de puissance ainsi qu'un arbre de polarisation. La correction se fait uniquement sur la durée de l'impulsion du courant. La Figure III-15 présente le principe général proposé, deux compteurs de $10\ bits$ sont utilisés pour calculer la probabilité de sortie de chaque séquence de $1\ Kbit$. La durée d'excitation dans le cas P->AP est ajusté avec un pas de $0.05\ ns$ en fonction de tous les autres paramètres pour chaque 1kbits généré. C'est donc un simple contrôle de rétroaction à un seul paramètre. Le

système de contrôle de la probabilité est géré au niveau logiciel. Une période d'initialisation appelé temps de verrouillage est nécessaire avant la convergence de la probabilité à 50%. Par contre La probabilité continue à fluctuée pour une correction sur des séquences de taille inférieure à $1\ Kbit$ alors que des corrections apportées sur des tailles de séquence beaucoup plus longue que $1\ Kbit$ augmente le temps de verrouillage.

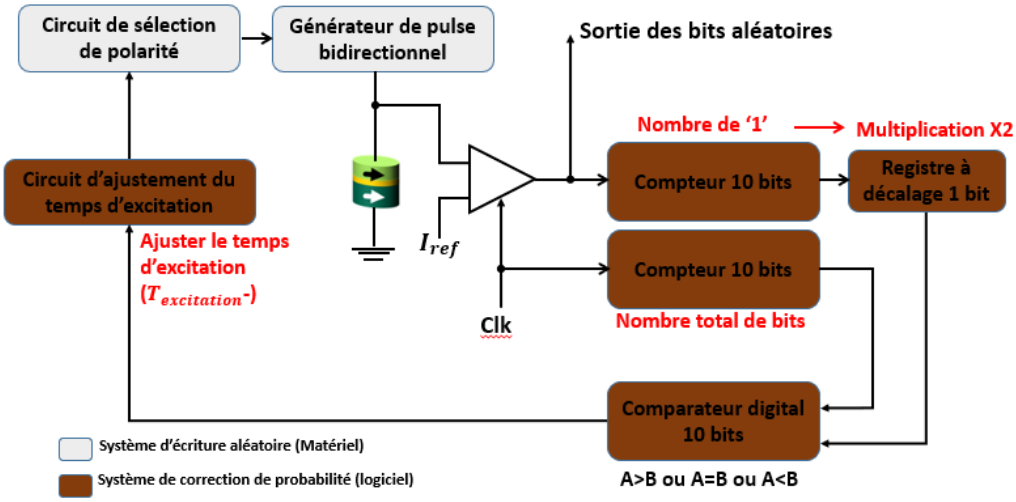


Figure III-15 : TRNG basé sur une jonction STT avec écriture d'excitation conditionnelle, suivi d'un contrôle de la probabilité de sortie en temps réel.

Cette technique bien qu'ayant été validé expérimentalement reste difficile quant à son implémentation dans les circuits intégrés notamment au niveau du circuit d'ajustement de la durée de l'impulsion d'écriture.

Les auteurs dans [98][99] ont également proposé une méthode de compensation de la probabilité basée sur le même principe, mais dont le paramètre d'ajustement est cette fois ci l'amplitude du courant d'écriture. Le circuit de correction est composé d'un DAC et d'un convertisseur tension/courant dans [98] tandis qu'il est basé sur le dimensionnent des transistors d'écriture en fonction des différents corners de la technologies CMOS (SS, TT, FF) dans [99]. Par contre le temps de verrouillage s'observait, même avec des corrections apportées sur des bits générés en dessous des $1Kbits$ et même sur des corrections apportées cycle par cycle dans [99].

4.1.3. Autres méthodes de génération de nombres aléatoires à base de MRAM

Parmi les études sur la génération de nombres aléatoires à base de MRAM, les auteurs dans [100] proposent une méthode d'implémentation par simulation du TRNG à base de jonction STT sur le même principe que dans [97] (c'est-à-dire une correction de la probabilité en

ajustant la durée du pulse). Le circuit permettant de moduler la durée du pulse est basé sur un générateur de délai inspiré de [101] dont le principe est de maintenir inchangé le signal d'écriture lors des fronts montant mais d'introduire des délais à travers des transistors lors des fronts descendant du signal d'écriture. La méthode proposée nécessite un soin particulier au dimensionnement du générateur de délai pour tenir compte des variations PVT (*process voltage temperature variations*). Aussi les nombres générés restent fortement corrélés quand la correction porte sur des séquences de taille inférieure à $512bits$. Les auteurs proposeront ensuite une amélioration de la méthode dans [102] en y ajoutant un post traitement arithmétique XOR sur plusieurs JTM écrites soit en parallèle ou en série. Cela permet de réduire considérablement le biais observé auparavant et de supprimer le temps de verrouillage. Dans leur étude, il ressort qu'un minimum de 24 JTM est suffisant pour passer tous les tests de NIST.

Dans [103], les auteurs étudient l'utilisation de plusieurs JTM en parallèle sans un système de correction de la probabilité. Une illustration de la méthode est présentée à la Figure III-16 (a). L'idée est de prendre avantage sur les variations des procédés de fabrication affectant chaque JTM individuellement. En appliquant une même tension à plusieurs JTM en parallèle, les courants traversant les jonctions seront légèrement différents en raison des différences de résistance dues aux variations de procédés de fabrication. La méthode consiste donc à déterminer la valeur de tension permettant d'écrire une jonction avec 50% de succès et de l'appliquer simultanément sur N JTM en parallèle. Ainsi en considérant chaque JTM comme étant un générateur indépendant, on divise par $\sqrt{N}$ le biais de probabilité des séquences aléatoires générées par l'ensemble de ces N JTM. Par contre, en plus des variations des procédées de fabrication, la précision de la probabilité de commutation dépend également de la tension réelle et de la durée de l'impulsion appliquée aux JTM ainsi que la température lors de l'écriture. Ces paramètres globaux (PVT) affecteront toutes les JTM du circuit de la même manière et dans la même mesure. En d'autres termes, chacune des JTM peut produire des nombres aléatoires avec un biais de probabilité dans la même direction (supérieure ou inférieure au 50% prévus). Afin de maintenir la probabilité précise, les impulsions appliquées aux JTM doivent être bien contrôlées et les variations PVT doivent être insignifiantes. Cette étude a été complété dans [104] par les auteurs en proposant une nouvelle méthode dite de *conception symétrique de pair JTM*. Cette dernière repose sur l'utilisation d'une paire de jonctions en série dont l'écriture se fait à travers le même courant. L'idée étant d'appliquer un courant d'écriture de 50% sur la structure et de couper le courant dès qu'une des JTM commute. Les deux JTM ayant théoriquement la même probabilité de commutation avec plus ou moins le même biais seront ainsi moins sensibles aux variations PVT. Une illustration de la méthode est présentée sur la Figure III-16 (b).

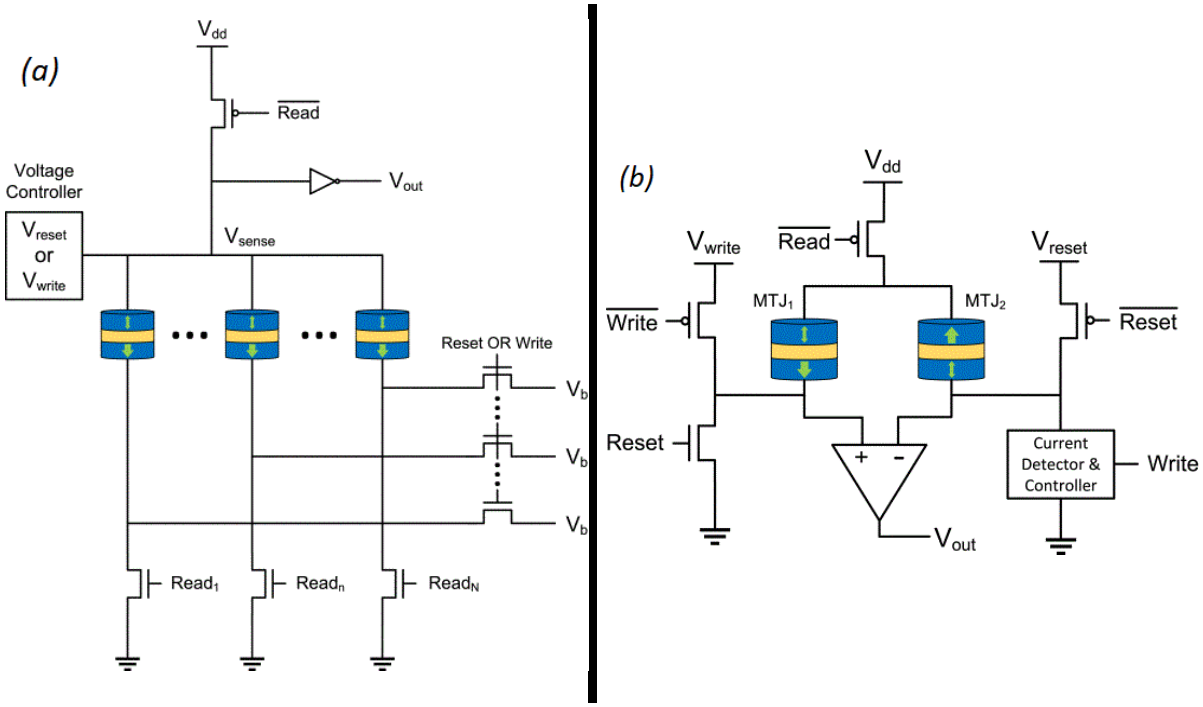


Figure III-16 : TRNG basé sur (a) plusieurs JTM en parallèle et (b) sur une paire de JTM en série[104].

Par contre le circuit de détection de la commutation des JTM doit être suffisamment rapide pour couper le courant d'écriture avant que la deuxième JTM ne commute.

Les auteurs dans [105], ont quant à eux conçu un prototype de TRNG basé sur une mémoire commerciale STT-MRAM conçu en technologie CMOS 28nm. Une modification du circuit d'écriture est effectuée pour permettre de passer du mode normal (fonctionnement mémoire) au mode TRNG. Dans le mode TRNG, l'écriture dans le sens P vers AP des JTM se fait normalement (avec un courant suffisamment élevé) c'est la phase de reset. Par contre l'écriture dans le sens AP vers P des JTM se fait avec un courant 2 à 3 fois plus faible que dans la configuration normale. Ceci a pour conséquence de rallonger le temps nécessaire à la commutation de la JTM. Un échantillonnage de ce temps de commutation est alors effectué à travers un oscillateur haute fréquence et un compteur. En répétant cette opération plusieurs fois sur la même JTM cela permet de mettre en évidence le caractère aléatoire du temps de commutation des JTM STT. Le LSB du compteur est utilisé comme nombre aléatoire. En d'autres termes, le compteur est activé lors de l'écriture de la JTM avec un faible courant, et il est arrêté juste quand la JTM commute et le LSB du compteur est utilisé comme nombre aléatoire. L'architecture général du circuit ainsi que la séquence d'activation et d'arrêt du compteur sont présenté à la Figure III-17.

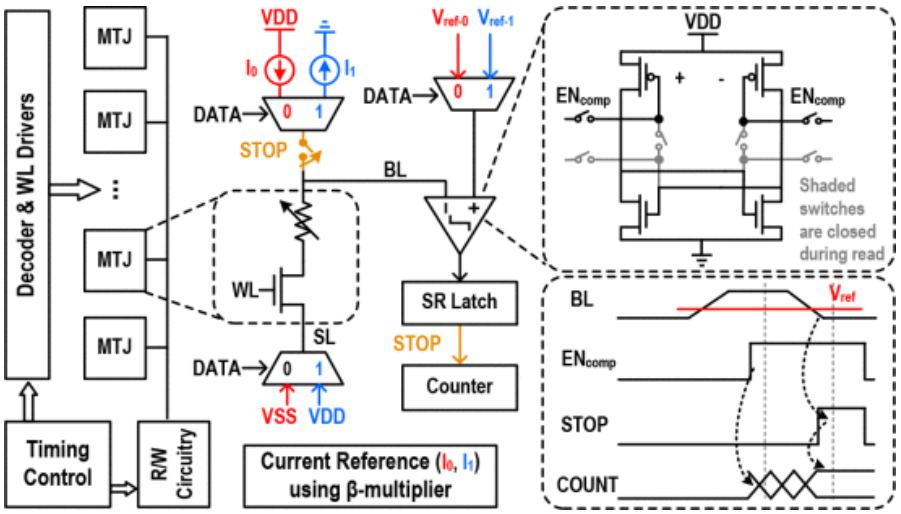


Figure III-17 : Circuit mémoire STT-MRAM utilisé comme TRNG à travers un échantillonnage du temps de commutation de la JTM lors de son écriture à courant faible [105].

Dans le même principe que précédemment, les auteurs dans [106] ont également étudié le temps de commutation d'une JTM STT sous un courant d'écriture faible. Leur méthode consiste à effectuer la différence entre deux temps de commutation successif pour générer un bit aléatoire. En effet le temps de commutation étant aléatoire sous écriture à faible courant, la probabilité d'avoir le même temps pour deux écritures successives est extrêmement faible. Ils ont mené une expérimentation pratique sur des JTM en procédant comme suit :

- Reset de la JTM à « 0 » avec un courant suffisamment élevé.
- Ecriture de la JTM à « 1 » avec un courant très faible (inferieure au courant critique) pendant un temps suffisamment long $(1us)$.
- Mesure du temps effectif de commutation de la JTM à « 1 »

Ce cycle d'écriture est répété deux fois et une différence entre les temps effectifs de commutation permet de générer un bit aléatoire.

Dans leurs études ils ont appliqué plusieurs valeurs de tension (courants) avec des pulses rectangulaires et triangulaires. Il ressort que pour n'importe quelles valeurs de courant inférieur au courant critique les bits générés par la méthode ont de bonnes propriétés statistiques. Par contre les meilleurs résultats ont été obtenu en utilisant des pulses d'écriture triangulaires.

Dans la littérature, il y'a également des études menées sur les autres technologies MRAM dans la conception des TRNG. Les auteurs dans [107] et [108] par exemple utilisent la méthode « spin dice » sur des jonctions SOT. Les bits générés dans ces études ont été validés par les

tests de NIST et l'intérêt majeur de l'utilisation de JTM SOT est leur endurance dans la mesure où le courant d'écriture des jonctions SOT ne traverse pas la JTM.

4.1.4. Synthèse

De nombreux travaux sur les TRNG basés sur les différentes technologies MRAM montrent l'intérêt réel de cette mémoire dans le domaine de la sécurité. La tendance commune dans ces études était de tirer parti du caractère aléatoire de l'écriture des JTM et de l'utilisé comme source d'entropie. Le Tableau III-1 fait un récapitulatif de ces travaux en les comparant en termes de la qualité des nombres aléatoires générés, leurs débits, leurs tailles ainsi que l'énergie consommé par bit généré.

Par contre la majorité de ces travaux a été effectuée en simulation. La validation de ces études à travers des prototypes est plus que nécessaire pour une utilisation future dans des circuits commerciaux.

Dans le cadre du projet Européen GREAT, nous avons décidé de concevoir un prototype de TRNG basé sur une JTM STT. Dans les paragraphes suivants nous détaillerons les travaux effectués dans ce sens.

Tableau III-1 : Comparaison des TRNG à base de jonctions magnétiques dans la littérature

	Source d'entropie	Type d'étude	Débit (Mbps)	Taille (µm²)	Energie (pJ/bit)	Post-traitement	Test de NIST
[93]	STT-JTM	Expérimentale	< 5	-	N/A	XOR	Validé
[97]	STT-JTM	Expérimentale	2 - 66	-	4,39	RTOPT	Validé
[105]	STT-JTM	Expérimentale	66	180	11 - 18	NON	Validé
[98]	STT-JTM	Simulation	66	-	2,63	RTOPT	Validé
[99]	STT-JTM	Simulation	66	5.88	1,25	RTOPT	-
[100] [102]	STT-JTM	Simulation	3 - 74	1,17 - 14,11	46,33	RTOPT & XOR	Validé
[103] [104]	STT-JTM	Simulation	66 - 178	3,84 - 7,64	0,64 - 0,81	NON	Validé
[107] [108]	SOT-JTM	Simulation	-	-	0,05 – 0,12	RTOPT	Validé

4.2. Environnement de conception hybride CMOS/STT-MRAM

Pour la conception d'un IP, l'utilisation d'un environnement spécifique appelé flot de conception est nécessaire. Dans un flot de conception "Full Custom" classique, les structures

sont conçues et simulées dans leur forme schématique. Ensuite, des analyses supplémentaires peuvent être effectuées pour vérifier le comportement du système suivant la température, les dispersions CMOS, etc. Ensuite les layouts des circuits sont conçus, les extractions de parasites (R, L, C) sont réalisées puis de nouvelles simulations incluant ces parasites sont exécutées. Une fois le bon fonctionnement du circuit assuré, les masques GDSII sont envoyés à la fonderie.

Dans notre cas, le flot de conception est un peu plus complexe puisque nous devons intégrer le process CMOS et magnétique. Les simulations avec les vues schématiques sont exécutées grâce au modèle compact de la JTM développé par le laboratoire SPINTEC [109] qui sera décrit dans le paragraphe suivant. La réalisation du layout nécessite la définition de règles DRC, de masques supplémentaires pour la conception de la partie magnétique le tout intégré dans un Design Kit qui sera présenté par la suite. La dernière section traitera des différentes étapes supplémentaires de fabrication.

4.2.1. Modèle compacte de la JTM

Pour les simulations, un modèle compact de la JTM réalisé par le laboratoire SPINTEC a été utilisé. Le modèle est entièrement paramétrable (choix de la taille de la JTM, du RA, de la TMR, etc.). Il est basé sur une modélisation des phénomènes physiques se produisant dans la JTM et des données recueillies sur la caractérisation des prototypes STT-MRAM pour être le plus réaliste possible. Il a été conçu de façon à être totalement intégrable à l'environnement CAO de cadence pour effectuer des simulations mixtes CMOS/MRAM. La Figure III-18 présente une séquence d'écriture d'une JTM-STT de $200\ nm$ de diamètre avec un RA de $7{,}5\ \Omega.um^2$.

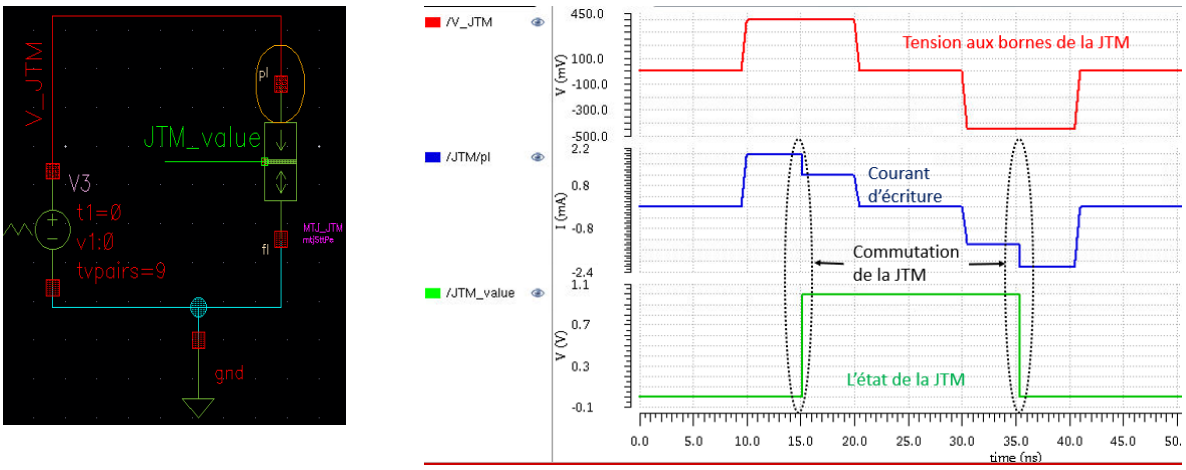


Figure III-18 : Simulation d'une séquence d'écriture d'une jonction STT

Les paramètres électriques de la jonction utilisée dans le projet sont résumés dans le Tableau III-2.

Tableau III-2 : Paramètres électriques de la jonction STT utilisé dans le projet GREAT

Paramètres	***Description***	***Valeur***
RA	Produit de la surface et de la résistance de la jonction	$7{,}5\ \Omega um^2$
$TMR0$	TMR quand la tension aux bornes de la JTM est à 0V	$1{,}5$
V_h	Tension de la JTM pour laquelle la TMR est à ½ $TMR0$	$0{,}6\ V$
D	Diamètre de la jonction	$200nm$
$Ic_{p\to Ap}$	Courant critique pour l'écriture P -> AP	$1{,}76\ mA$
$Ic_{Ap\to p}$	Courant critique pour l'écriture AP -> P	$1{,}22\ mA$
R_p	Résistance de le JTM à l'état parallèle	$239\ \Omega$
Δ	Le coefficient de stabilité thermique	38

Le modèle tient compte également des variations de procédés de fabrication pouvant affecter la JTM. Ces variations peuvent affecter principalement trois paramètres à savoir le RA, la TMR et le courant critique d'écriture en raison de $\pm 3\sigma$. Les pires cas étant une augmentation de $+3\sigma$ sur ces trois paramètres simultanément ou une diminution de -3σ .

Le Tableau III-3 présente les valeurs de simulations de la jonction sous les trois principaux corners.

Tableau III-3 : Valeurs des paramètres électriques de la JTM sous les 3 principaux corners

Paramètres	$variation\ de\ -3\sigma$	$valeur\ nominale$	$variation\ de + 3\sigma$
R_p	$203\ \Omega$	$239\ \Omega$	$275\ \Omega$
$Ic_{p\to Ap}$	$1{,}496\ mA$	$1{,}76\ mA$	$2{,}024\ mA$
$Ic_{Ap\to p}$	$1{,}037\ mA$	$1{,}22\ mA$	$1{,}403\ mA$

A partir des valeurs des Tableau III-2 et Tableau III-3 ainsi que de l'équation $(III-4)$, la probabilité de commutation de la JTM peut être calculer. La Figure III-19 présente la

probabilité de commutation de la JTM dans le sens P à AP en fonction du courant d'écriture pour les trois principaux corners et pour des durées de pulse de $10ns$ et $20ns$.

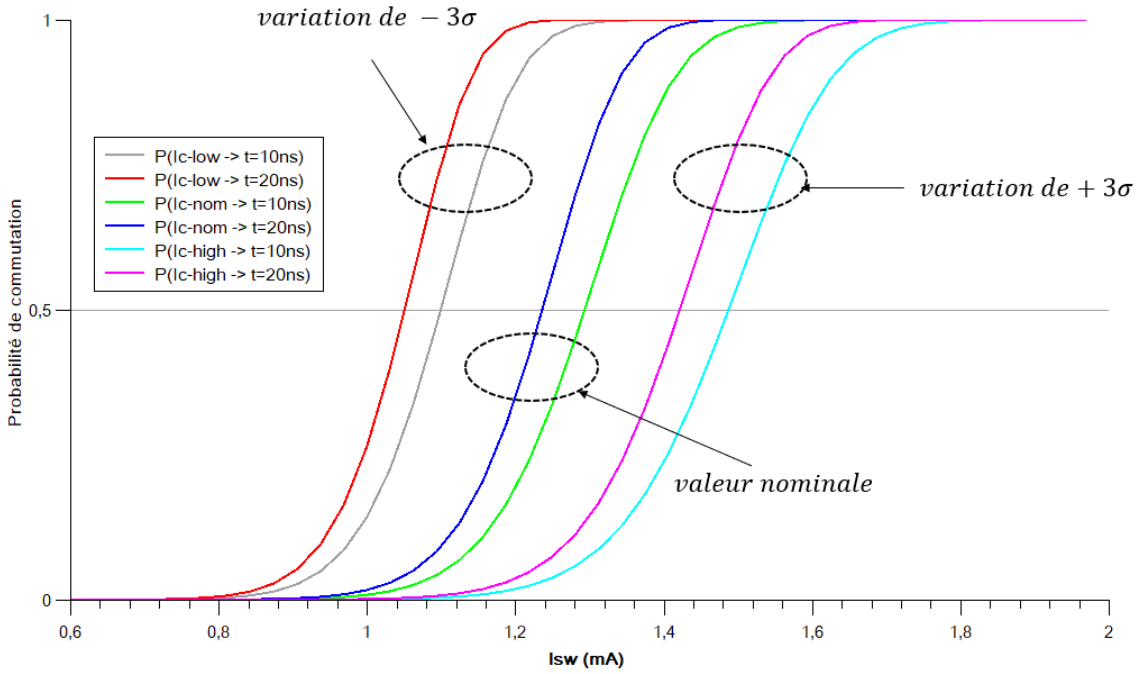


Figure III-19 : Probabilité de commutation de la JTM en fonction du courant d'écriture

4.2.2. « Design Kit » magnétique

Pour la conception et les vérifications du processus MRAM, un « Design Kit » (DKit) magnétique a été réalisé dans le cadre du projet GREAT par le laboratoire SPINTEC. Il repose sur le DKit standard de la technologie CMOS 180nm de la fonderie TOWER sur lequel des modifications ont été apportées permettant la compatibilité des deux procédés technologiques. Pour la conception du layout, les couches de dessins correspondant à la partie magnétique ont été ajoutées à l'environnement Virtuoso standard). Ces différentes couches correspondent au surcoût en termes de masque de fabrication supplémentaire à l'intégration de la technologie STT-MRAM avec le CMOS. Des règles de dessin ont été définies par le laboratoire SPINTEC de façon à minimiser les risques de défaillance lors de la gravure des JTM.

Toutes ces règles ont été incluses à l'environnement Cadence afin de pouvoir exécuter toutes les fonctions possibles d'un DKit standard telles que les vérifications DRC, LVS (CMOS et/ou MRAM) et l'extraction parasite des niveaux magnétiques. La Figure III-20 présente le layout de la JTM ainsi que les *layers* magnétiques du DKit.

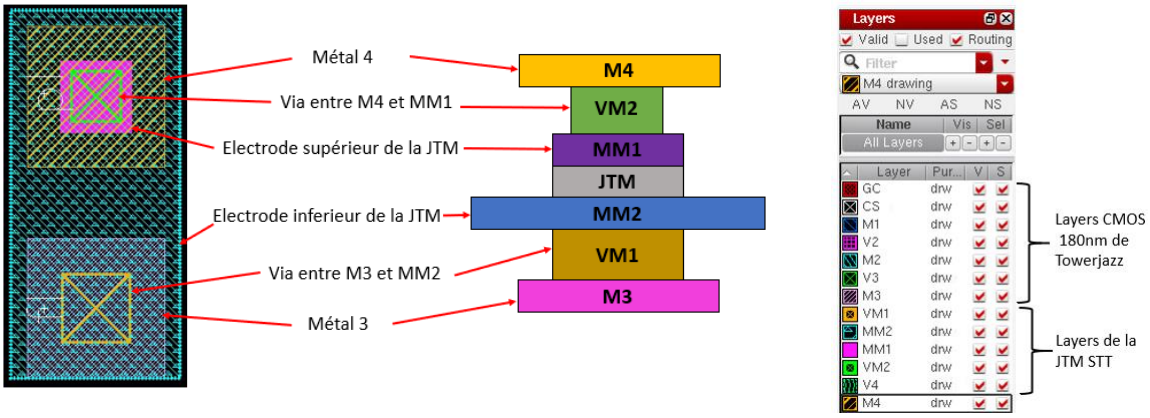


Figure III-20 : Vue layout de la JTM ainsi que les layers magnétiques du DKIT

La fabrication du CMOS est arrêtée après le niveau de métal 3 de la technologie Towerjazz (avant le dépôt de la couche de passivation), de sorte que les contacts sur le dessus de la plaque soient accessibles pour le dépôt de la partie magnétique. Une fois la partie CMOS réalisée, les wafers sont expédiés à SPINTEC pour le dépôt de la partie magnétique et enfin transférés au CMP pour séparer les puces du wafer et effectuer leur mise en boitier.

Après avoir vue les flots de conception utilisés, l'étape suivante consiste à définir les éléments basic de notre IP à savoir les circuits de lectures et d'écritures permettant de réaliser les fonctions souhaitées.

4.3. Les amplificateurs de lecture

Dans le contexte de la réalisation d'un prototype de générateur de nombre aléatoire à base de STT-MRAM, nous allons présenter dans les paragraphes suivants les différentes structures mises en œuvre à ce jour pour convertir l'information résistive d'une jonction tunnel magnétique en une information électrique (0V ou VDD) compatible avec les circuits électriques CMOS. Puis nous les comparerons sur la base de certains critères afin de déterminer laquelle de ces structures est la plus apte à être intégrée dans notre design.

4.3.1. Cellule de lecture "Black and Das"

Dans cette structure pionnière [110], les auteurs ont apporté de la non volatilité à la cellule SRAM classique volatile en rajoutant deux JTM en dessous des inverseurs rebouclés comme le montre la Figure III-21. La séquence de lecture de la donnée non volatile commence avec l'activation du signal de lecture (RD). Le transistor MN3 est alors passant, et les sorties Q et Qb sont court-circuitées, le verrou (latch) est alors dans un état métastable. La phase de lecture se termine par la désactivation du signal RD. Par conséquent, le verrou est poussé vers

l'un de ses états stables, qui est strictement déterminée par l'état des résistances des deux JTM. De cette façon, les données non volatiles stockées dans les jonctions deviennent disponibles sous forme électrique, aux sorties du verrou Q et Qb.

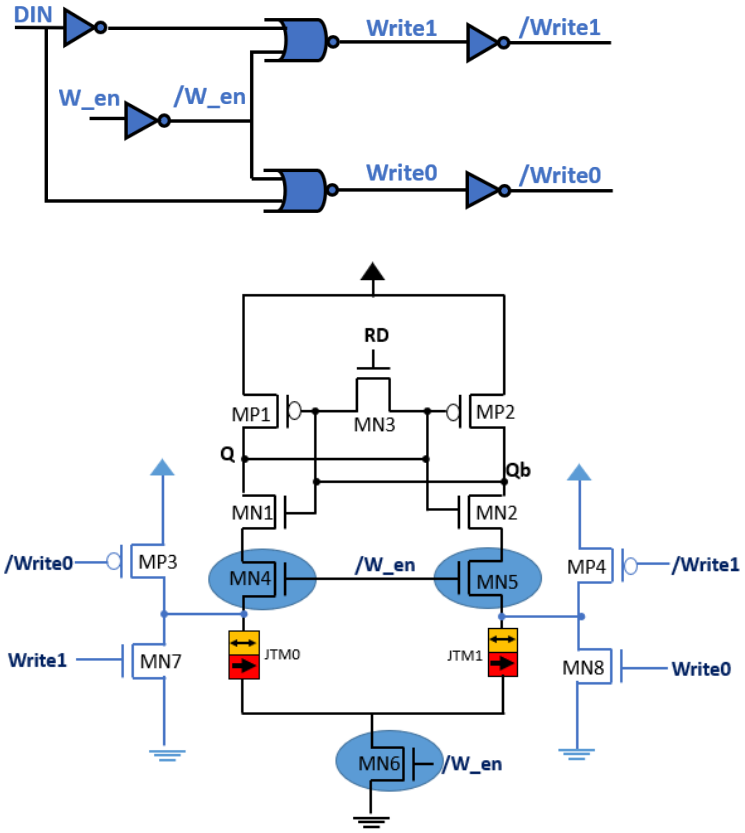


Figure III-21 : Structure de Black & Das avec deux JTM STT et le circuit d'écriture

Comme déjà mentionné dans le chapitre II, l'écriture des JTM dépend du modèle ou du type de JTM utilisé. Dans le cas des STT, le courant d'écriture doit pouvoir circuler dans les deux sens à travers la JTM. Ainsi les transistors MP3, MP4, MN7 et MN8 sont utilisés lors de la phase d'écriture. Ces transistors sont contrôlés par pair à travers un circuit logique présenté à la Figure III-21. Le signal DIN correspond à la donnée (0 ou 1) à écrire et le signal W_en permet d'activer la séquence d'écriture.

Quand $W\_en = 0$, tous les 4 transistors d'écriture (MP3, MP4, MN7 et MN8) sont bloqués, les transistors MN4, MN5 et MN6 (qui sont des transistors d'isolement des JTM pendant la phase écriture) sont passant. Les JTM sont donc accessible pour une lecture. Par contre quand $W\_en = 1$, les transistors d'isolement MN4, MN5 et MN6 sont bloqués et les JTM sont accessibles pour une écriture. Alors en fonction de la donnée à écrire à travers DIN, le courant circulera soit à travers les transistors MP3 et MN8 ou soit à travers MP4 et MN7.

4.3.2. Cellule de lecture de type « Hass »

Conçu à l'origine pour les applications spatiales en raison de son immunité aux radiations [111], cette cellule a sa partie non volatile (les JTM) placée entre les drains des transistors constituant les inverseurs comme le montre la Figure III-22. Si on remplace les JTM par des fils on retrouve les deux inverseurs rebouclés.

L'opération de lecture des JTM se fait à travers le signal RD qui commande les transistors MP3 et MN3. En partant d'un état stable du verrou par exemple $Q = 0$ et $Qb = 1$, quand RD est à VDD, le transistor MP3 est bloqué pendant que le transistor MN3 est passant. Il n'y a essentiellement aucun courant qui traverse les JTM dans ces conditions, donc les nœuds $L\ et\ Q$ sont quasiment au même potentiel (gnd) et les nœuds R et Q sont également à la même tension (VDD). Lorsque le signal RD est désactivé, le transistor MP3 devient passant tandis que MN3 est bloqué. Le circuit étant coupé de la masse, tous les nœuds internes commencent à s'élever vers la tension d'alimentation. Les nœuds $Q\ et\ Qb$ seront alors au même potentiel. Le transistor MP3 qui court-circuite les deux nœuds permet d'égaliser rapidement $Q\ et\ Qb$. Sans MP3, une période de temps beaucoup plus longue serait nécessaire pour égaliser ces tensions, ce qui est essentiel pour la lecture de la différence de résistance entre les deux JTM. Lorsque RD est à nouveau activé, MP3 est désactivé tandis que MN3 est activé rétablissement ainsi le chemin vers la masse. Cela amène les sources des transistors MN1 et MN2 à la masse. Les potentiels aux nœuds $L\ et\ R$ et donc la tension VGS des transistors MN1 et MN2 sont beaucoup plus élevée que la tension de seuil des NMOS, de sorte que ces transistors commencent à conduire le courant. À ce stade, le courant traversant chaque JTM est limité par sa résistance. En supposant que la JTM0 a une résistance plus élevée que la JTM1, moins de courant circule dans le nœud L et son potentiel tombe légèrement plus rapidement que celui du nœud R. Lorsque le potentiel au nœud L tombe, il a tendance à réduire le VGS de MN2, ce transistor va donc diminuer son courant, renforçant la différence de résistance ΔR des JTM avec une rétroaction positive pour maintenir le potentiel au nœud R élevé. Puisque MN1 reste actif, il continuera à faire circuler le courant à travers MTJ1 et commencera à tirer le nœud Q vers le bas, tandis que peu de courant circulera à travers MTJ2 à cause du faible VGS de MN2. Lorsque Q tombe à gnd, le transistor PMOS MP2 conduit plus de courant et amène le nœud Qb à VDD. À mesure que le nœud Qb augmente, le transistor MP1 sera désactivé, de sorte que les transistors PMOS à couplage croisé fournissent également une rétroaction positive pour renforcer ΔR. Tous les nœuds de la cellule atteignent rapidement VDD ou tombent à gnd, et le verrou atteint un état stable.

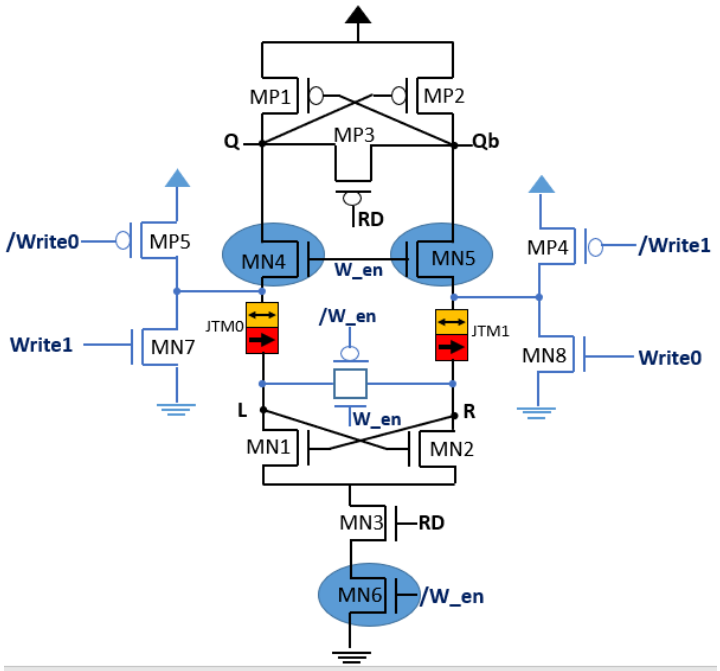


Figure III-22 : Structure de lecture Hass adaptée au circuit d'écriture STT

Pour l'écriture, le transistor MN6 avec les transistors de passage MN4 et MN5 isolent les JTM de la masse et des nœuds de sortie ($Q\ et\ Qb$), respectivement. La porte de transmission PG relie les extrémités inférieures des MTJ, ce qui permet au courant polarisé en spin de passer par les JTM dans les deux sens. Par contre, contrairement à la cellule Black & Das, le fait que les JTM soient connectés au plus près des nœuds de sorties ($Q\ et\ Qb$), rend l'écriture des JTM plus sensible à des perturbations sur ces nœuds.

Il convient également de mentionner que l'utilisation d'une porte de transmission plutôt que d'un simple transistor de passage permet d'avoir des résistances ON plus faibles. Des résistances ON plus faibles diminuent la résistance dans le trajet du courant d'écriture, ce qui permet de réduire la taille des transistors d'écriture.

4.3.3. Amplificateur de détection préchargé (PCSA)

Proposée en 2009 [112], cette cellule a une topologie similaire à celle d'un amplificateur de détection SRAM typique, comme le montre la Figure III-23. La lecture des données non volatiles des JTM comprend deux phases. Lorsque le signal RD est à 0, les transistors $MP3$ et $MP4$ préchargent les nœuds Q et Qb à Vdd. Pendant la phase de précharge, la cellule est déconnectée de la masse à travers le transistor $MN3$ qui est bloqué, ce qui réduit la consommation d'énergie au minimum. La phase de précharge est suivie par la phase de conversion. Le signal RD passe à 1, les des transistors de précharge PMOS MP3 et MP4 sont

bloqués, les tensions préchargées aux nœuds Q et Qb commencent alors à se décharger à travers les deux JTM et le transistor $MN3$. Comme les JTM sont dans des états opposés avec une résistance élevée et une résistance faible, la stabilité du verrou sera strictement déterminée par les résistances des JTM. La branche avec la plus grande résistance de JTM se déchargera moins rapidement et donc prendra la logique « 1 » à sa sortie (Q/Qb), tandis que l'autre (Qb/Q) sera à l'état « 0 ».

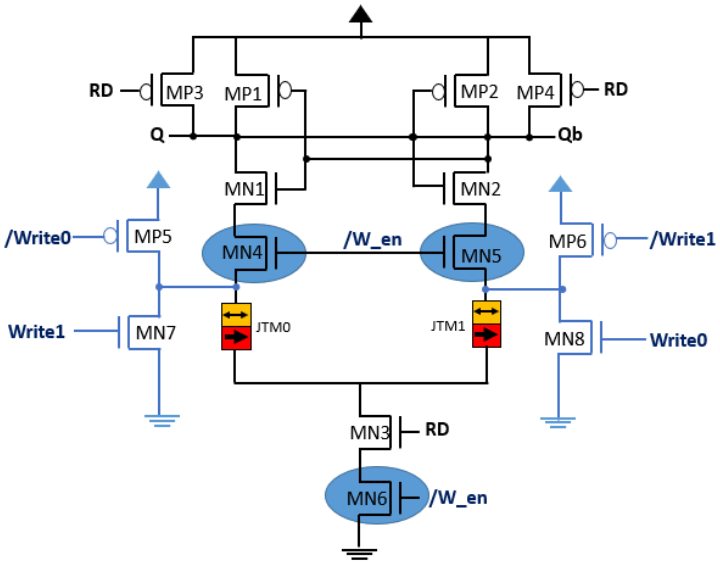


Figure III-23 : Structure de lecture PCSA

L'interface d'écriture est le même que dans le cas de la cellule Black & Das. Les transistors d'isolement MN4, MN5 et MN6 sont désactivés pendant la phase d'écriture.

4.3.4. SRAM rémanente (RSRAM)

Dans cette configuration, les deux jonctions complémentaires sont placés en dehors du cœur de la SRAM pour éviter la dégradation de ses performances, comme le montre la Figure III-24 [113]. Les inverseurs da la SRAM et les MTJ sont connectés par le biais des deux transistors d'accès.

Selon que les JTM soient connectés à l'alimentation ou à la masse les transistors d'accès sont de type P ($RSRAMp$) ou N ($RSRAMn$), respectivement. La phase de lecture commence par l'activation du signal RD, les transistors d'accès sont alors passant. Cela permet de connecter les deux JTM aux nœuds $Q\ et\ Qb$ du verrou, le mettant ainsi dans son état métastable. La lecture se termine en désactivant le signal RD. Les transistors d'accès sont alors désactivés et le verrou bascule dans un de ses états stables qui est strictement déterminé par les résistances des JTM.

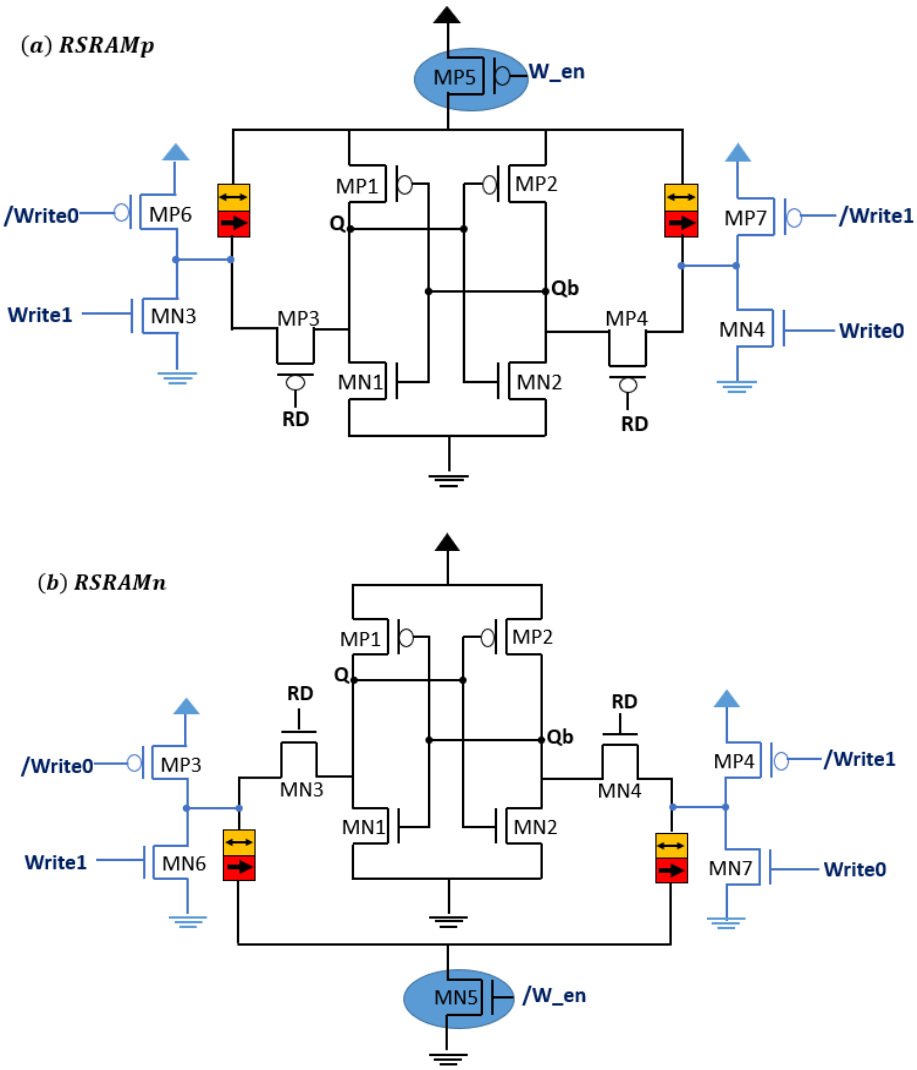


Figure III-24 : Structure de lecture SRAM Rémanent (RSRAM) (a) avec transistor d'accès PMOS et (b) avec transistor d'accès NMOS

Pendant la phase d'écriture, les transistors d'accès sont désactivés pour éviter de perturber les données volatiles (Q, Qb).

4.4. Comparaison des différentes cellules de lecture

Une comparaison de toutes ces cellules vues précédemment, en termes de surface, de temps de conversion (lecture), de consommation d'énergie et de robustesse a déjà été menée dans [114] en technologie *FreePDK 45nm* et *STT* $100nmx50nm$. Le Tableau III-4 en fait un résumé. En termes de surface, la cellule Hass est celle qui occupe le plus de surface silicium (environ 7 fois la taille des autres cellules). Par contre elle est celle qui a le temps de conversion le plus faible. La cellule PCSA est la plus efficace en termes de consommation énergétique.

Enfin les cellules $RSRAMp$ et $RSRAMn$ sont les plus robustes. La robustesse étant déterminée à travers le paramètre de marge statique au bruit (SNM) et celui de la dégradation du niveau logique des inverseurs lors de l'écriture des JTM. Plus la différence entre ces deux paramètres est supérieure à zéro (SNM-LLD > 0) plus la cellule est robuste.

Tableau III-4 : Caractéristiques des différentes cellules de lecture à base de SRAM dans [114]

Cellule de lecture (SA)	$\boldsymbol{Black}$ & $\boldsymbol{DAS}$	$\boldsymbol{Hass}$	$\boldsymbol{PCSA}$	$\boldsymbol{RSRAMp}$	$\boldsymbol{RSRAMn}$
Surface ($\boldsymbol{F^2}$)	< 25	> 175	< 25	< 25	< 25
Temps de conversion ($\boldsymbol{ps}$)	< 500	< 100	< 2000	< 2000	< 2000
Consommation statique ($\boldsymbol{nW}$)	< 100	< 700	< 50	< 400	< 100
Consommation dynamique ($\boldsymbol{fJ}$)	< 100	< 75	< 75	< 250	< 350
Marge statique au bruit - SNM ($\boldsymbol{mV}$)	304	431	323	342	342
Dégradation de niveau logique - LLD ($\boldsymbol{mV}$)	153	1100	109	0	0
Robustesse (SNM – LLD)	Modéré	Très faible	Modéré	Elevé	Elevé

Pour notre application de TRNG qui nécessite plusieurs écritures et lectures et dont l'objectif principal est la validation d'études théoriques à travers un démonstrateur, le critère principal que nous avons retenu est la robustesse de la cellule. Ainsi le choix s'est porté sur la cellule $RSRAMn$. La Figure III-25 présente la cellule de lecture réadaptée pour notre design. Une des JTM est remplacée par quatre JTM pour former la référence. La résistance équivalente formée par ces 4 JTM doit se situer entre la résistance minimale (Rp) et la résistance maximale (Rap) de la JTM0. Elles sont donc initialisées de sorte à avoir les JTM1 et JTM2 en configuration P ($Rmin$) et les 2 autres JTM (JTM3 et JTM4) en configuration AP ($Rmax$). Ainsi la résistance équivalente est $\frac{Rap+Rp}{2}$. Le circuit d'initialisation illustré en vert sur la Figure III-25 est utilisé à cet effet. La mise à VDD du signal INIT permet d'activer les 4 transistors MP3, MP4, MN6 et MN7, provoquant une circulation de courant de part et d'autre des JTM permettant d'écrire les JTM1 et JTM2 dans la configuration parallèle et les JTM3 et JTM4 dans la configuration antiparallèle.

Aussi les transistors MN6 et MN7 en plus des deux transistors d'accès (MN3 et MN4) permettent d'isoler les JTM pendant les phases d'écriture et d'initialisation.

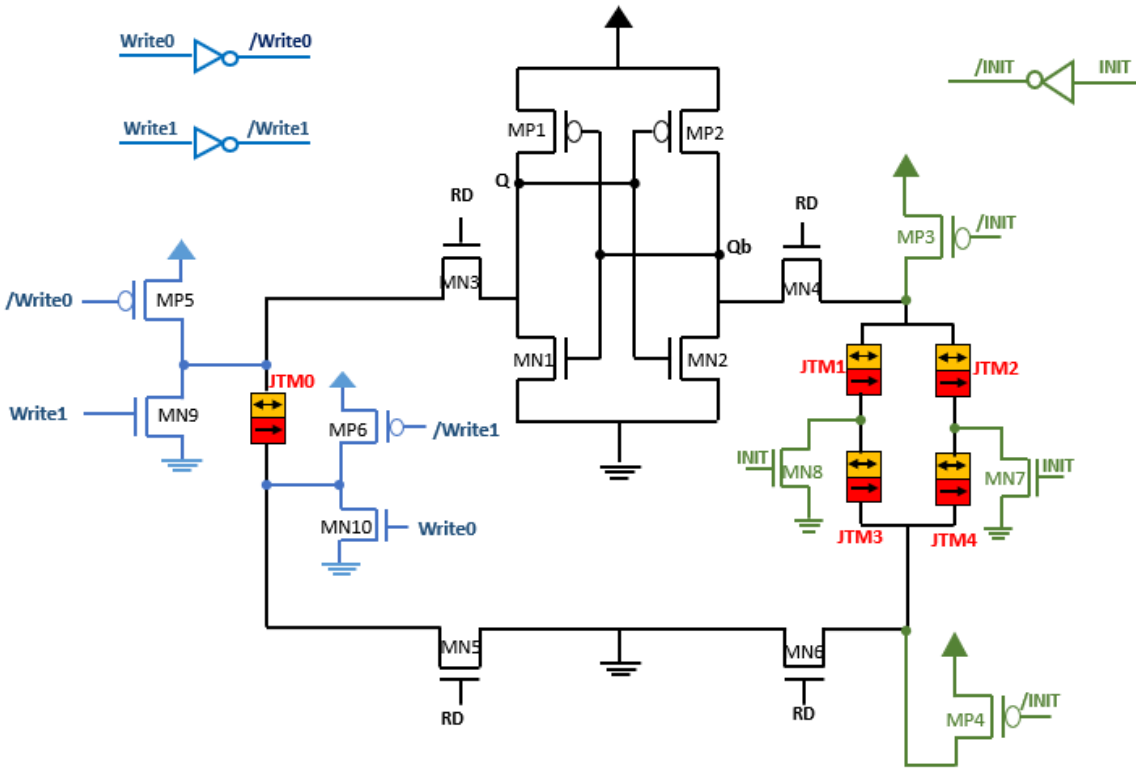


Figure III-25 : Cellule RSRAMn adaptée à un système d'écriture non complémentaire (référence composée de 4JTM avec le circuit d'initialisation)

4.5. Le circuit d'écriture

Comme nous l'avons vue dans les sections précédentes, pour générer des RBs à travers l'écriture des jonctions STT, on peut soit agir sur la durée de l'impulsion du courant d'écriture, soit sur son amplitude. Dans le cadre de la conception d'un circuit intégré, il est plus facile de contrôler l'amplitude du courant que la durée de l'impulsion. Une modification du circuit d'écriture est de ce fait nécessaire.

Nous avons choisi d'intégrer une source de courant programmable au circuit d'écriture de la JTM. La Figure III-26 présente la source de courant à base de miroir de courant simple. Les transistors MP1 et MN1 sont montés en diode et permettent de générer le courant de référence I_R. Le système de miroir est composé des transistors MN2 à MN6 qui ont tous le même V_{gs} que MN1. Cela implique que le courant au niveau de chacun des drains de ces transistors est fonction du courant I_R. La relation entre les courants drains des transistors MN2 à MN6 et le courant I_R est donnée par l'équation suivante :

$$I_R = I_{D1} = K.\frac{W_1}{L_1}.(V_{gs} - V_{th})^2$$

$$I_{Dx} = K.\frac{W_x}{L_x}.(V_{gs} - V_{th})^2$$

$$I_{Dx} = \frac{\frac{W_x}{L_x}}{\frac{W_1}{L_1}}.I_R \ (III-1)$$

Où K est une caractéristique du transistor, W et L la largeur et la longueur du canal, I_{Dx} ($2 \leq x \leq 6$), le courant drain des transistors MN2 à MN6.

Les drains des transistors du système miroir (MN2-MN6) sont reliés entre eux à travers des portes de transmissions. Ce qui permet de faire varier le courant I_{sw}. En considérant la même longueur de canal L pour tous les transistors et avec les largeurs W telles que définit sur la Figure III-26, le courant I_{sw} varie entre I_R et $2I_R$ avec un pas de $\frac{I_R}{16}$.

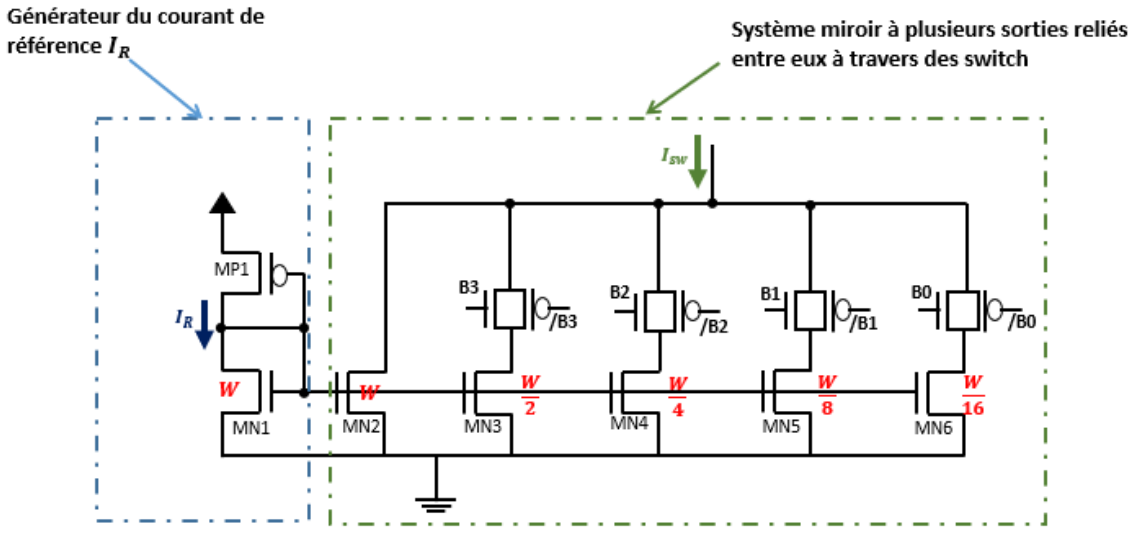


Figure III-26 : Miroir de courant programmable comme source de courant pour l'écriture de la JTM

La Figure III-27 présente le résultat de simulation de la source de courant programmable avec un courant nominal $I_R = 1\ mA$ et un I_{sw} qui varie entre $1\ mA$ et $2\ mA$ avec un pas de $62{,}5\ uA$. Comme on peut le voir, à chaque incrémentation binaire des signaux B0 à B3, il y'a une augmentation du courant I_{sw} traversant la JTM.

Nous avons choisi 4 bits dans le but d'avoir une plage de compensation beaucoup plus large ($2^4\ valeurs\ de\ courant\ possible$). Cela permet d'avoir une meilleure compensation des variations de procédé de fabrication, de tension et de température (PVT).

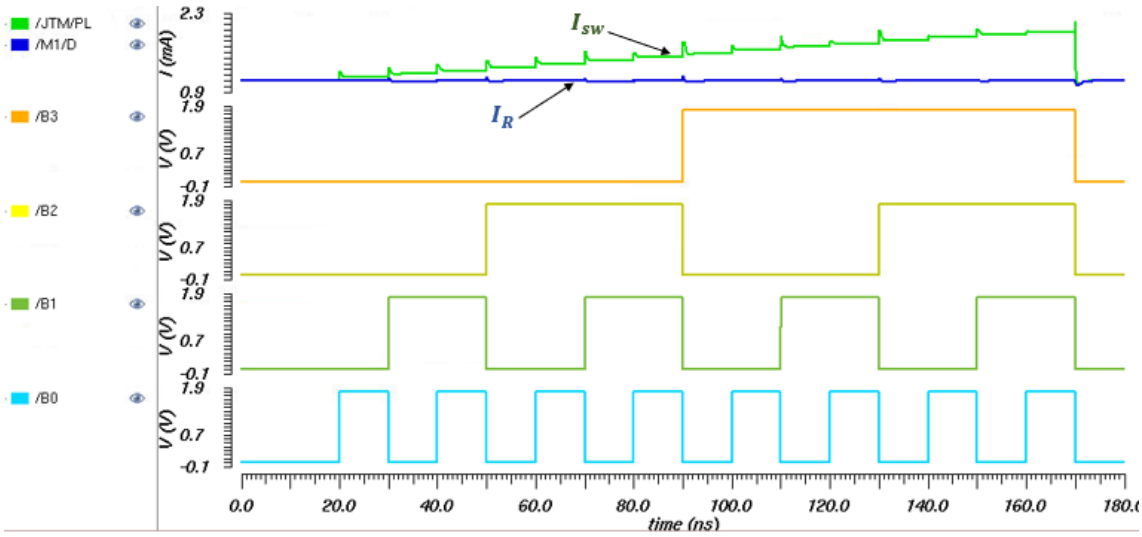


Figure III-27 : Simulation comportementale de la source de courant programmable

4.6. Système complet basé sur le spin dice

Le système de base final est présenté sur la Figure III-28. Il permet de réaliser la méthode « spin dice ». Le circuit d'écriture comporte la source de courant programmable dans le sens d'écriture à « 1 » de la JTM. Cela permet ainsi de faire varier l'amplitude du courant lors de l'écrire de la JTM à « 1 ». La phase de reset est assurée par les transistors en rouge. Ces derniers fournissent un courant suffisant (supérieur à $2I_{sw}$) pour s'assurer de l'effectivité de l'écriture de la JTM à « 0 ».

L'algorithme permettant de contrôler et de gérer le courant d'écriture à « 1 » de la JTM en fonction de la probabilité de sortie est illustré à la Figure III-29. Il débute par une étape d'initialisation de la JTM à « 0 » (la phase de reset). Ensuite une première écriture de la JTM à « 1 » avec un courant relativement faible (écriture aléatoire avec $I_{sw} = I_{sw}$) est effectué suivi d'une lecture de la JTM. A ce stade trois scenarios sont possible :

1) La probabilité de sortie à 1 de la JTM est égale à 50% ($P_{roba} = 50\%$), alors la JTM subit une phase de reset puis une nouvelle écriture aléatoire est effectué avec la même valeur de courant que précédemment.

2) La probabilité de sortie à 1 de la JTM est inférieur à 50% ($P_{roba} < 50\%$), alors la JTM subit une phase de reset suivi d'une nouvelle écriture aléatoire avec une incrémentation de la valeur de courant d'écriture à $I_{sw} = I_{sw++}$).

3) La probabilité de sortie à 1 de la JTM est supérieur à 50% ($P_{roba} > 50\%$), alors la jonction subit une étape de reset suivi d'une nouvelle écriture aléatoire avec cette fois ci une décrémentation de la valeur du courant d'écriture aléatoire ($I_{sw} = I_{sw--}$).

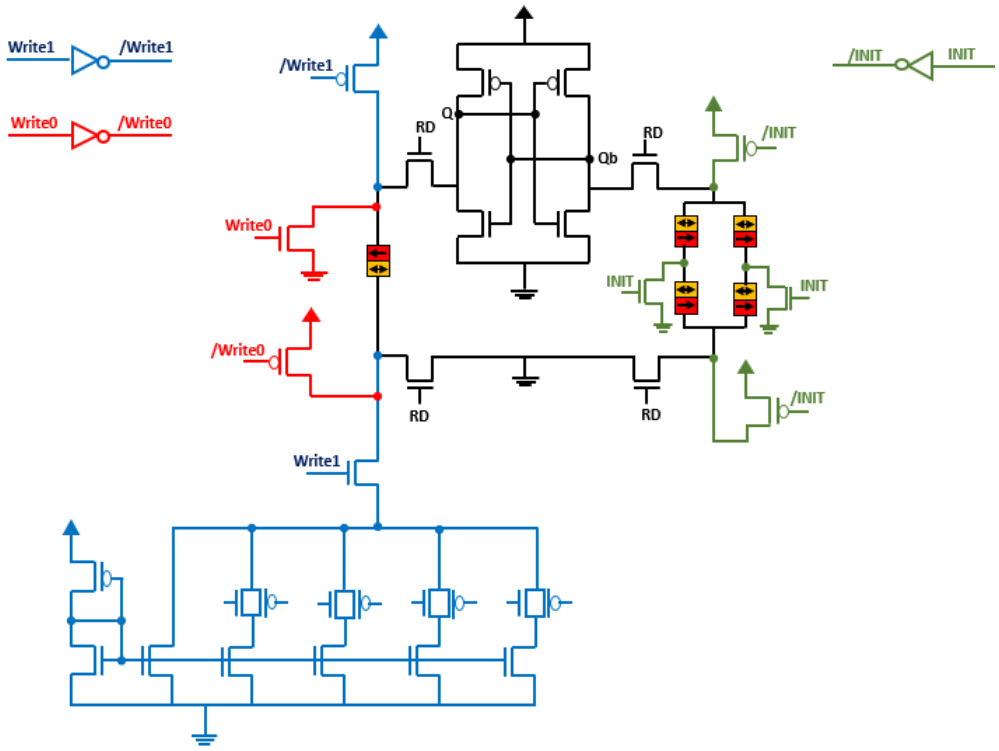


Figure III-28 : Circuit d'écriture et de lecture d'une JTM avec la possibilité de faire varier le courant d'écriture a « 1 » de la JTM.

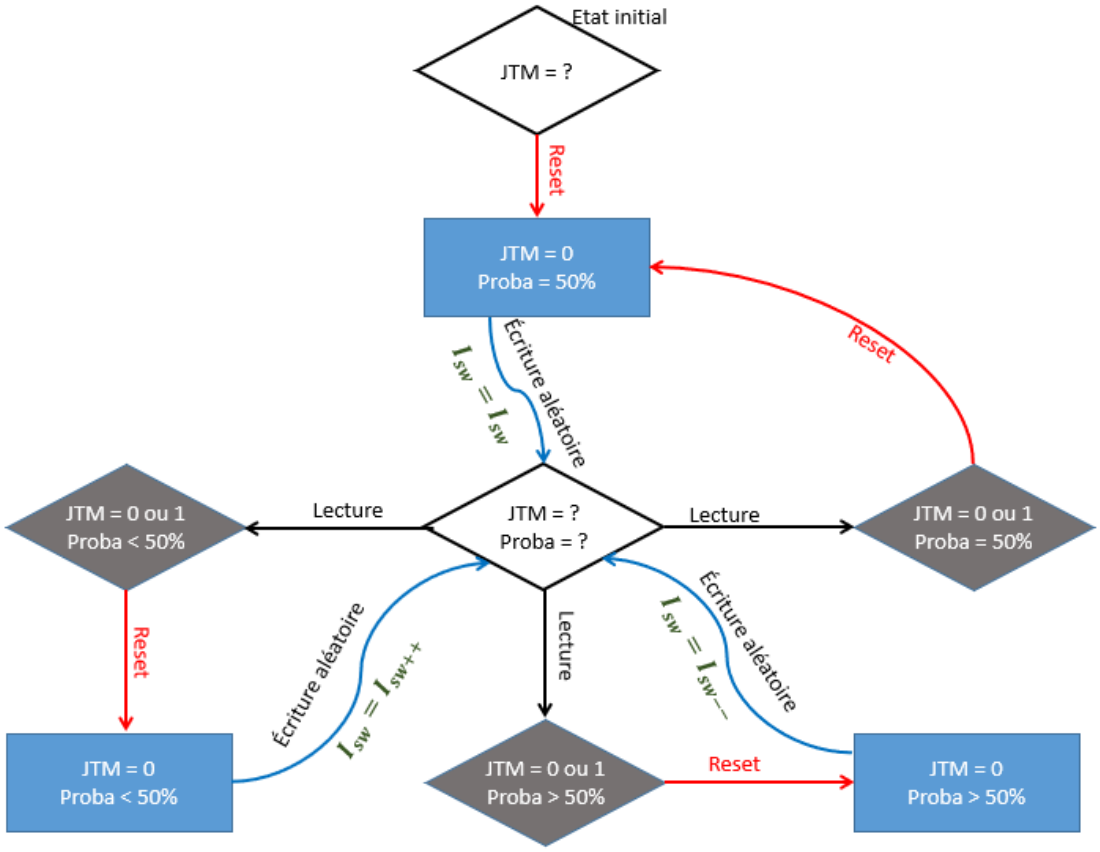


Figure III-29 : Algorithme de contrôle et de gestion de la probabilité de sortie à « 1 » de la JTM. Les trois états en rectangle correspondent aux probabilités de sortie calculées après chaque reset. L'état au milieu indique l'état inconnu de la JTM après chaque écriture aléatoire et les 3 états en gris indiquent l'état de la JTM ainsi que la probabilité de sortie actuelle après chaque lecture.

Une implémentation de l'algorithme de contrôle du circuit a été effectué en VHDL. Ce code est basé sur la méthode de contrôle de la probabilité RTOPT vue dans la section 4.1.2. Il utilise également 2 compteurs de 10bits et d'un comparateur pour le calcul de la probabilité de sortie.

4.7. Flot de conception du TRNG

Nous avons utilisé un flot « semi-custom » composé à la fois d'un flot analogique et d'un flot numérique. La conception du circuit utilise le flot analogique (full-custom), les transistors de la structure sont dessinés manuellement en respectant le dimensionnement des cellules (hauteur, pitch, etc.), ensuite les fichiers nécessaires à son intégration dans un environnement numérique sont générés (fichier LEF correspond à la géométrie de la structure, fichier LIB donnant les informations liées aux temps de propagation). La synthèse logique du code VHDL de l'algorithme de contrôle et son placement routage utilisent un flot de conception numérique classique.

Le layout du système ainsi que le packaging dans un boitier de type DIL40 sont présentés à la Figure III-30.

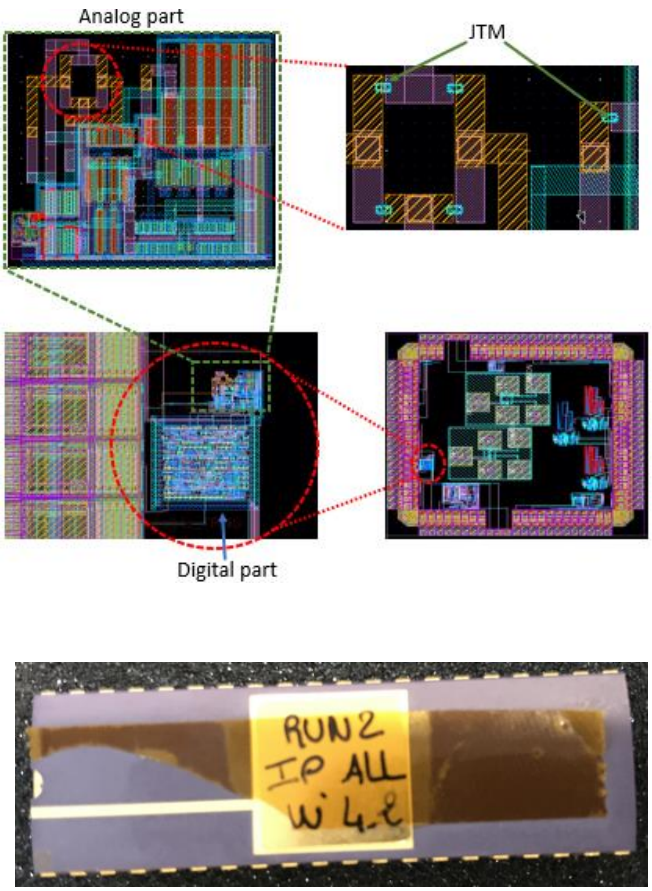


Figure III-30 : Layout du circuit et son packaging dans un boitier de type DIL40

4.8. Caractérisation et test du démonstrateur

Nous avons reçu quatre circuits provenant de deux wafers différents pour nos tests.

Une caractérisation des JTM a d'abord été faite par SPINTEC sur des matrices de JTM se trouvant sur les mêmes wafers que nos circuits. Les résultats de cette caractérisation ont montré une variation assez importante des paramètres attendus. En effet une résistance parallèle moyenne supérieur à $500\ \Omega$ a été observé au lieu des $239\ \Omega$ attendu. Aussi la TMR varie fortement (de 0% à 165% environ). Un problème de rétention a également été observé. Les JTM perdent leur valeur (passent à AP) après $1us$ pour certains wafers et $25us$ pour

d'autres. Les résultats de mesure ses les wafers sélectionnés pour la découpe de notre circuit sont présentés sur la Figure III-31.

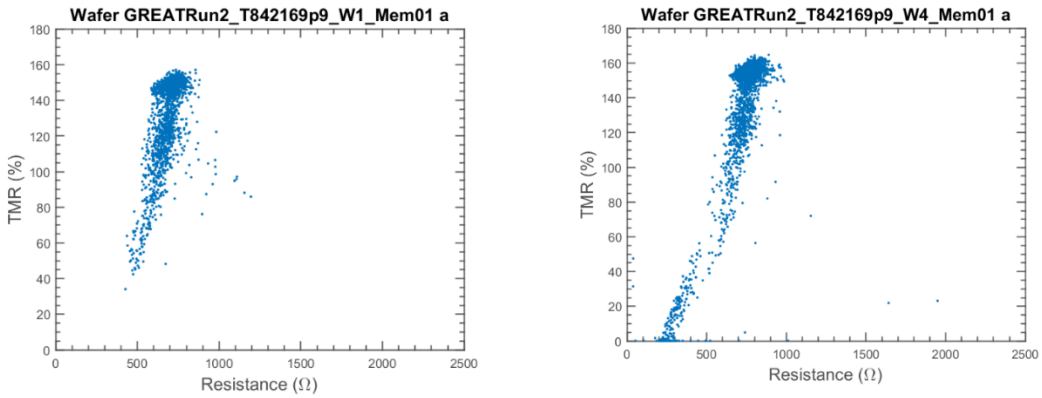


Figure III-31 : Mesure de la résistance parallèle et de la TMR sur deux wafers contenant notre circuit (résultats de SPINTEC)

Ces résultats nous laissent penser que nos circuits pourraient avoir un problème lié au niveau du processus de dépôt des JTM.

Nous avons néanmoins testé le fonctionnement de la partie CMOS à travers les signaux de contrôle de la source de courant programmable. En effet pendant le layout nous avons redirigés ces signaux vers des plots de sortie afin de pouvoir déterminer lors des tests la valeur de courant théorique pour laquelle nous aurions les 50% de probabilité de commutation.

La Figure III-32 présente les résultats de mesure d'un des circuits. En entrée du circuit on a le signal « INIT » permettant d'initialiser les JTM composant la référence, le signal d'horloge « CLK », le signal d'activation du TRNG « EN » et le signal de reset de tout le système « RST ». En sortie du circuit on a les signaux de contrôle ou de programmation de la source de courant « BIT0 » à « BIT3 » ainsi que le signal de sortie du TRNG « DOUT ».

Comme on peut le voir, quand le signal EN passe à « 1 » on a les signaux BIT0 à BIT3 qui qui varient de «0000 » à « 1111 » traduisant une augmentation du courant d'écriture en interne. Par contre la sortie DOUT reste collée à « 0 ».

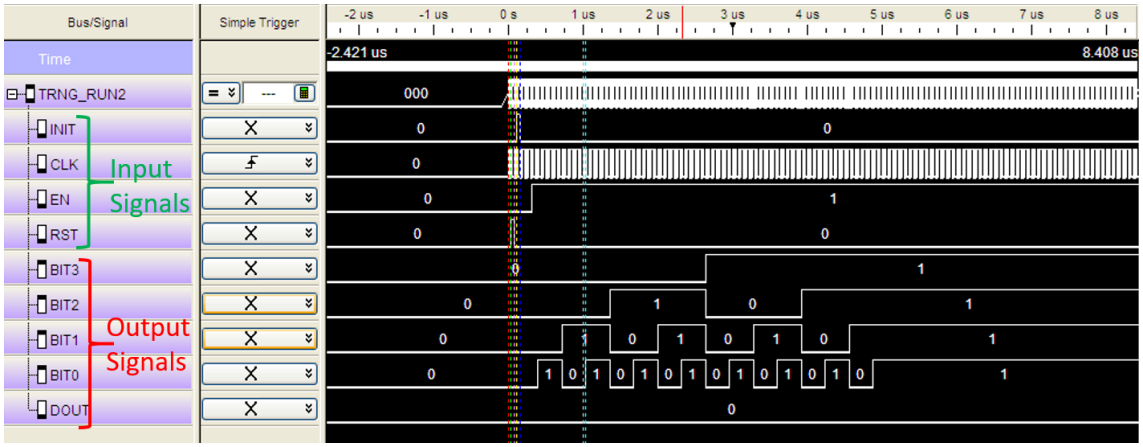


Figure III-32 : Test d'un circuit validant le fonctionnement de la partie digital du TRNG

Par contre les mesures d'un deuxième circuit nous ont présenté des phénomènes qui paraissent aléatoire à la sortie du TRNG. Ces comportements ont été observés pendant un certain nombre de fois puis la sortie est restée collé à « 1 ».

Les résultats obtenus de ce circuit sont présentés à la Figure III-33.

Figure III-33 : Résultats de mesures d'un circuit présentant un phénomène qui semble aléatoire

Nous avons obtenu très peu de résultats lors de nos tests sur les quatre circuits. Plusieurs hypothèses peuvent expliquer cela. D'abord les fortes dispersions de résistances et de TMR entre les JTM d'un même wafer, ensuite le problème de rétention également observé pourrait nécessiter un rafraichissement fréquent des JTM de référence ainsi qu'à fonctionner avec une fréquence assez élevée. Mais cela au risque de ne pourvoir véritablement bien écrire la JTM lors du reset par exemple. Enfin un problème de claquage des JTM est aussi possible dans la mesure où au début de nos tests il nous a fallu détruire les court-circuits résiduels en appliquant des tensions de l'ordre de 3,6 V à tous nos signaux d'entrées pour que les circuits

« fonctionnent ». Cependant le comportement observé sur quelques cycles semble indiquer un fonctionnement possible de cette structure, notamment avec une stabilité plus importante de la technologie.

5. Conclusion

Dans ce chapitre nous avons menés des expérimentations sur la génération de nombre aléatoire à base de jonctions STT-MRAM. Nous avons d'abord présenté une revue des méthodes de génération de nombres aléatoire à travers les JTM STT puis nous avons conçu un démonstrateur à base de technologie hybride CMOS/STT-JTM. Les différentes étapes de fabrication du procédé hybride ont été présentées. La réponse magnétique des circuits conçus n'a pas été satisfaisante. Cela est dû en partie à la non maitrise du processus de fabrication des JTM entrainant des dispersions importantes qui n'ont pas été prévues lors des simulations. Cependant le comportement observé sur quelques cycles d'un des circuits conçus semble indiquer un fonctionnement possible de cette structure, notamment avec une stabilité plus importante de la technologie.

Le chapitre suivant va aborder l'utilisation des TAS-MRAM dans la conception d'un TRNG. Les résultats d'expérimentations pratiques sur des JTM TAS à travers un circuit de test de 1 kbits en particulier seront présentés.

Chapitre IV Générateur de nombre véritablement aléatoire à base de TAS-MRAM

1. Introduction

Comme vue dans le chapitre précédent, notamment dans l'état de l'art sur les méthodes de TRNG basés sur les jonctions MRAM. Des études théoriques et des expérimentations ont été menées sur la plupart des technologies MRAM. Cependant à notre connaissance aucune étude ou expérimentation de ce type n'a été mené sur les jonctions de type TAS.

Dans le cadre du projet Multismart (FUI AAP18) nous avions eu accès à un testchip de la société Crocus sur lequel nous avons décidé de mener cette expérimentation. Nous détaillerons dans les lignes suivantes le principe de fonctionnement dans le cas des jonctions TAS ainsi que les outils et la méthode d'extraction d'aléa utilisés puis nous présenterons les résultats obtenus en termes de la qualité des bits aléatoires générés à travers les tests de NIST.

2. Principe de fonctionnement

La source d'aléa ici est basée sur le principe d'écriture de la JTM. En effet dans la technologie TAS-MRAM, l'écriture d'une jonction consiste à appliquer une tension (ou courant) de chauffage jusqu'à ce que la température interne de la JTM dépasse la température de blocage de la couche de stockage tout en ayant un champ magnétique externe suffisant. Le champ magnétique est appliqué dans la direction de commutation voulu de la couche de stockage.

Le principe d'extraction de l'aléa dans ce contexte consiste à mettre la JTM dans un état métastable soit en appliquant une tension (courant) de chauffage insuffisant soit en appliquant un champs magnétique faible. En raison de la variation du processus de fabrication, le seuil de l'amplitude et de la durée de la tension de chauffage ou de la valeur du champ magnétique n'est pas identique pour tous les bits mémoires d'une puce. De plus, du fait de la variation environnementale et des effets de vieillissement, ce seuil n'est pas toujours le même pour une même jonction. Ainsi, il faut utiliser une tension de chauffage ou un champ magnétique supérieure au seuil afin de pouvoir écrire correctement et de manière fiable une cellule TAS-MRAM. L'utilisation d'une valeur proche du seuil peut entraîner des échecs dans les opérations d'écriture. Si l'on peut trouver une valeur de tension de chauffage ou de champ magnétique qui provoque une probabilité d'échec d'écriture de 50%, cette valeur peut être

utilisée pour concevoir un TRNG. Cependant, avoir exactement 50% n'est pratiquement pas réalisable.

3. Outils et méthodes d'extraction d'aléa dans les TAS-MTJ

Nous avons utilisé un circuit de test composé de TAS-MRAM fabriqué par la société Crocus technology. Chaque puce est composée de $1Kbits$ (1024 bits) de JTM TAS qui sont disposés en un réseau matriciel de 32 × 32 de telle sorte que chaque jonction soit individuellement adressable et accessible. L'une des puces utilisées est illustrée sur la Figure IV-1. L'architecture des matrices est présentée à la Figure IV-2. IOF, IOM et IOR sont des pins de sortie utilisés pendant l'opération de lecture. Comme le montre la figure, le pin IOF est au-dessus de la JTM, IOM est connecté juste en dessous de la JTM, et IOR est entre une résistance poly de 500Ω et un transistor de sélection. La résistance totale d'une cellule quand elle est dans la configuration parallèle (état «0 ») peut être modifiée de R_{min} à R_{max} à travers une opération d'écriture d'un « 1 » ($W1$). De même, l'écriture d'un « 0 » ($W0$) modifie la résistance de la JTM de R_{max} à R_{min}. Les deux opérations d'écriture nécessitent trois tensions : V_{Heat}, V_{Field1} et V_{Field2}. Le premier est nécessaire pour chauffer localement la JTM sélectionnée, tandis que le deuxième et le troisième permettent de générer le champ magnétique nécessaire pour changer l'orientation magnétique de la couche de stockage dans l'état souhaité pendant le chauffage.

Pour écrire correctement la JTM à « 0 » ($CW0$), il faut appliquer les valeurs de tensions suivantes :

$V_{Heat} = 2V$

$V_{Field1} = 3{,}3V$

$V_{Field2} = 0V$

De même, pour écrire correctement la JTM à « 1 » ($CW1$), il faut appliquer les tensions suivantes :

$V_{Heat} = 2V$

$V_{Field1} = 0V$

$V_{Field2} = 3{,}3V$

La durée de ces trois signaux, T_{Heat}, T_{Field1} et T_{Field2}, doit être de $30ns$ environ chacune. Dans ce cas, on peut être sûr que les opérations d'écriture se font sans échec. Nous appelons ces opérations « Certain Write 0 » ($CW0$) et « Certain Write 1 » ($CW1$).

L'opération de lecture consiste à mesurer la résistance totale de la jonction. Pour ce faire on applique les tensions suivantes :

$V_{Field1} = V_{Field2} = 0V$

$V_{Heat} = 0{,}3V$ avec $T_{Heat} = 30ns$

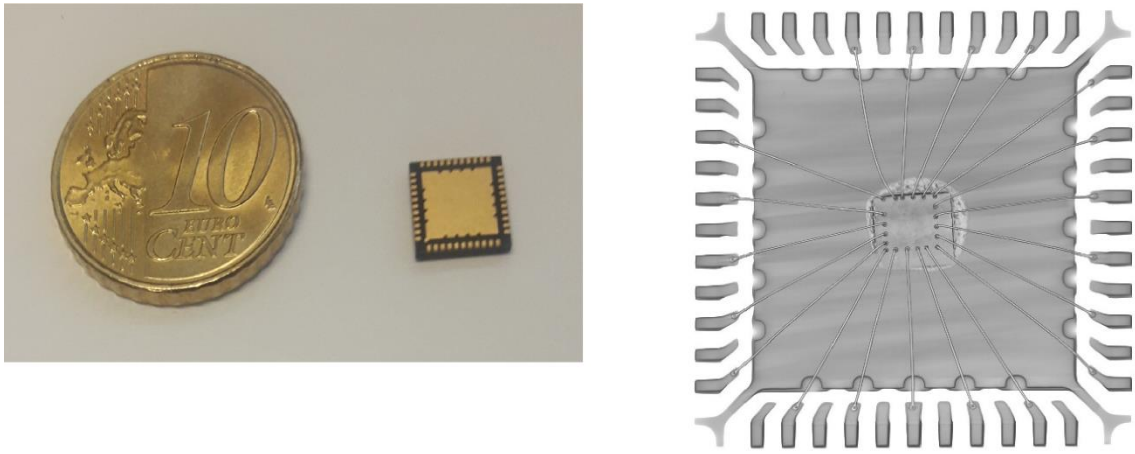


Figure IV-1 : (a) 1kbits TAS-MRAM dans un Package QFN44 (b) image microscopique de l'intérieur du QFN44

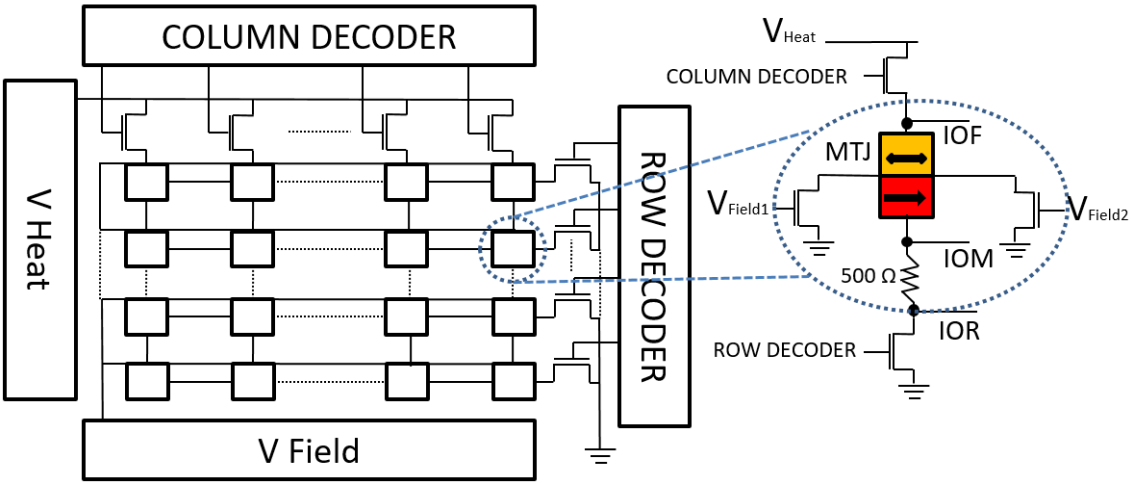


Figure IV-2 : Architecture de la puce composée de 1024 bits TAS-MRAM

Nous avons cherché à explorer l'effet de la tension de chauffage lors de l'écriture d'une jonction TAS. L'idée étant d'obtenir des tensions de chauffage pour lesquelles la JTM présente un phénomène aléatoire lors de son écriture. Ainsi si nous arrivons à déterminer la valeur de tension de chauffage qui provoque des échecs des opérations d'écriture avec une probabilité autour de 50%, nous aurions une source d'aléa potentielle pour un TRNG.

4. Instrumentation et expérimentation

Pour nos expérimentations, nous avons utilisé une carte FPGA Xilinx NEXYS2 et une carte d'acquisition du *National Instrument*, pour écrire et lire les matrices de JTM-TAS. Le FPGA permet de programmer 3 potentiomètres numériques se trouvant sur le PCB. Ces derniers sont utilisés pour fournir les tensions nécessaires lors de l'écriture et de la lecture d'une JTM. La carte d'acquisition est quant à elle utilisée pour lire la valeur de la tension aux bornes de la JTM pendant la phase de lecture. La Figure IV-3 présente le banc test utilisé pour nos expérimentations.

Une écriture correcte d'une JTM nécessite l'application d'une tension de champs de 3,3V et d'une tension de chauffage de 2 V. Par contre la lecture consiste à appliquer uniquement une tension de chauffage de 0,3 V et de mesurer la tension correspondante aux bornes de la JTM. Une valeur moyenne de 240 mV est obtenue dans le cas d'une lecture de la JTM à « 0 » et une valeur moyenne de 290 mV est obtenue pour une lecture de la JTM à « 1 ».

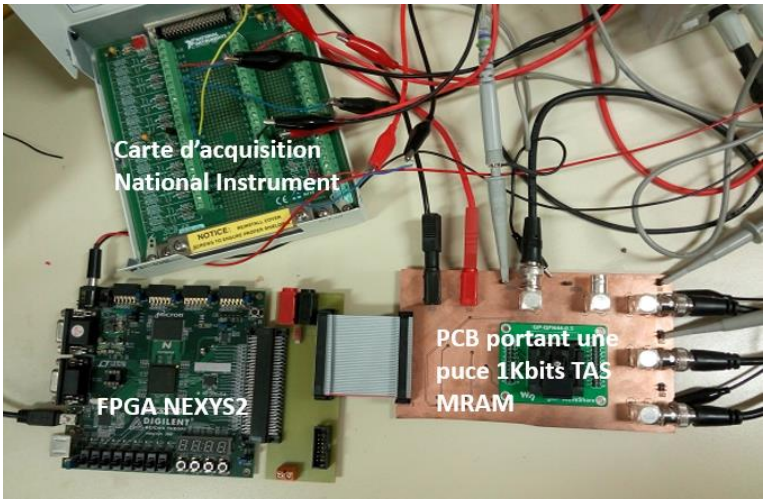


Figure IV-3 : Banc de test comportant une carte FPGA pour la programmation des composants utilisés pour l'écriture de la JTM TAS, une carte d'acquisition pour la lecture et le PCB contenant une puce 1kbit de TAS-MRAM

4.1. Caractérisation de l'écriture des jonctions TAS-MTJ

Comme nous l'avons expliqué dans la section 3, les signaux V_{Heat}, V_{Field1} et V_{Field2} sont utilisés dans les matrices TAS-MRAM pour les opérations d'écriture $W0$ et $W1$. Nous avons introduit également les propriétés requises de ces signaux pour avoir des opérations d'écriture correctes ($CW0$ et $CW1$). Ainsi, pour générer des échecs lors de ces opérations d'écriture on a 4 possibilités : agir sur la tension de chauffage ou la tension de champs magnétique ou agir sur la durée des pulses de ces tensions (chauffage et champs).

Notre dispositif nous permettait d'agir sur la tension de chauffage. Afin de générer des échecs lors des opérations $W1$, nous procédons à une séquence d'écriture « $CW0 - UW1$ » où $CW0$ correspond à une écriture correcte de la JTM (la phase de reset) et $UW1$ (*pour Uncertain Write 1*) correspond à une opération $W1$ pendant laquelle la tension de chauffage est dans la plage [0V, 2V]. Dans ce cas, la probabilité pour que la jonction soit écrite à « 1 » après avoir exécuté la séquence de $CW0 - UW1$, dépend de la tension sélectionnée et des dimensions réelles de la jonction. Afin d'avoir une source d'aléa idéale pour un TRNG, il faut sélectionner une valeur de tension de chauffage dans la plage mentionnée de telle sorte qu'elle provoque une probabilité d'échec de 50% lors de l'opération UW1. Pour trouver cette valeur de tension, nous procédons à une caractérisation comme suit :

10^3 séquences de $CW0 - UW1$ sont exécutées pour chaque valeur de V_{Heat} comprise entre [0V, 2V] et la probabilité de commutation de la JTM à « 1 » est calculé. Cette opération est répétée avec un pas de $10mV$ pour chaque valeur de V_{Heat}.

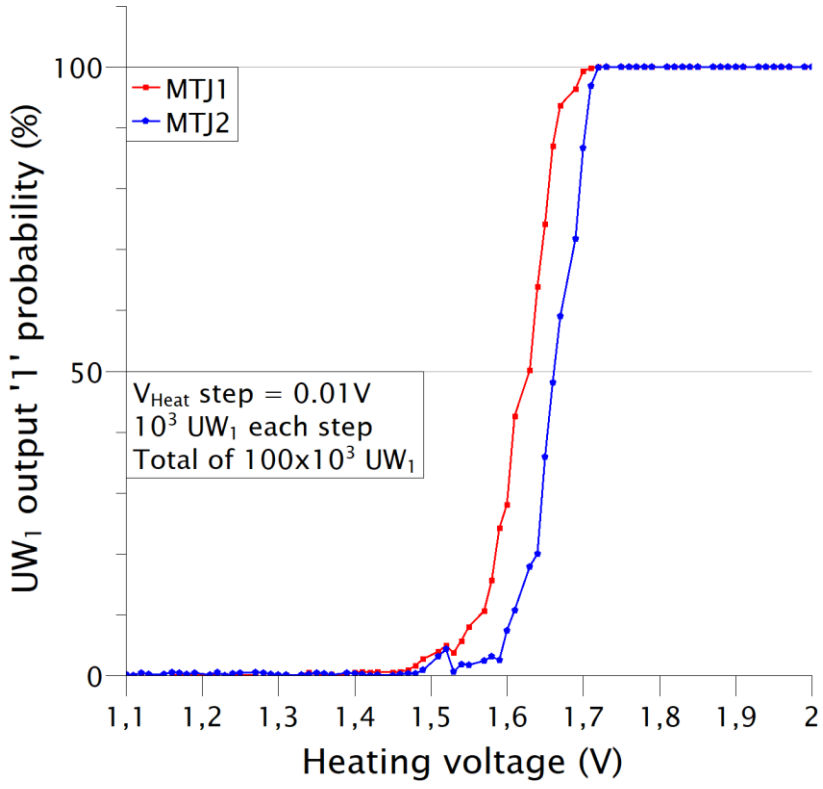


Figure IV-4 : Probabilité de commutation des JTM TAS en fonction de l'amplitude de la tension de chauffage

Les résultats de cette expérimentation sur deux jonctions (MTJ1 et MTJ2) d'une même puce sont présentés sur la Figure IV-4. Comme on peut l'observer, la probabilité de commutation de la jonction à « 1 » dépend de la valeur de la tension de chauffage. Elle augmente rapidement pour des tensions comprises entre 1,5 V et 1,7 V. La différence entre les courbes

bleu et rouge de cette figure est due aux variations de procédés de fabrication mentionnées plus haut.

4.2. Exploration de TRNG sur une jonction TAS

Les résultats observés sur la Figure IV-4 permettent d'envisager la possibilité de générer des bits aléatoires (*Rb, random bit*) avec une probabilité proche de 50% en ajustant la tension de chauffage. Pour cela, après avoir déterminé la tension de chauffage qui donne approximativement une probabilité d'échec d'écriture de 50%, nous l'utilisons pour générer une longue séquence de bits (10^7 bits) pour l'analyser ensuite en termes de nombre de « 0 » et de « 1 ». Les 10^7 bits sont ainsi regroupés en 1000 groupes de 10^4 bits chacun. Et pour chaque groupe on détermine le nombre de « 0 » et de « 1 ». Ces résultats sont présentés sur la Figure IV-5 à travers la distribution statistique de la probabilité de commutation des 1000 groupes.

Bien qu'une activité de commutation de 50% soit nécessaire pour avoir des nombres véritablement aléatoires elle n'est cependant pas suffisante. Des nombres aléatoires de haute qualité nécessitent une équiprobabilité de chaque événement de commutation. En d'autres termes, la distribution statistique de la probabilité de commutation des bits générés devrait être la même que la distribution binomiale. Cependant, la probabilité de commutation fluctue beaucoup en raison d'effets environnementaux tels que les fluctuations thermiques et de tension. Cette fluctuation de la probabilité de commutation est observée pour les deux jonctions (MTJ1) et (MTJ2). Dans les histogrammes, chaque point de l'axe des x présente une probabilité de commutation ; et l'axe des y montre combien de fois une probabilité de commutation se produit parmi les 1000 groupes. Pour MTJ1, la moyenne et l'écart type sont respectivement de 54,22% et 8,66% ; et pour MTJ2, ces valeurs sont respectivement de 53% et 6,79%. Les mêmes expériences sur d'autres jonctions de la même puce ont donné des résultats presque identiques à MTJ1 et MTJ2.

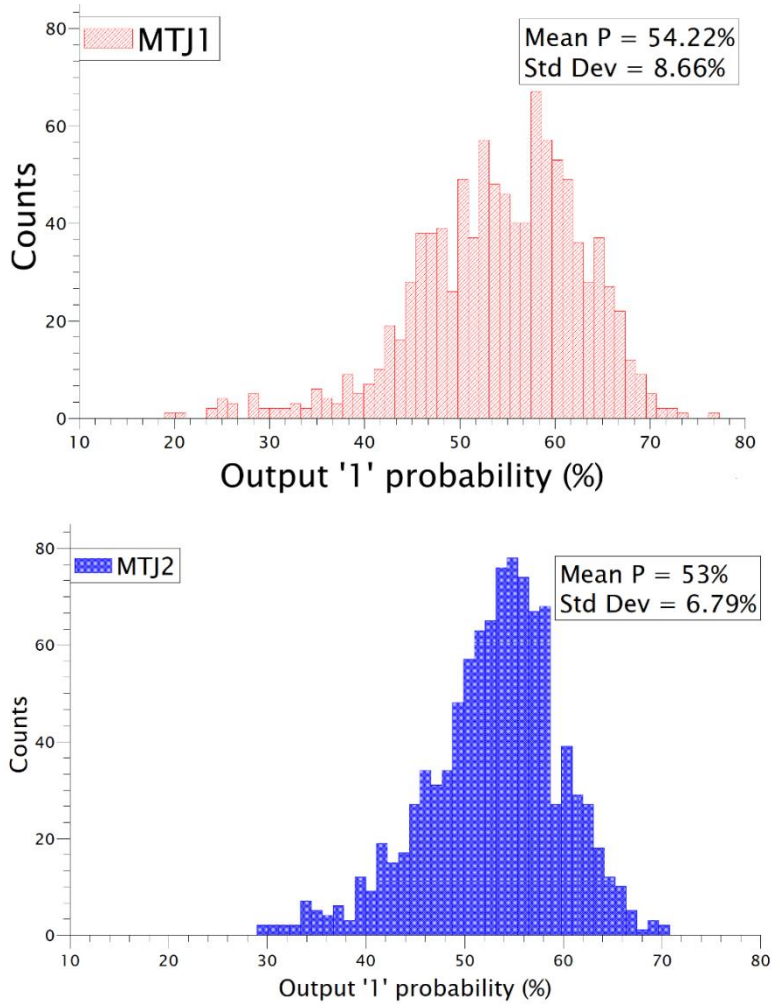


Figure IV-5 : Histogramme de la probabilité de commutation obtenu à partir de chaque 10^4 Rbs sur un total de 10^7 Rbs générés par les jonctions MTJ 1 (rouge) et MTJ 2 (bleu).

4.3. Utilisation des procédures de post-traitement pour améliorer la probabilité de sortie

Comme mentionné dans la section Chapitre III3.2 du Chapitre III, la qualité des nombres aléatoires bruts ($Raw\ Rbs$) générés par les TRNG peut être améliorée à travers des méthodes de post-traitement arithmétiques, comme celles de Von Neumann et du XOR. Les résultats présentés sur la Figure IV-5 montrent clairement la nécessité d'utiliser de telles procédures pour améliorer l'entropie.

En appliquant le correcteur de Von-Neumann aux nombres aléatoires bruts générés à la fois par MTJ1 et MTJ2, nous obtenons moins de $9X10^5\ Rbs$ à partir d'un total de $10^7\ Rbs$ bruts générés par chaque MTJ. Les Rbs traités représentent moins de 9% des Rbs bruts initiaux. Ce qui confirme la faiblesse en termes de débit de la technique de correction de Von

Neumann. Par contre, comme on peut l'observé sur la Figure IV-6, le biais est considérablement réduit. En effet, la probabilité de commutation obtenue pour chaque 8×10^3 Rbs à partir d'un total de 8×10^5 Rbs bruts pour les deux jonctions après correction de Von-Neumann est comprise entre 48,5% et 51,5%. La valeur moyenne de la probabilité de commutation et les écarts types sont respectivement de 49,97% et 0,48% pour MTJ1 et 50,008% et 0,49% pour MTJ2.

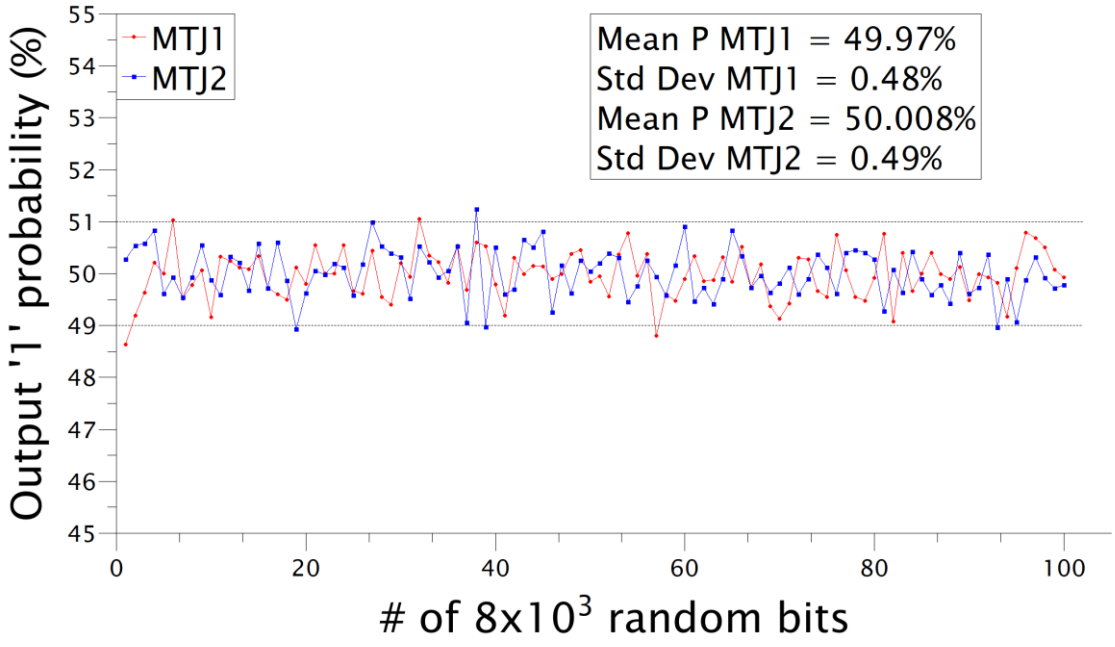


Figure IV-6 : La probabilité de commutation après un post-traitement de Von Neumann

En appliquant un post-traitement XOR sur les Rbs bruts générés par MTJ1 et MTJ2, nous obtenons une distribution des probabilités de commutation beaucoup plus centrée autour de 50% avec une moyenne de 49,5% et un écart type de 1,37%. Ces résultats confirment une bonne amélioration apportée par un post-traitement XOR. L'histogramme de la probabilité de commutation est représenté sur la Figure IV-7.

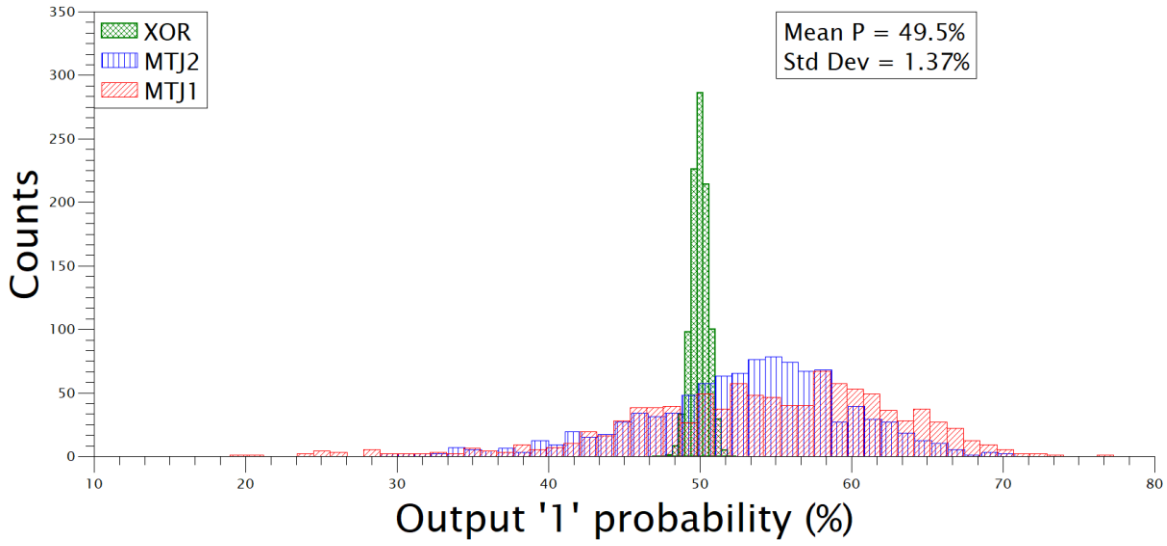


Figure IV-7 : Histogramme de la probabilité de commutation après un post-traitement XOR sur deux JTMs (MTJ1 et MTJ2)

4.4. Amélioration du biais à travers une boucle de contre réaction

Les auteurs dans [97] ont développé un TRNG basé sur la STT-MRAM (qui a été présenté en détail dans la section 4.1.2 du Chapitre III) où ils proposent une méthode de correction du biais à travers une boucle de contre-réaction appelée RTOPT (*Real time output probabilité tracking*). Cette boucle de contre-réaction permet de réajuster en temps réel la valeur de la tension de chauffage lors de l'écriture de la jonction. La Figure IV-8 présente l'architecture du RTOTP qui est composé d'un comparateur, de deux compteurs, et d'un circuit d'ajustement de la tension de chauffage. Le comparateur et les deux compteurs permettent de calculer la probabilité de commutation en temps réel du générateur et de la corriger en ajustant la tension de chauffage. Pour cela il faut au préalable définir le nombre de bits à générer sur lequel la probabilité sera calculée.

Nous avons implémenté cette méthode et nous avons générer 10^7 bits avec une correction de la probabilité pour chaque 10^3 bits générés. En d'autres termes, pour chaque 10^3 bits générés on calcule le nombre de « 0 » et de « 1 » puis on décide de diminuer ou d'augmenter la tension de chauffage en fonction du nombre de « 0 » et « 1 » calculé. Cette méthode compense en quelque sorte le faible débit de sortie du correcteur Von-Neumann et du post-traitement XOR. Elle présente également une meilleure résilience aux variations du PVT (*Process Voltage Température*).

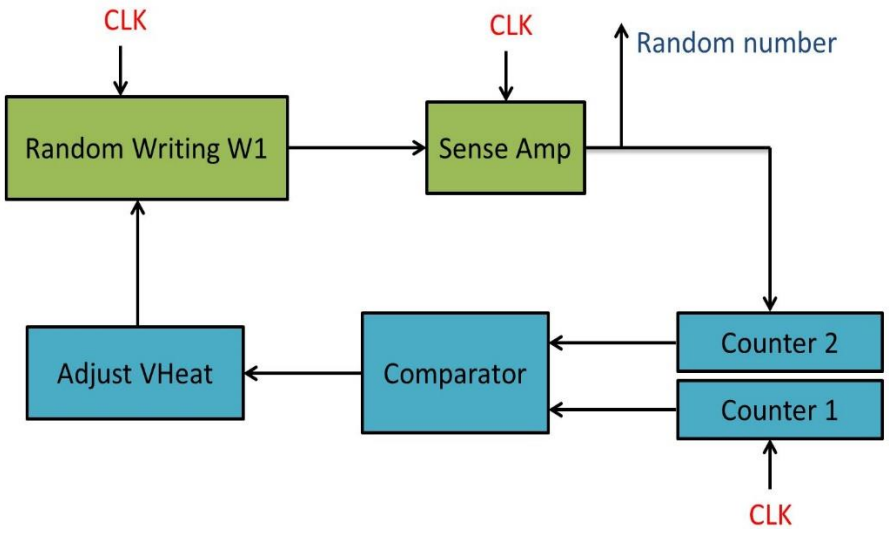


Figure IV-8 : Architecture de la boucle de contre réaction

La probabilité de commutation obtenue pour les 10^3 Rbs à partir d'un total de 10^7 Rbs générés par la méthode RTOPT est représentée sur la Figure IV-9. Cette figure montre l'amélioration de la distribution de probabilité de commutation qui fluctue autour de 47,45% et 52,47% avec une valeur moyenne de 50,12% et un écart type de 0,78%.

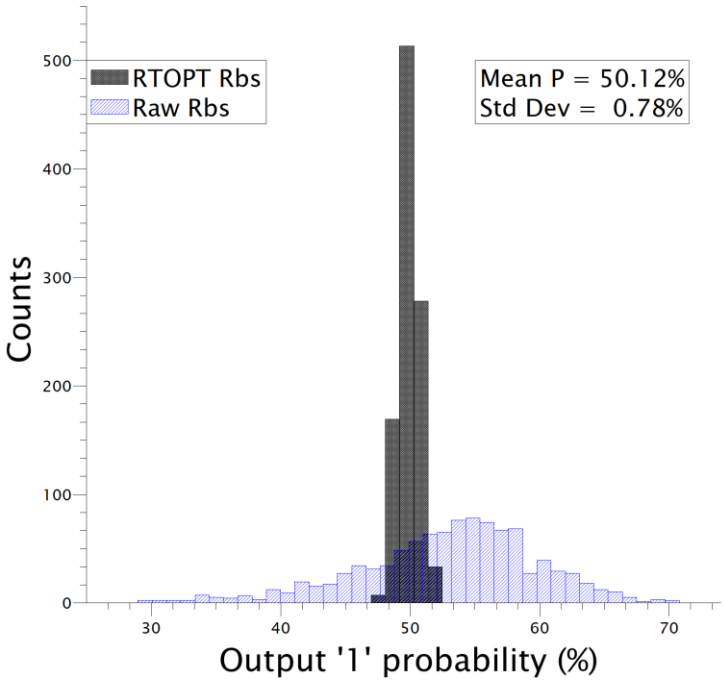


Figure IV-9 : Histogramme de la probabilité de commutation après un post-traitement basé sur la RTOPT

5. Evaluation statistique des nombres aléatoires générées

Pour évaluer la qualité d'un générateur de nombre aléatoire, le caractère aléatoire de ses nombres générés est étudié à l'aide de tests statistiques comme mentionné dans la section 3.3 du Chapitre III qui comparent les sorties des RNG avec celles qui, théoriquement, ont une séquence de nombre véritablement aléatoire. Nous avons décidé d'évaluer notre générateur à travers les tests de NIST dans la mesure où ils sont spécifiques aux applications de type cryptographiques.

Pour cela, nous avons généré un certain nombre de bits que nous avons organisé comme suit :

- 100 séquences de 10^5 bits générées par un seul MTJ (Raw Rbs) ;
- 100 séquences de 8×10^3 bits obtenues après correction de Von Neumann ;
- 100 séquences de 10^5 bits générées par la méthode RTOPT sur une seule jonction ;
- 100 séquences de 10^5 bits obtenues en effectuant un XOR sur 8 jonctions (XOR^3) ;
- 100 séquences de 10^5 bits obtenues en combinant la méthode RTOPT sur 4 jonctions puis en appliquant un XOR sur les 4 jonctions (XOR^2 sur RTOPT Rbs).

En raison du faible nombre de nos séquences de bits, nous n'avons pas pu tester 5 tests de la suite des 15 tests de NIST. Seuls 10 tests étaient applicables à nos séquences.

Les résultats sont présentés dans le Tableau IV-1. On constate que les séquences binaires brutes (Raw Rbs) du générateur (sans correction) ne passent aucun des 10 tests. Cependant, avec une correction XOR appliquée sur la sortie de quatre JTM, il y'a une amélioration de l'entropie se traduisant par le passage d'un certain nombre de tests de la série des tests de NIST. On remarque également que l'on passe tous les tests à partir d'une correction XOR sur 8 JTM.

Aussi l'application de la correction Von Neumann permet de passer tous les 10 tests mais avec une réduction importante du débit.

Quant à la méthode de correction RTOPT, elle améliore la probabilité mais ne permet pas pour autant de valider une grande partie des tests de NIST. Par contre en la combinant avec une correction XOR, on améliore considérablement le résultat des tests. Une combinaison RTOPT avec XOR sur au moins 4 JTM permet de passer tous les tests de NIST.

Tableau IV-1 : Résultat des tests de NIST sur les nombres aléatoires générés sans correction (Raw) ainsi qu'après les corrections XOR^3, Von Neumann, RTOPT et une combinaison de RTOPT et XOR^2. Les tests sont validés si P-value ≥ 0,0001 et Prop ≥ 96/100

Méthode	Raw		XOR3		Von Neumann		RTOPT		RTOPT & XOR2	
Test	P-value	Prop	P-value	Prop	P-value	Prop	P-value	Prop	P-value	Prop
Frequency	0,000000	5/100	0,122325	99/100	0,637119	99/100	0,202265	97/100	0,549284	100/100
Block Freq	0,000000	0/100	0,616305	98/100	0,514124	98/100	0,000012	88/100	0,699313	100/100
Run	0,000000	0/100	0,779188	98/100	0,102526	100/100	0,125552	96/100	0,999438	100/100
Longuest Run	0,000000	1/100	0,554420	98/100	0,275709	99/100	0,000052	97/100	0,779188	100/100
Cumul Sum 1	0,000000	0/100	0,224821	100/100	0,236810	99/100	0,000023	90/100	0,474986	99/100
Cumul Sum 2	0,000000	0/100	0,739918	100/100	0,437274	99/100	0,000045	87/100	0,249359	98/100
Binary Rank	0,000000	99/100	0,015589	1000/100	0,415226	100/100	0,000009	100/100	0,775562	100/100
FFT	0,000022	57/100	0,014550	100/100	0,025256	99/100	0,000170	95/100	0,125456	100/100
Serial 1	0,000000	0/100	0,719747	100/100	0,0965578	99/100	0,000587	33/100	0,935716	99/100
Serial 2	0,000032	30/100	0,779188	100/100	0,051942	99/100	0,002881	95/100	0,289667	99/100
Approx. Entropy	0,000000	0/100	0,964295	98/100	0,699313	97/100	0,000005	9/100	0,456656	99/100
Non-Overl. Temp	Fail	Fail	Pass	Pass	Pass	Pass	Fail	Fail	Pass	Pass

6. Conclusion

Des expériences pratiques sur des circuits de tests composés de TAS-MRAM réelles ont été réalisées pour analyser l'efficacité d'un TRNG basé sur les JTM-TAS. Une étude sur le comportement stochastique de la commutation des JTM lors de l'écriture des jonctions TAS en tant que source d'aléa a été menée. Les résultats ont montré que la qualité aléatoire des nombres générés sans correcteur d'entropie n'est pas suffisante pour passer des tests d'évaluations statistiques standard, comme le NIST SP 800-22. La méthode de post-traitement à base de filtre de parité (XOR) et la correction de Von Neumann améliorent la qualité aléatoire des bits générés permettant de passer la suite de tests de NIST. De plus, une méthode adaptative incluant une boucle de rétroaction a été étudiée dans ce travail. Cette méthode a donné de meilleurs résultats en termes de nombre de jonctions nécessaire, en la combinant avec la méthode de correction XOR pour passer les tests de NIST. En conclusion il a été mis en évidence que l'on peut facilement concevoir des TRNG dans un circuit s'il comprend des éléments de mémoire basés sur la technologie TAS-MRAM.

Les travaux présentés dans ce chapitre ont fait l'objet d'une publication dans la revue IEEE Access [115].

Le chapitre suivant va aborder l'utilisation des MRAM dans la conception d'une autre primitive de sécurité, les PUF. Les résultats de l'utilisations des JTM TAS en particulier seront présentés.

Chapitre V Fonctions Non Clonables Physiquement à base de MRAM

1. Introduction

Le concept de fonction non clonable physiquement a été introduit pour la première fois par K. Lofstrom et al. dans les années 2000[116]. Ils proposent d'exploiter les variations de procédés de fabrication affectant les transistors pour l'identification des circuits intégrés (IC). L'acronyme PUF (*pour physically uncloonable function*) désignant les fonctions physiques non clonable sera présenté par la suite par Gassend et al. [117].

Une PUF est une fonction non mathématique intégrée dans un dispositif afin d'extraire un secret d'un système physique complexe. Elle peut être décrit comme étant une fonction qui renvoie une valeur caractéristique (une signature ou une empreinte digitale) d'un circuit intégré soumis à un défi ou challenge donné. Par conséquent, elle peut être définit comme une procédure de challenge-réponse physique pour extraire la signature d'un circuit intégré. Une PUF est similaire aux empreintes digitales humaines car elle produit une signature spécifique d'un appareil pour permettre l'authentification de cet appareil. En outre, elle tire principalement parti des variations des procédés de fabrication inhérentes à la conception des circuits intégrés, car ces variations sont difficilement reproductibles.

La MRAM étant également soumis à ces variations, est un candidat potentiel pour la conception des PUF. En effet une exploitation astucieuse des variations affectant les dimensions des jonctions d'une matrice de JTM permet d'en extraire des signatures uniques pour chaque matrice.

Différentes PUF basées sur les MRAM ont été étudiées et proposées dans la littérature. Cependant, la plupart des PUF à base de MRAM étudiées dans la littérature ont été réalisés dans des environnements de simulation. Malgré les connaissances précieuses obtenues dans le cadre de ces études, le manque d'expériences pratiques est très tangible. Dans ce chapitre, nous présentons des résultats d'expérimentations et des analyses sur le comportement de commutation stochastique des TAS-MRAM pour la conception de PUF. Ces travaux ont été menés avec un collègue Post-Doctorant Arash Nejat au sein du LIRMM et en collaboration avec l'équipe du Pr. Bertrand Cambou de la Northern Arizona University.

La suite du chapitre est organisée comme suit : la section 2 présente les propriétés et les paramètres généraux associés aux PUF, puis la section 3 fait une classification des différentes PUF, les méthodes d'évaluation des PUF sont ensuite abordées dans la section 4 suivi par un

état de l'art des PUF à base de MRAM dans la section 5 nous permettant ensuite de présenter les résultats expérimentaux sur la conception de PUF à base de TAS-MRAM dans la section 6 enfin, la section 7 conclut ce chapitre.

2. Propriétés et paramètres des PUF

Deux propriétés principales furent d'abord données lors de la première définition des PUF par les auteurs dans [117], **l'imprévisibilité** et l'**unicité physique**. La PUF est reconnue comme imprévisible lorsqu'un attaquant pouvant utiliser une quantité limitée et fixe de ressources ne peut extraire qu'une quantité négligeable d'informations de sa réponse secrète. Quant à l'unicité physique, cela signifie qu'il devrait être techniquement très difficile, pour ne pas dire impossible, de produire deux PUF identiques. Les PUF profitent des caractéristiques du circuit, qui sont liées à la variation aléatoire incontrôlable du processus de fabrication. Par conséquent, moins le contrôle de la précision du circuit est important pendant le processus de fabrication, plus la reproduction à l'identique d'une PUF est difficile. Plus tard, Maes et al. [118] étendent la liste des propriétés qu'un PUF doit avoir. Ils présentent alors quatre propriétés qui ont été identifiées à partir de plusieurs définitions de PUF proposées dans la littérature :

- **Evaluable** : Cela signifie que, d'un point de vue pratique, le circuit de mesure utilisé doit être facile à mettre en œuvre et très peu coûteux (par exemple en utilisant des composants standard). D'un point de vue théorique, la réponse du PUF doit être facile à évaluer ou à produire (évaluable en utilisant un temps fixe et limité).
- **Unicité** : La réponse d'une PUF est extraite comme étant l'identité du dispositif physique. Ainsi, en théorie, un ensemble de paires challenge-réponse (*CRP, challenge response pair*) devrait suffire pour identifier de manière unique une PUF au sein d'une population donnée.
- **Reproductible** : Cette propriété distingue les PUF des TRNG. En effet, la réponse d'une PUF doit être reproductible (avec une tolérance de quelques petites erreurs) lorsqu'on lui soumet le même défi ou challenge, même dans des conditions environnementales différentes.
- **Sécurisé** : cette propriété peut être divisée en trois propriétés ;
  - ✓ *Non modélisable* : cela signifie qu'il doit être difficile voire impossible de construire une fonction capable de reproduire toutes les paires de challenge-réponses d'une PUF.
  - ✓ *Unidirectionnel :* la réponse d'une PUF ne doit pas permettre de prévoir ou deviner le challenge à appliquer.

- ✓ *L'inviolabilité* : Les PUF devraient produire une réponse d'erreur lorsqu'elles sont soumises à une attaque invasive pour éviter le clonage par exemple.

Handshuh et ses collègues ont proposé une autre propriété intéressante appelée "*caractère aléatoire*" dans [119]. Elle permet d'évaluer la qualité statistique de la réponse secrète générée. Afin d'extraire une clé secrète de haute qualité d'une PUF, une quantité suffisante de hasard est nécessaire dans les réponses.

Il convient de noter que parmi toutes les PUF proposés dans la littérature, aucun ne vérifie toutes ces propriétés à 100%. La plupart d'entre elles répondent à la majorité des propriétés requises jusqu'à un certain pourcentage. Cependant, il ressort que les deux propriétés les plus importantes à vérifier sont l'unicité et la reproductibilité. Ainsi on distingue deux paramètres importants que sont les variations *intra-distantes* et *inter-distantes*, pour évaluer respectivement l'unicité et la reproductibilité d'un dispositif PUF.

2.1. Intra-distance ou la reproductibilité

L'intra-distance est un paramètre qui décrit la différence entre les réponses d'une même PUF soumise à l'application d'un même challenge. Pour mieux caractériser la stabilité d'une instance PUF, il est très important de connaître la répartition de ce paramètre dans des conditions environnementales normales et anormales. Ainsi, les statistiques de cette répartition sont souvent utilisées comme mesure pour évaluer la reproductibilité ou la fiabilité d'une instance PUF. Il est à noter que la valeur moyenne de la variation intra-distance est plus élevée lorsque l'évaluation est faite dans des conditions environnementales anormales. Par conséquent, pour mieux évaluer la stabilité des PUF, il convient de prendre en compte le pire cas qui est représenté par la plus grande valeur de l'intra-distance.

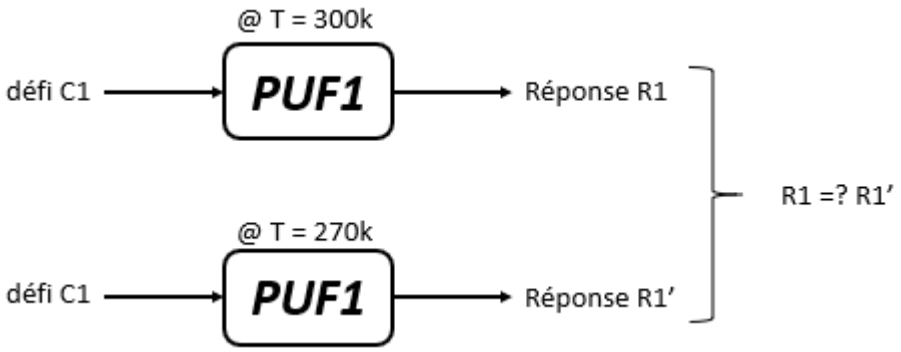


Figure V-1 : Evaluation de la reproductibilité de la réponse d'une PUF sous différentes températures de fonctionnement

La Figure V-1 illustre un processus d'évaluation de la reproductibilité de la réponse d'une PUF fonctionnant à deux températures différentes.

Une grande fiabilité est souhaitable dans toutes les applications, mais elle est absolument essentielle dans des applications telles que la génération de clé cryptographique. La méthode classique pour améliorer la fiabilité consiste à utiliser des codes de correction d'erreur (*ECC pour error correction code*) pour corriger la réponse brute des PUF [120]. Malheureusement, ces blocs ECC engendrent des coûts d'implémentation supplémentaire et augmentent rapidement au fur et à mesure que la capacité du bloc de correction d'erreur augmente. De plus, certaines techniques de correction d'erreur doivent être mises en œuvre avec soin pour éviter l'introduction d'une vulnérabilité de sécurité par la fuite des informations sur les bits de réponse [120].

2.2. Inter-distance ou l'unicité

L'inter-distance est également un paramètre permettant de mesurer la différence entre les réponses de plusieurs PUF (inclus dans une même puce ou dans des puces différentes) lors de l'application d'un même défi ou challenge. C'est la mesure de la capacité d'une instance PUF à avoir un comportement unique et distinct par rapport à d'autres PUF de même structure, implémentés sur des puces différentes. L'inter-distance mesure donc l'absence de corrélation entre les bits de réponse des différentes instances PUF, et idéalement les bits de réponse devraient différer avec une probabilité de 50%.

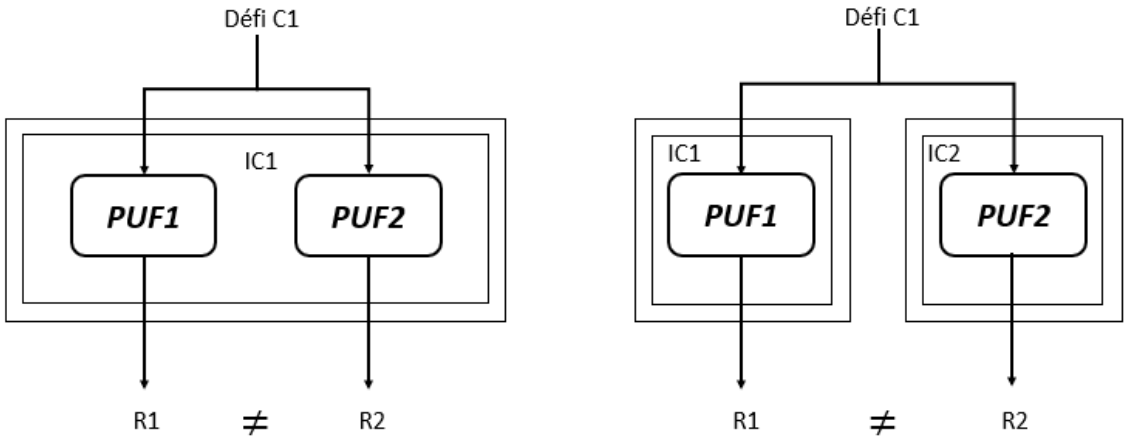


Figure V-2 : Evaluation de l'unicité de la réponse des PUF

3. La Classification des PUF

Dans la littérature, les PUF sont classées différemment en fonction des critères considérés. Une classification plus détaillée des PUF a été étudiée dans [121] en considérant les mécanismes physiques mises en œuvres dans leurs conceptions. De façon générale on peut distinguer les PUF en fonction du matériel utilisé (dispositif silicium ou non) et de leur niveau de sécurité (PUF fort ou PUF faible).

3.1. Les PUF à base de dispositif silicium et non-silicium

3.1.1. Les PUF à base de dispositif non-silicium

Cette catégorie regroupe tous les PUF dont le matériel de base de sa conception n'est pas un dispositif silicium. Il convient alors de noter ici que les PUF pourraient, en principe, être construits à partir de n'importe quelle entité physique. Les PUF optiques qui exploitent la diffusion aléatoire de la lumière [122][123], et les conceptions à radiofréquence (RF) qui sont basées sur la caractéristique unique des ondes électromagnétiques émises par un dispositif ou un circuit intégré en fonctionnement [124][125] sont des exemples de conceptions de PUF ne reposant pas sur des dispositifs silicium. Il existe également des conceptions plus exotiques de méthode d'identification d'objets sécurisés, comme ceux basés sur l'exploitation des imperfections microscopiques de la surface d'un morceau de papier pour générer une signature unique [126][127].

3.1.2. PUF à base de dispositif silicium

Les PUF silicium utilisent les variations des procédés de fabrication incontrôlables pour générer une signature unique pour chaque circuit intégré. Selon les différentes sources de variation, les PUF sous silicium peuvent être classées en PUF à base d'électroniques analogiques, en PUF basées sur le délai, et en PUF basées sur la mémoire.

3.1.3. PUF à base d'électronique analogique

Les PUF basées sur les composants analogiques exploitent les variations aléatoires pouvant affecter les paramètres de ces composants.

Lofstrom et al. [116] ont proposé d'exploiter les fluctuations des courants de drain pour l'identification du circuit intégré (ICID), qui est une identification unique pour chaque circuit fabriqué. Cependant, cette méthode présente des CRP limités. Elle n'est donc pas sûre, car les adversaires peuvent lire l'ID de manière exhaustive et « cloner » le circuit intégré. Tuyls et al. [128] ont introduit en complément des PUF, un revêtement qui les rend résistants aux

attaques invasives puisque ce revêtement protecteur est opaque et chimiquement inerte. Le revêtement protecteur est constitué d'un matériau dopé avec des particules diélectriques aléatoires recouvrant le circuit. La couche métallique supérieure du circuit intégré contient un ensemble de capteurs qui sont utilisés pour mesurer les valeurs de capacité locale du revêtement. Un défi ou challenge correspond à une tension d'une certaine fréquence et d'une certaine amplitude appliquée aux capteurs à un point donné du réseau de capteurs. La réponse est la valeur de capacité mesurée qui est convertie en une chaîne de bits sous forme de clé ou d'identifiant. B. Škorić et al. ont proposé une nano-clef en silicium qui exploite la variabilité des paramètres électriques des transistors MOS de taille minimale, en particulier la tension de seuil, pour générer un PUF utilisable dans un système d'authentification dans [129].

3.1.4. PUF basée sur le délai

Les PUF basées sur le délai (certainement les plus classiques dans la littérature) comparent le temps de propagation de deux chemins nominalement identiques. L'idée de base est présentée sur la Figure V-3. Les deux chemins $CH1$ et $CH2$ ont le même délai nominal $T_{nominal}$.

Toutefois, leurs délais réels, T_1 et T_2, sont différents en raison de leurs composantes aléatoires $T_{random1}$ et $T_{random2}$. Un bit de réponse dans un PUF de délai est généré en comparant (directement ou indirectement) T_1 et T_2, ce qui revient en fait à comparer $T_{random1}$ et $T_{random2}$. Si $T_{random1}$ et $T_{random2}$ qui sont distribués de manière aléatoire et identique, le bit de réponse généré en comparant T1 et T2 sera un bit aléatoire. Il existe de nombreuses implémentations de PUF basés sur les délais mais les plus couramment utilisés sont les PUF d'arbitrage et les PUF basées sur les oscillateurs en anneau.

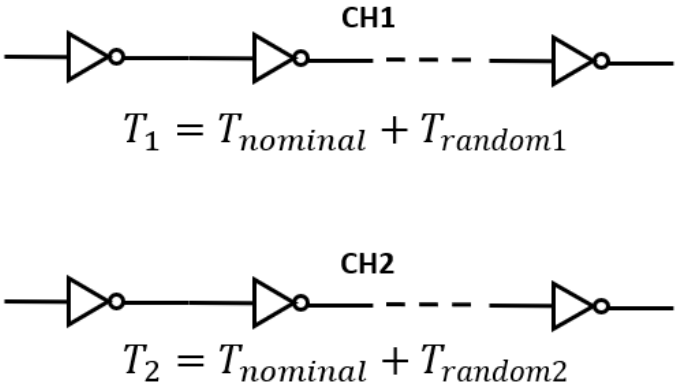


Figure V-3 : Principe de base des PUF à base de délai à travers la comparaison de deux temps de propagation de deux chemins identiques

Les PUF d'arbitrage (Arbiter PUF) : ce sont des PUF basées sur le délai et qui génèrent leurs bits de réponse en comparant le temps de propagation de deux chemins nominalement identiques, mais avec des délais légèrement différents dus aux variations des procédés de fabrication. L'idée de base est celle qui est présentée dans la Figure V-4. Les chemins d'une PUF d'arbitrage se connectent à un bloc décisionnel (l'arbitre), qui en général est un comparateur de phase, et qui génère un « 0 » ou un « 1 » selon le chemin le plus rapide. Mais en pratique, pour avoir un grand nombre de CRP, les inverseurs de la Figure V-4 sont remplacés par des multiplexeurs configurables comme le montre la Figure V-5.

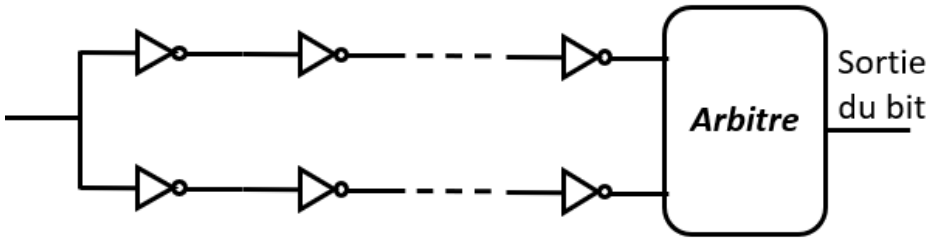


Figure V-4 : Principe de base de la PUF d'arbitrage

L'arbitre peut être implémenté en utilisant une bascule D. Idéalement, une PUF d'arbitrage devrait être capable de générer 2^N bits de réponse non biaisés et aléatoires pour chacune des 2^N combinaisons possibles de challenge.

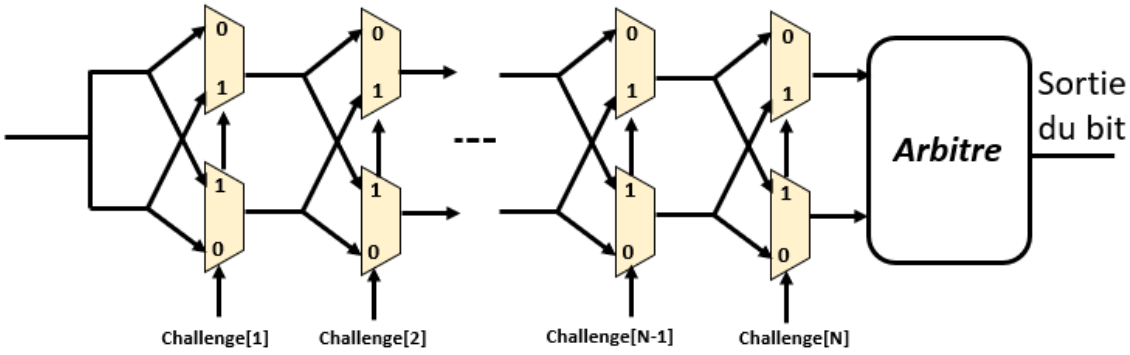


Figure V-5 : Une PUF d'arbitrage typique avec deux chemins de propagation nominalement identiques suivis par un arbitre pour générer un bit de réponse en fonction du chemin le plus rapide

Cependant, les PUF d'arbitrage sont vulnérables aux attaques de modélisation : si l'on suppose que les deux chemins sont composés de N composants de commutation (multiplexeurs), alors 2^N possibilités différentes de bits de contrôle (challenges) peuvent être appliquées. La réponse binaire du PUF est linéairement dépendante des N éléments de retard. Cela signifie que les réponses à 2^N bits ne peuvent pas être indépendantes. Il est alors possible de modéliser le circuit comme un modèle additif en considérant le délai total du chemin

comme étant la somme de tous les délais élémentaires des multiplexeurs et une fois que l'on connait ces délais élémentaires et la relation avec les bits du challenge appliqué, on est capable de prédire le bit de réponse à un challenge donné, même si on n'a pas accès à la PUF. La structure de la PUF d'arbitrage est de ce fait mathématiquement clonable. Plusieurs modifications de l'arbitre ont été proposées pour améliorer la résistance aux attaques par modélisation [130][131]. Malheureusement, il a été démontré que de simples algorithmes d'apprentissage machine peuvent être utilisés pour construire des modèles précis, même dans le cas où le circuit est modifié [132].

Les PUF basées sur les oscillateurs en anneau (RO-PUF) : Les RO-PUF comparent la fréquence de deux oscillateurs en anneau nominalement identiques pour la génération d'un bit de réponse. L'idée de base est présentée à la Figure V-6.

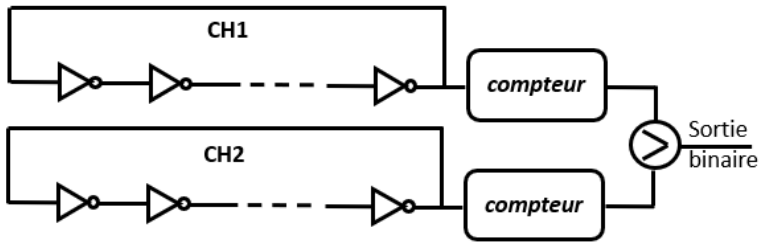


Figure V-6 : Principe de base d'une PUF à base d'oscillateur en anneau

La fréquence des oscillateurs en anneau sélectionnés est comptée à l'aide d'un compteur, puis les valeurs du compteur sont comparées pour générer un bit de sortie.

La Figure V-7 [133] illustre une implémentation typique du RO-PUF composé de plusieurs RO identiques, suivis de deux grands multiplexeurs pour sélectionner les deux RO à comparer. Conceptuellement, un RO-PUF fonctionne d'une manière très similaire à un PUF d'arbitrage. Tandis qu'une PUF d'arbitrage dispose d'un comparateur de phase (c'est-à-dire l'arbitre) pour amplifier la différence de délai entre deux chemins identiques, le RO-PUF amplifie la différence de fréquence des oscillateurs en anneau sélectionnés en leur permettant d'osciller pendant un temps suffisamment long pour que les compteurs de mesure de fréquence génèrent au moins un bit de différence.

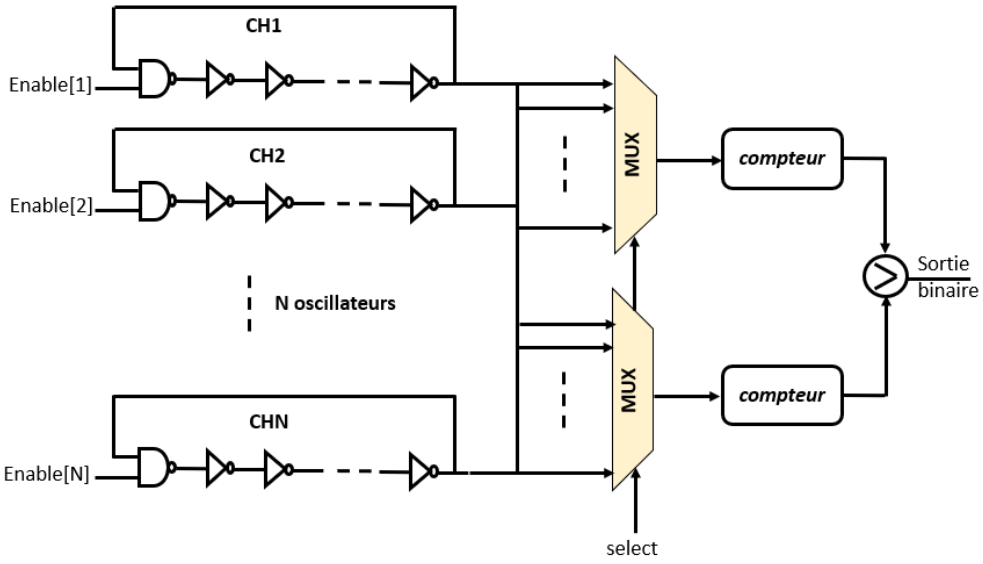


Figure V-7 : PUF à base d'oscillateur en anneau présenté dans [133]

3.1.5. Les PUF basées sur les mémoires

Les PUF basées sur les mémoires utilisent la résolution de la métastabilité des composants bistables qui les composent (« 0 » ou « 1 »). Pour générer un bit aléatoire, le composant bistable est initialisé à un état intermédiaire ou métastable (ni « 0 » ni « 1 »), puis son basculement à un état stable est laissé aux effets des variations de procédés de fabrication. Si les variations de procédés de fabrication sont principalement de nature aléatoire, alors l'état final du composant est aléatoire et imprévisible.

Les PUF à base de SRAM ainsi que les PUF dite « Butterfly » sont les plus utilisés dans la catégorie des PUF basés sur les mémoires.

Les PUF à base de SRAM : les PUF à base de SRAM sont l'une des premières conceptions à apparaître dans cette catégorie, elles ont été initialement proposées pour sécuriser les conceptions FPGA en générant une clé de cryptage spécifique à chaque FPGA pour brouiller les flux binaires avant de les stocker dans une mémoire externe [134].

La Figure V-8 (a) montre une cellule SRAM typique, cette dernière est composée de deux inverseurs rebouclés (P1, N1, P2, N2) et de deux transistors d'accès de type N (N3, N4). Les deux inverseurs ont deux états stables, logique « 1 » ou logique « 0 », auxquels on peut accéder par N3 et N4 par deux lignes de bits, à savoir « BL » et « BLB » (le complément de « BL »). Chaque inverseur pilote l'un des deux nœuds d'état, « Q » ou « QB ». L'opération d'accès est sous le contrôle d'une ligne de mots appelée « WL ». Lorsque cette cellule est mise sous tension, les deux inverseurs rebouclés entrent dans une « compétition » ; le « vainqueur » sera finalement déterminé par la différence de force de commande des MOSFET

dans les inverseurs. La cellule SRAM a essentiellement trois états possibles, dont deux sont « stables » et le troisième est « métastable », comme le montre la Figure V-8 (b). Si les transistors qui composent les inverseurs sont parfaitement adaptés, la SRAM peut rester dans un état métastable « indéfiniment » lors de la première mise sous tension. En réalité, bien que ces transistors soient conçus pour avoir des tailles nominales identiques, des variations aléatoires dans le processus de fabrication du silicium garantissent qu'un inverseur aura un courant d'attaque plus fort que l'autre inverseur, ce qui permet d'avoir un état stable de la cellule à la mise sous tension.

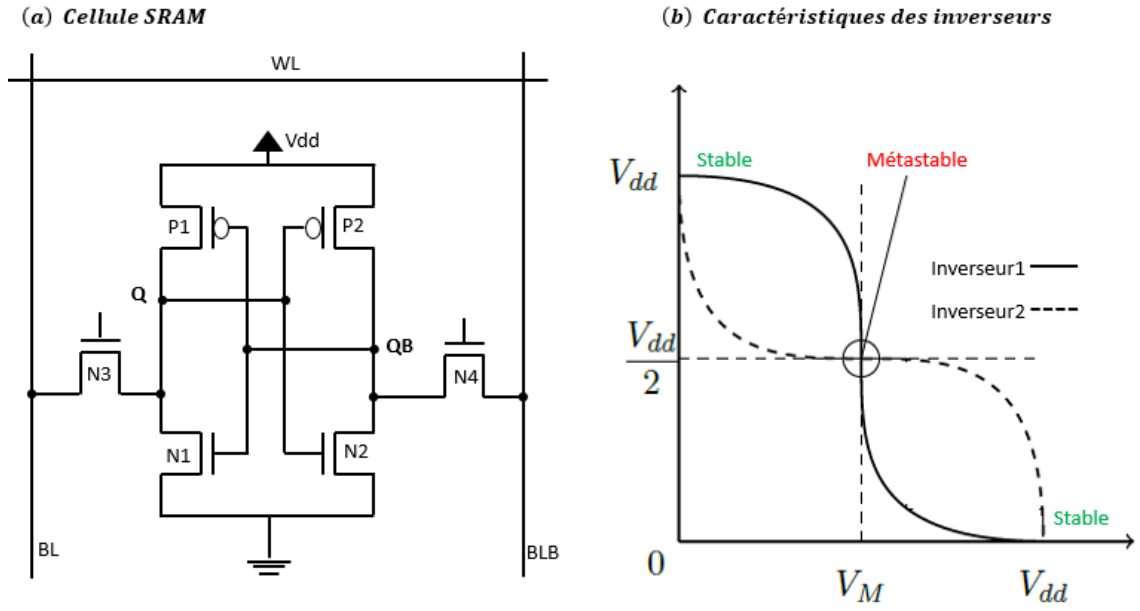


Figure V-8 : (a) Cellule SRAM typique à 6 transistors et ses caractéristiques en tension des inverseurs rebouclés

Majoritairement, les cellules SRAM ont un état initial préféré spécifique à la cellule, qu'elles prennent systématiquement à la mise sous tension, cette caractéristique des mémoires SRAM leur permet d'être utilisées pour des conception des PUF. Le nombre de CRP obtenues à partir d'une PUF basée sur la SRAM est proportionnel à la taille de la mémoire, dans ce cas, le défi sera l'adresse de lecture de la mémoire et la réponse sera les valeurs des cellules adressées à la mise sous tension. Par exemple, une mémoire adressable de 64 mégabits par octet comporte 8 mégabits de CRP.

PUF papillon (Butterfly PUF) : Une amélioration des PUF basées sur la SRAM appelée Butterfly-PUF a été proposée par Guajardo et al. [135]. Alors que les SRAM-PUF sont basées sur des inverseurs rebouclés, les Butterfly-PUF utilisent un circuit à couplage croisé instable, remplaçant les inverseurs par des bascules ou des verrous, comme le montre la Figure V-9. Le verrou du Butterfly-PUF, en tant que circuit de stockage d'information, peut être vide (la sortie

est à 0) ou être réinitialisé (la sortie est à 1) conférant ainsi l'avantage de ne pas nécessiter la mise hors tension et la mise sous tension pour son fonctionnement comme dans le cas des PUF à base de SRAM.

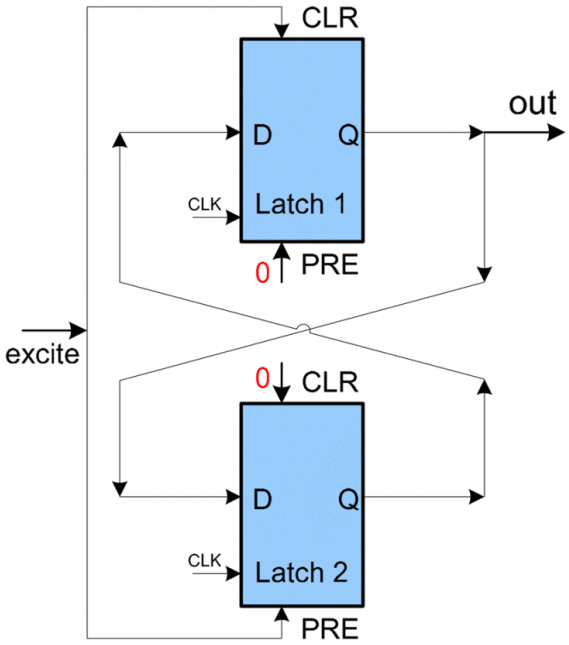


Figure V-9 : Butterfly-PUF basé sur le rebouclage de deux verrous [135]

La structure de la cellule Butterfly-PUF est construite de manière aussi symétrique que possible par l'acheminement manuel des fils du signal. Elle se compose de deux verrous, chacun avec un signal PRE actif niveau haut (qui met la sortie Q à 1) et un signal CLR actif niveau haut (qui met la sortie Q à 0). Les données D sont transférées à la sortie Q lorsque le signal CLK est au niveau haut. Lors de la conception, le signal PRE du verrou 1 et le signal CLR du verrou 2 sont toujours réglés sur le niveau bas. Le signal d'excitation est connecté au CLR du verrou 1 et au PRE du verrou 2. Les sorties des verrous sont rebouclées. Pour démarrer le fonctionnement de la PUF, le signal d'excitation est réglé au niveau haut. Cela amène le circuit Butterfly-PUF à un point de fonctionnement instable (car les deux verrous ont des signaux opposés sur leurs entrées et leurs sorties). Après quelques cycles d'horloge, le signal d'excitation est réglé au niveau bas. Cela démarre le processus du circuit PUF pour atteindre l'un des deux états stables possibles, 0 ou 1, sur le signal de sortie. L'état stable dépend des légères différences dans les délais des fils de connexion qui sont conçus en utilisant des chemins symétriques sur la matrice FPGA. Par conséquent, ces légères variations sont uniquement basées sur les caractéristiques intrinsèques du circuit intégré et varient d'un dispositif à l'autre et d'une position à l'autre dans le FPGA.

3.2. PUF fort et PUF faible

La distinction entre une PUF forte et une PUF faible a été introduite dans [134] et a ensuite été réellement définie dans [136]. Fondamentalement, la distinction est basée sur les propriétés de sécurité des paires challenge-réponse (CRP). En termes simples, une PUF est dite PUF forte si elle a un grand nombre de CRP. Les PUF faibles, en revanche, ne peuvent générer que quelques CRP. Du point de vue de la sécurité, cela signifie que pour une PUF forte, un adversaire ne peut pas prédire la réponse d'un défi aléatoire avec une forte probabilité, même avec une connaissance préalable d'un grand nombre de CRP. La plupart des implémentations de PUF, y compris les PUF à base d'oscillateur en anneau et les PUF à base de SRAM ne génèrent qu'un petit nombre de bits aléatoires et sont donc classées comme des PUF faibles. Bien que les PUF d'arbitrage puissent avoir un nombre exponentiel CRP, il a été démontré que de simples algorithmes d'apprentissage machine peuvent être utilisé pour construire des modèles précis de son comportement en connaissant un petit nombre de CRP [132][137][138].

4. Les méthodes d'évaluation des PUF

Aucun test statistique standard n'a encore été défini comme le NIST SP 800-22, ou le DIEHARD, utilisés pour évaluer la qualité des structures TRNG. Pour évaluer les structures de leurs PUF, les auteurs proposent souvent des méthodes et mesures plus ou moins différentes. Dans cette section, nous présentons une sélection des différentes méthodes proposées pour évaluer la structure des PUF. On peut les diviser en deux classes : une basée sur la distance et le poids de Hamming et l'autre basée sur des calculs statistiques.

4.1. Méthode d'évaluation basé sur le calcul de la distance ou du poids de Hamming

Comme présenté dans la section 2, les propriétés les plus importantes auxquelles une PUF doit répondre sont l'unicité et la reproductibilité de ses réponses. Le caractère aléatoire de la PUF est également souligné comme une propriété importante qui doit être vérifiée pour éviter les réponses biaisées. La plupart des PUF proposées dans la littérature sont évaluées en utilisant différents calculs de distance/poids de Hamming. Par exemple les méthodes proposées dans [139] sont les suivantes :

4.1.1. Unicité ou les variations inter-PUF (UQ) :

L'unicité d'une PUF montre à quel point les signatures générées par la PUF à partir de différentes puces sont uniques. La distance de Hamming moyenne (HD) entre les réponses de plusieurs instances d'une PUF est une estimation de la propriété d'unicité.

Parmi N puces existantes, toutes les combinaisons possibles par paires sont testées. Considérons deux puces différentes, $u\ et\ v$, ayant des réponses de $n\ bits$ $R_u\ et\ R_v$ respectivement, la distance de Hamming moyenne pour le groupe de N puces est définie comme suit :

$$UQ = \frac{2}{N.(N-1)} \sum_{u=1}^{N-1} \sum_{v=u+1}^{N} \frac{HD(R_u, R_v)}{n} \text{ X } 100\% \ (V-1)$$

La meilleure performance d'unicité est obtenue lorsque UQ est proche de 50%.

4.1.2. La reproductibilité ou les variations intra-PUF (RE)

La reproductibilité (RE) évalue la stabilité des réponses d'une instance PUF dans les conditions normales ou anormales. Une réponse de référence R_i^{ref} de $n\ bits$ est extraite de la puce i dans les conditions normales de fonctionnement. La même réponse R_i^t est extraite de la même puce i dans des conditions environnementales différentes. Ensuite, les deux réponses sont comparées en utilisant la distance de Hamming. La distance de Hamming moyenne intra-PUF est utilisée comme une estimation de la propriété de stabilité du PUF sur la puce i telle que définie ci-dessous :

$$RE = \frac{1}{T} \sum_{t=1}^{T} \frac{HD(\ R_i^{ref}, R_i^t)}{n} \text{ X } 100\% \ (V-2)$$

L'idée est de réaliser T fois le même défi pour la même PUF dans les mêmes conditions de fonctionnement et de comparer la réponse résultante à celle de référence extraite dans les conditions normales de fonctionnement. La même opération est répétée autant de fois que les conditions de fonctionnement sont différentes. La distance de Hamming moyenne obtenue est alors considérée comme la performance de fiabilité de la PUF. La valeur idéale de RE est 0%. Dans la pratique une valeur de $RE \leq 10\%$ est acceptable dans la mesure où des mécanismes de correction peuvent être utilisés pour réduire la valeur du RE.

4.1.3. L'uniformité ou le caractère aléatoire des réponses du PUF (UF)

Une méthode très simple est proposée pour évaluer le caractère aléatoire de la réponse de la PUF. L'uniformité de la réponse est évaluée en calculant le poids de Hamming de

chaque réponse R de la PUF obtenue lorsqu'un défi k est appliqué. Le caractère aléatoire d'une réponse R_k de $n\ bits$ est donné par :

$$UF = \frac{1}{n}\sum_{t=1}^{n} r_{i,t}\ \text{X}\ 100\%\ (V-3)$$

Où $r_{i,t}$ correspond au bit t des $n\ bits$ de réponse. Pour une réponse uniformément distribuée, RE doit être égale à 50 % du nombre total de bits de réponse.

4.2. Méthode d'évaluation statistiques de la qualité des PUF

En 2010, une méthode statistique pour évaluer la performance des PUF est introduite par Y. Hori et al. [140]. Ils proposent également d'évaluer d'autres indicateurs de performance en plus de ceux vu dans la section 3.1 (la diffusion, l'exactitude, l'identification correcte, etc.). La plupart des métriques proposées sont basées sur des calculs statistiques. Les performances idéales sont obtenues lorsque les résultats des mesures sont de 100%. Ci-après, nous présenterons les méthodes d'évaluation de l'unicité, de la stabilité et du caractère aléatoire telles que proposées dans [140].

4.2.1. Unicité ou les variations inter-PUF (UQ) :

Tout comme dans [139], les calculs de Hamming sont utilisés pour l'évaluation de l'unicité. La distance de Hamming de toutes les combinaisons possibles des dispositifs PUF est prise en compte pour évaluer le caractère unique d'une réponse PUF. Les auteurs proposent de procéder petit à petit en variant le défi appliqué. L'unicité moyenne de la PUF UQ est donnée par :

$$UQ = \frac{4}{K.L.N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\sum_{i=1}^{N}\sum_{j=i+1}^{N}(R_{i,k,l} \oplus R_{j,k,l})\ (V-4)$$

4.2.2. La reproductibilité ou les variations intra-PUF (RE)

Telle que définit par les auteurs, la régularité d'une réponse PUF donnée est considérée comme l'inverse de sa probabilité d'erreur. Nous indiquons par $RE_{n,k,l}$ la stabilité d'un bit l sélectionné sur la puce n lorsque nous appliquons le même défi k ($RE_{n,k,l} = 1 - error_{prob}$. Y. Hori et al. proposent donc, dans un premier temps, de calculer la probabilité d'obtenir « 1 » en appliquant le même défi k pour un bit l sélectionné. Ensuite, la probabilité d'erreur est obtenue en calculant l'entropie minimale de la valeur résultante. Elle peut être exprimée sous la forme :

$$P_{n,k,l} = \frac{1}{T}\sum_{t=1}^{T} b_{n,k,l,t} \; et \; RE_{n,k,l} = 1 + log_2 \max\left(P_{n,k,l}, 1 - P_{n,k,l}\right)$$

Alors la stabilité RE_n de la réponse d'une puce n à un défi k est définit par la moyenne de $RE_{n,k,l}$ comme suit :

$$RE_n = \frac{1}{K.L}\sum_{k=1}^{K}\sum_{l=1}^{L} RE_{n,k,l} \;\; (V-5)$$

Contrairement à la mesure de la stabilité basée sur la méthode de Hamming, ici il n'y a pas besoin d'une réponse de référence.

4.2.3. L'uniformité ou le caractère aléatoire des réponses de la PUF (UF)

Pour évaluer la répartition des « 0 » et « 1 » dans la réponse de la PUF, les auteurs proposent de calculer :

- ✓ Le nombre de « 1 » dans les réponses d'une puce n soumis à K différents défis et répéter l'opération T fois pour chaque défi.
- ✓ L'entropie minimale de la probabilité résultante

Cela peut être exprimé comme suit :

$$P_n = \frac{1}{K.T.L}\sum_{k=1}^{K}\sum_{t=1}^{T}\sum_{l=1}^{L} b_{n,k,t,l} \; et \; UF = -\, log_2 \max\left(P_n, 1 - P_n\right) \; (V-6)$$

Nous constatons que, par rapport à la métrique du poids de Hamming, cette méthode tient compte de la stabilité de la réponse de la PUF, ce qui la rend beaucoup plus précise. Elle évalue également le caractère aléatoire de tous les défis appliqués.

5. Revues sur les PUF à base de MRAM

De nombreuses études ont été menées sur l'utilisation des MRAM dans la conception des PUF. La plupart de ces circuits sont basés sur des jonctions STT et les méthodes d'extraction utilisées peuvent être classées en deux grandes catégories : celle basée sur l'écriture des JTM et celle basée sur les variations aléatoires de la résistance des JTM.

5.1. Extraction basée sur l'écriture des JTM

J. Das et al. ont été les premiers à s'intéresser à la conception de PUF à base de JTM-MRAM dans [141]. Le principe repose sur la variation des procédés de fabrication affectant la géométrie de chaque jonction individuellement. Les JTM des MRAM sont structurées avec une

anisotropie de forme leur donnant une géométrie rectangulaire ou elliptique. Cependant, en raison des variations de fabrication, la forme des Jonctions varie par rapport à celle d'un rectangle parfait ou d'une ellipse parfaite. Cela entraine une répartition énergétique dans la couche de stockage différente d'une JTM à l'autre. Cette variation est aléatoire et n'est pas contrôlable. Par conséquent, elle ne peut pas être clonée. L'idée de base est d'utiliser ces variations géométriques aléatoires des couches libres des JTM pour générer une signature numérique unique d'une matrice de MRAM. La Figure V-10 montre le paysage énergétique d'une JTM pour (Figure V-10 (a)) une géométrie parfaite ; et (Figure V-10 (b)), (Figure V-10 (c)) des variations de la géométrie de la couche libre. Pour (Figure V-10 (b)), (Figure V-10 (c)) la magnétisation au point A' développe une préférence pour l'un des deux états.

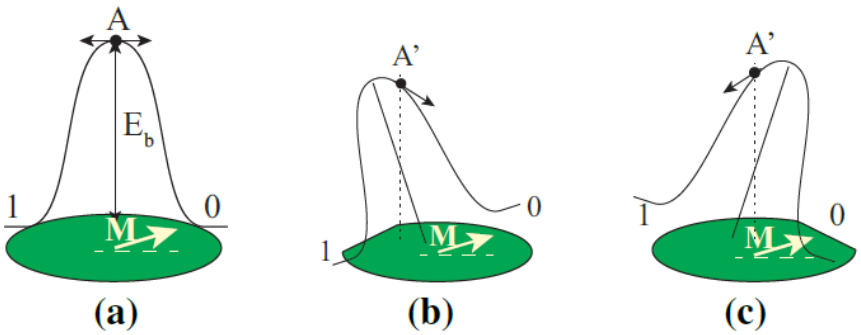


Figure V-10 : Paysage énergétique dans (a) une JTM à géométrie parfaite, (b) et (c) une JTM avec des variations dans la geometrie de la couche de stockage

En relâchant la magnétisation M au point A', celle-ci se relaxera alors vers l'état le plus proche (ou l'état préféré). Pour cela ils appliquent un champ magnétique externe relativement faible à une matrice de 10X20 JTM et ils observent la répartition des JTM écrites à « 0 » et celles à « 1 ».

Les auteurs dans [142] et [143] utilisent la dépendance temporelle associée à la probabilité de commutation lors de l'écriture des jonctions STT. L'idée est d'appliquer une impulsion d'écriture avec une probabilité de commutation théorique de 50% à une matrice de JTM-STT. En raison des variations de procédés de fabrication affectant les JTM, la probabilité de commutation réelle de chaque JTM est différente. Ainsi on aura trois catégories de JTM dans la matrice :

- JTM dont la probabilité de commutation est inférieure à 50% ($P_{sw} \leq 50\%$)
- JTM dont la probabilité de commutation est autour de 50% ($P_{sw} \approx 50\%$)
- JTM dont la probabilité de commutation est supérieure à 50% ($P_{sw} \leq 50\%$)

En appliquant cette impulsion plusieurs fois à la matrice on peut isoler les JTM ayant leur probabilité de commutation très supérieure à 50% (proche de 100%) et celle ayant leur probabilité de commutation très inférieure à 50% (proche de 0%) et exclure le reste.

5.2. Extraction basée sur les variations aléatoires de la résistance des JTM

La résistance de chaque JTM dans les mémoires MRAM varie de l'une à l'autre à cause des variations des procédés de fabrication. Ainsi, le courant mesuré pendant la lecture de la résistance de la jonction varie également. Comme cette variation de résistance est une propriété intrinsèque de la MRAM, elle peut être exploitée pour générer des bits aléatoires spécifiques à une puce. En effet en comparant deux JTM initialisées au même état (« 0 » ou « 1 »), on peut générer une carte de « bit aléatoire » qui dépendra principalement des variations de procédés de fabrication affectant ces JTM.

Les auteurs dans [144] ont été les premiers à expérimenter cette méthode sur des Jonction STT. Ils utilisent pour cela des cellules différentielles (2 transistors / 2 JTM) associées à un circuit de lecture (PCSA) telle que vu dans la section Chapitre III4.3.3 du Chapitre III. Cependant les deux JTM sont écrites dans le même état et l'extraction du bit aléatoire est fait à travers le circuit de lecture dans ce cas. Ainsi pour une matrice composée de ces cellules, l'adresse mémoire et les données de la cellule sont respectivement le défi et la réponse de la PUF. Après extraction des bits aléatoires, une réécriture différentielle, cette fois-ci, des cellules est effectuée pour renforcer la fiabilité et la répétabilité des réponses générées.

D'autres travaux ont également utilisés cette méthode d'extraction [145], [146] et [147].

5.3. Synthèse

Le Tableau V-1 présente un récapitulatif des travaux effectués dans la littérature sur la conception des PUF à base de MRAM en les comparant en termes de la qualité de leurs stabilités, de leurs unicités et du caractère aléatoire des bits de réponse générés.

Cependant, presque tous ces travaux ont été réalisés dans des environnements de simulation. Malgré les connaissances précieuses obtenues dans le cadre de ces études, le manque d'expériences pratiques est très tangible. Toujours dans le cadre du projet FUI Multismart nous avons effectué des expériences d'extraction de signatures sur les JTM TAS dont nous disposions.

Tableau V-1 : Comparaison des differents PUF à base de MRAM dans la littérature

Référence	Type d'étude	Reproductibilité	Unicité	Uniformité
[141]	Simulation	0,0225	0,47	0,99
[144]	Simulation	$6{,}6\times10^{-6}$	0,51	0,985
[148]	Simulation	$6{,}6\times10^{-6}$	0,498 - 0,51	0,95
[143]	Simulation	0,05	0,458	-
[149]	Simulation	0,051 - 0,057	0,498 - 0,5	-
[150]	Expérimentale (256 bits)	-	0,50	0,99

6. Fonction non clonable physiquement basée sur la TAS-MRAM

Nous avons utilisé les jonctions TAS vu dans le chapitre précédent pour nos expérimentations. Contrairement à la génération de nombres aléatoire sur des JTM individuelles, ici l'idée est de pouvoir générer des signatures de chacune des puces que nous avons. Nous avons au total 26 puces qui ont été fabriquées et testés.

6.1. Principe d'extraction

Pour rappel nous disposons d'un circuit de test de 1kbits composé de TAS-MRAM fabriqué par la société Crocus technology où les jonctions sont individuellement adressables et accessibles pour écriture et lecture. Les JTM sont indépendantes les unes des autres en termes de lecture. Ce qui impose le choix de la méthode d'extraction basée sur l'écriture des JTM plutôt que la comparaison directe des résistances de paires de JTM écrite dans le même état. Ainsi le principe d'extraction de la signature des circuits consiste à déterminer la valeur des tensions d'écriture (tension de chauffage et tension de ligne de champ) pour lesquelles une signature unique correspondant à la réponse des JTM est obtenue.

6.2. Instrumentation et Expérimentation

Nous avons utilisé le même banc de test présenté à la Figure IV-3 du Chapitre IV. Il est composé d'une une carte FPGA Xilinx NEXSYS2 et d'une carte d'acquisition du *National Instrument*, pour écrire et lire les matrices.

Une écriture correcte d'une JTM nécessite l'application d'une tension de champ magnétique de 3.3V et d'une tension de chauffage de 2V. Par contre la lecture consiste à appliquer uniquement une tension de chauffage de 0.3V et de mesurer la tension correspondante aux bornes de la JTM. Une valeur moyenne de 240mV est obtenue dans le cas d'une lecture de la JTM à « 0 » et une valeur moyenne de 290mV est obtenue pour une lecture de la JTM à « 1 ».

Le déroulement des expériences commence par la distinction entre les JTM claquées ou défectueuses et celles intactes. Les JTM claquées ont toujours une valeur de tension inférieure à 140 mV lors de la lecture et les JTM défectueuses sont toujours collées dans l'un des états P ou AP (« 0 » ou « 1 »), et aucune opération d'écriture ne peut changer leurs états.

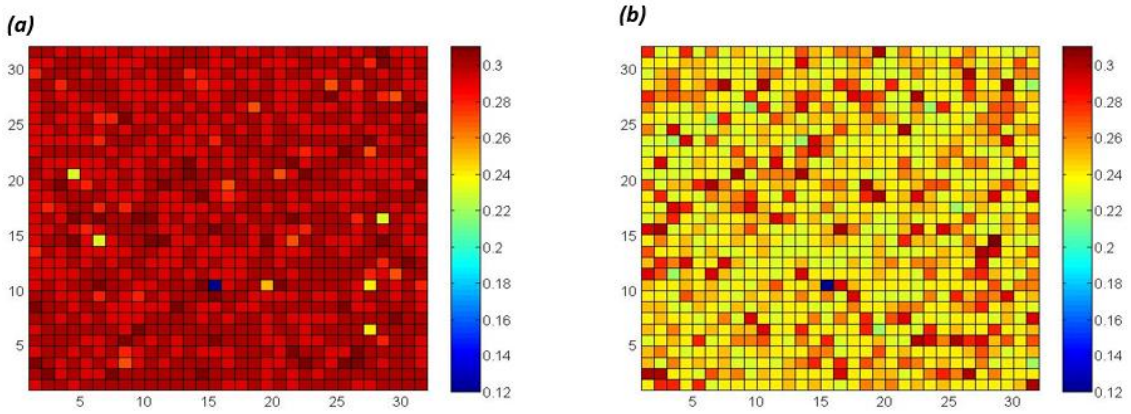


Figure V-11 : Lecture d'une matrice 32X32 de JTM TAS (a) après écriture à "1" de toutes les JTM et (b) après écriture de toutes les JTM à "0"

Les Figure V-11 montrent une répartition spatiale de l'état des JTM d'une puce suite à l'écriture correcte (tension de champs et de chauffage maximale) des jonctions à « 1 » puis à « 0 ». Comme le montre cette figure, il y a une JTM claquée (en bleue). Six JTM qui sont collées à « 0 » (en jaune) avec une valeur comprise entre 220 mV et 250 mV et qui sont visibles sur la Figure V-11 (*a*), et enfin 21 JTM qui sont collées à « 0 » (rouge) avec une valeur autour de 300 mV et qui sont visibles sur la Figure II-1 (*b*).

La deuxième étape de nos expériences a consisté à analyser les différentes tensions de chauffage afin de trouver celles pour lesquelles on obtient une distance de Hamming de 50% pour la variation inter-PUF. Pour ce faire, nous appliquons des séquences d'écritures CW1-UW0 correspondant à une écriture correcte à « 1 » (CW1) suivi d'une écriture aléatoire à « 0 » (CW0). L'expérience commence à partir d'une tension de chauffage initiale à 1,8 V. La séquence CW1-UW0 est appliquée à toutes les JTM des 26 puces fabriquées. Ensuite, 0,1 V

est déduit de la tension de chauffage, et à nouveau, la séquence est répétée. Cette procédure est poursuivie jusqu'à ce que la tension de chauffage atteigne 1 V.

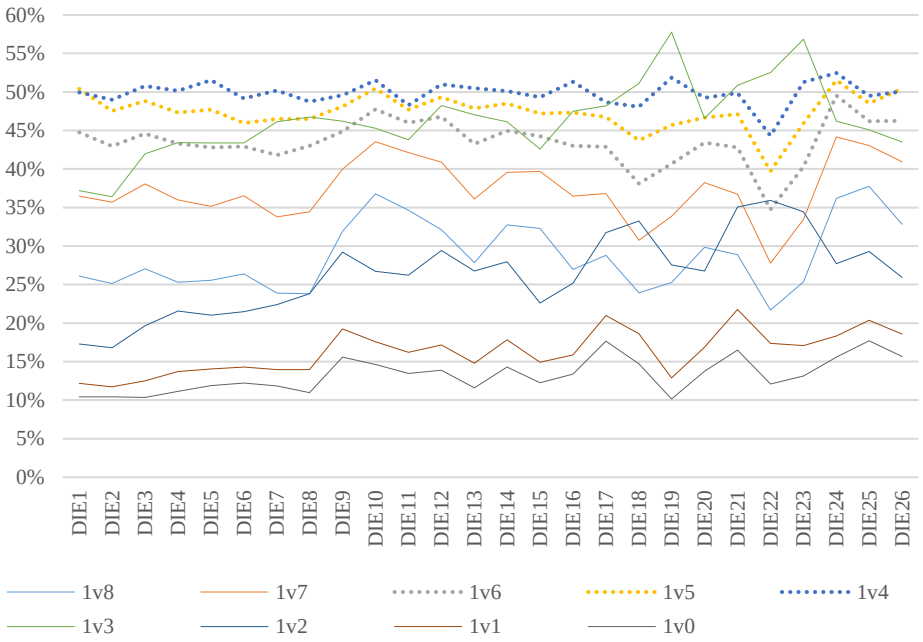


Figure V-12 : Variations inter-PUF correspondant à chaque puce pour différentes tensions de chauffage

La Figure V-12 montre la variation inter-PUF, obtenue à partir de l'équation ($V-1$), pour chacune des 26 puces sous différentes tensions de chauffage. Les lignes en pointillés de cette figure montrent que pour les tensions de chauffage comprises entre 1,4 V et 1,6 V, les variations inter-PUF pour la majorité des puces sont comprises entre 45% et 55%, ce qui est une fourchette acceptable. La Figure V-13 montre les variations inter-PUF moyenne de toutes les 26 puces pour les tensions de chauffage utilisées. La barre verte en pointillée de la figure représente le meilleur résultat de la variation inter-PUF qui est de 49,8% pour 1,4 V.

La troisième étape de nos expériences a consisté à mesurer la variation intra-PUF. Pour ce faire, nous avons appliqués 1000 fois la séquence de CW1-UW0 sur toutes les JTM des 26 puces. Comme précédemment, 9 tensions de chauffage comprises entre 1,8 V et 1 V sont analysées. Les résultats sont présentés sur la Figure V-14 et la Figure V-15. La première montre les variations intra-PUF correspondant à chaque puce pour différentes tensions de chauffage. La Figure V-15 montre quant à elle, les variations intra-PUF moyennes de toutes les puces pour les différentes tensions de chauffage. Dans ces figures, les lignes et les barres correspondant à la tension de chauffage de 1,4 V sont en pointillés (en vert). Pour cette valeur de tension, la variation intra-PUF, obtenue à partir de l'équation ($V-2$), est de 7,7%. Ce taux est acceptable car il peut être corrigé en utilisant des méthodes de correction d'erreurs.

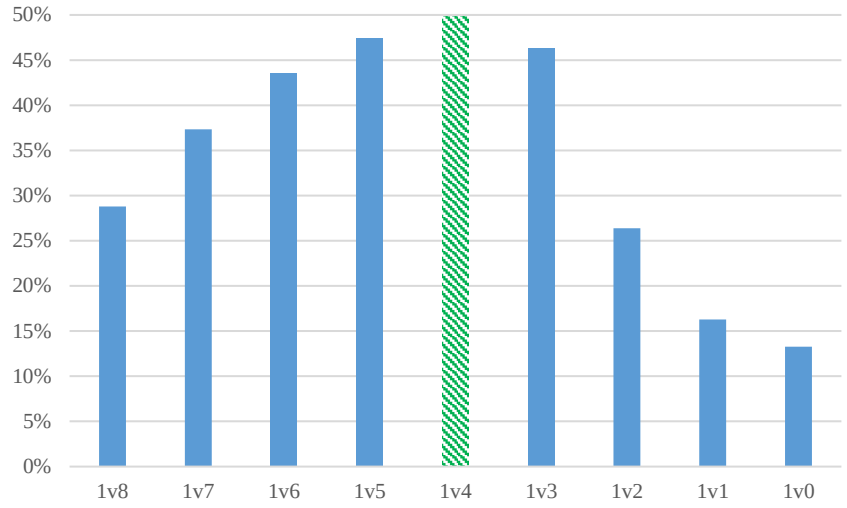


Figure V-13 : Variation inter-PUF entre toutes les 26 puces pour différentes tensions de chauffage

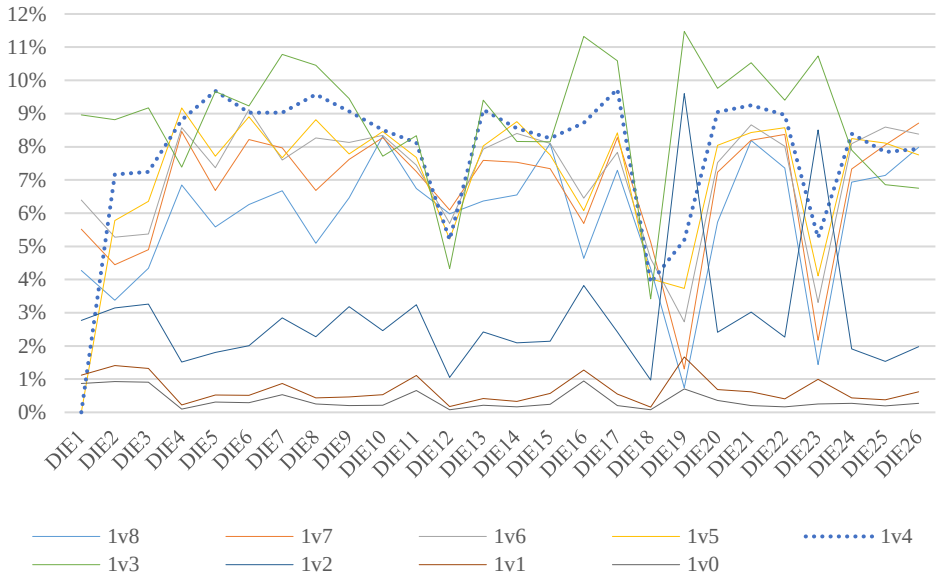


Figure V-14 : Variation intra-PUF correspondant à chaque puce pour différentes tensions de chauffage

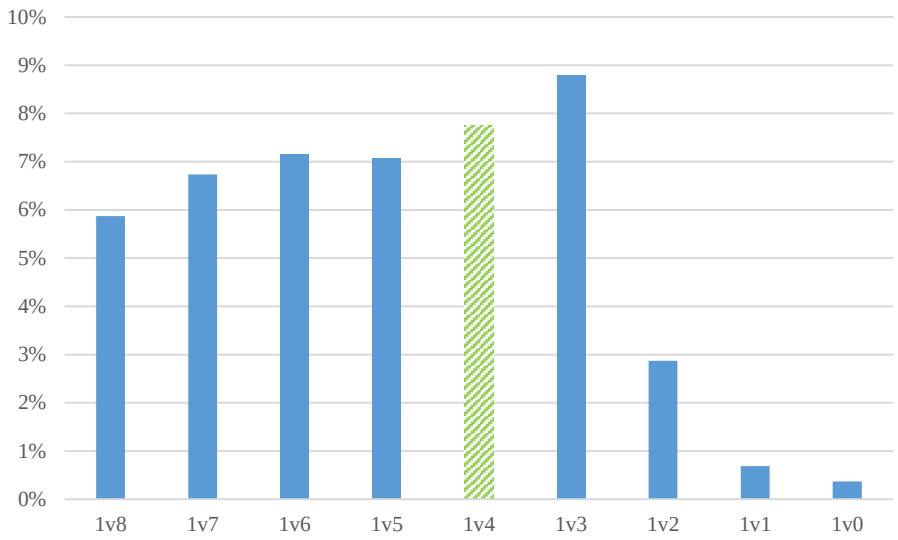


Figure V-15 : Variations intra-PUF moyenne entre toutes les puces pour différentes tensions de chauffage

Enfin pour le calcul de l'uniformité, nous avons déterminer le rapport entre le nombre de "0" et le nombre de "1" dans chaque réponse. La Figure V-16 montre l'uniformité moyenne entre toutes les 26 puces pour les différentes valeurs de tensions de chauffage. Comme le montre la figure, le meilleur résultat qui est égal à 0,94, est obtenu en utilisant la tension de chauffage de 1,4v.

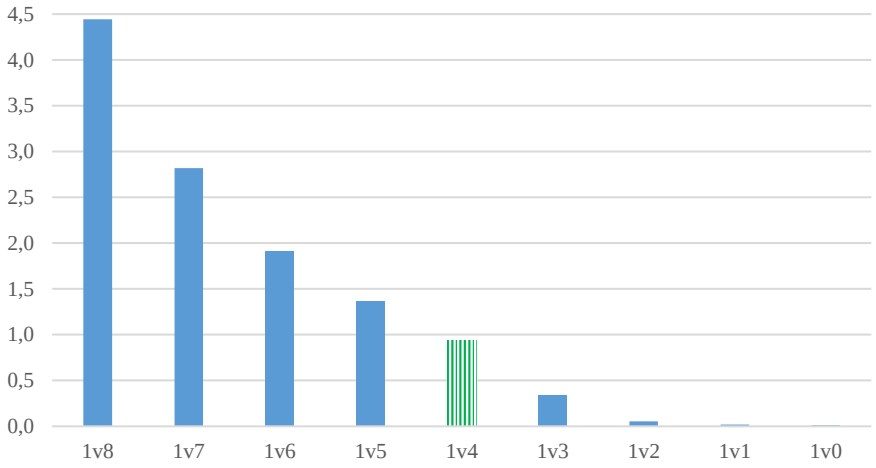


Figure V-16 : Uniformité moyenne entre toutes les matrices pour différentes valeurs de tensions de chauffage

En comparant nos résultats à ceux des précédentes publications sur les PUF à base de MRAM présentés dans le Tableau V-1. On remarque que dans [140], des résultats prometteurs de PUF à base de MRAM sont présentés. Cependant, les résultats rapportés sont basés sur des études de simulation, qui ne rendent pas pleinement compte de la réalité physique. En

comparaison, nos résultats sont basés sur des expériences pratiques sur des matrices TAS-MRAM fabriquées. Dans [148], les résultats sont basés sur des expériences sur des mémoires STT-MRAM de 256 bits tandis que nos résultats sont basés sur des analyses faites sur une taille de mémoire un peu plus grande (1 Kbits). Nos résultats montrent une bonne valeur de l'inter-PUF (49.8%) entre différentes matrices PUF basées sur les TAS-MRAM. Cependant, la valeur de l'intra-PUF (7.7%) obtenu est supérieure à celles des autres études rapportées.

7. Conclusion

Dans ce chapitre nous avons abordé la conception des PUF qui sont des primitives de sécurité importante dans le domaine de la cryptographie. Nous avons d'abord fait une brève historique des PUF puis nous nous sommes focalisés sur l'utilisation des PUF à base de MRAM. Plusieurs travaux ont déjà été menés dans ce sens dans la littérature à travers des simulations et ont montrés l'intérêt et l'avantage que pourrait avoir une telle implémentation. Néanmoins, le manque d'expérimentations pratiques sur des puces réelles est assez tangible. Nous avons ainsi mené des expériences pratiques sur la conception de PUF à base des jonction TAS sur 26 puces de $1Kbits$ chacun à notre disposition. Les expériences visaient à analyser l'efficacité d'une PUF conçue avec ces puces. Le comportement de commutation stochastique des JTM comme source d'extraction a fait l'objet de l'étude. Nos résultats expérimentaux ont montré que nous pouvons disposer d'une PUF basée sur les TAS-MRAM avec une variation inter-PUF de près de 50% et une uniformité de 94% en acceptant un taux d'erreur d'environ 7,7% correspondant à la variation intra-PUF, sans modification du circuit. Pour y parvenir, nous devons exécuter sur les $1kbits$ de TAS-MRAM, une séquence d'initialisation de toutes les JTM à « 1 » puis appliquer une séquence d'écriture aléatoire avec une valeur de la tension de chauffage à 1,4 V. La principale conclusion est qu'il est possible d'utiliser des matrices de jonctions TAS pour la conception de PUF en y associant un code correcteur d'erreur.

Les travaux présentés dans ce chapitre ont fait l'objet d'une publication dans la revue IEEE Access[151].

Chapitre VI Conclusion générale et perspectives

1. Conclusion générale

Face à la difficulté pour faire face à l'évolution de la loi de Moore ces dernières années, l'industrie de la micro-électronique explore des alternatives à travers des conceptions hybrides combinant les dispositifs CMOS aux dispositifs émergeants et compatibles CMOS. L'une de ces technologies prometteuses est la MRAM qui est une mémoire non volatile émergeante à très fort potentiel et offrant des possibilités dans plusieurs domaines d'applications.

Un des domaines d'applications de cette technologie est la sécurité qui est un enjeu majeur pour l'industrie électronique. En effet nous avons vu que les propriétés internes des MRAM les rendent adaptées à la conception de primitives cryptographiques émergentes telles que les TRNG et les PUF. La majeure partie des études sur l'utilisation des MRAM dans la conception de ces primitives de cryptographie, dans la littérature, a été menée dans des environnements de simulation présentant ainsi un manque criard d'expérimentations pratiques pour valider ces différentes études. Dans ce travail de thèse nous avons proposé de mener des expérimentations pratiques sur des dispositifs MRAM réels dans le but de palier à ce manque.

Dans un premier temps, nous avons effectué un état de l'art des différentes technologies MRAM développées jusqu'à nos jours en présentant les principaux phénomènes liés à la magnétorésistance et ayant abouti à la conception des MRAM. La différence majeure entre ces technologies réside dans leurs méthodes d'écriture.

Ensuite une étude sur la conception de TRNG basée sur les jonctions STT est présentée dans le Chapitre III. L'état de l'art de cette nouvelle conception de TRNG a montré une insuffisance au niveau des expérimentions pratiques sur des jonctions réelles. Nous avons ainsi proposé de concevoir un circuit basé sur cette technologie de MRAM, les étapes de la conception hybride CMOS/STT-MTJ ont été présentées jusqu'à la réalisation du démonstrateur. Les résultats de mesure sont mitigés en raison notamment des dispersions importantes observées sur les caractéristiques des jonctions. Néanmoins des comportements observés sur un des circuits semblent indiquer un fonctionnement possible de notre structure avec une amélioration du process magnétique.

Dans le Chapitre IV, nous avons également mené des expérimentations pratiques sur une autre technologie MRAM à savoir la TAS-MRAM développée par la société CROCUS technologie. Le comportement stochastique de la commutation des jonctions TAS au cours de

leur écriture a été mise en évidence. Nous avons détaillé le processus d'extraction d'aléa permettant de générer par la suite des nombres aléatoires. Pour évaluer la qualité des nombres aléatoires, le test standard NIST SP 800-22 a été utilisé. Nous avons ainsi observé des défauts statistiques sur les nombres générés sans aucun post-traitement. Cependant, ces défauts statistiques se corrigent facilement à travers un filtre de parité (XOR) ou une correction de Von Neumann. Une méthode adaptative consistant à contrôler à temps réel la probabilité de sortie du générateur permet également de réduire les défauts statistiques si on la combine avec le filtre de parité.

Enfin dans le Chapitre V, nous nous sommes intéressés à la conception de PUF basée sur les MRAM, nous avons procédé à des expérimentations sur 26 circuits de $1kbits$ chacun à base de jonction TAS-MRAM. La méthode d'extraction de la signature unique de chaque circuit est présentée. Les résultats des expériences ont montré que nous pouvons atteindre une variation inter PUF de près de 50% avec une uniformité des réponses autour de 94% et une variation intra PUF (ou taux d'erreur de la réponse du PUF) de 7.7% sans aucune correction. Ces résultats sont plutôt encourageants et permettent d'envisager l'utilisation de ces jonctions dans la mise en place de PUF dans les dispositifs électroniques comportant des jonctions de type TAS.

2. Perspectives

Les travaux réalisés pendant cette thèse montrent l'intérêt qu'offrent les MRAM en termes de primitives cryptographiques à faibles coûts. Nous avons proposé des solutions d'implémentions de ces primitives à travers des expérimentations pratiques sur des dispositifs MRAM. Néanmoins des études complémentaires sont nécessaires pour une évaluation globale de nos solutions.

Dans l'immédiat, pour le cas du TRNG, une évaluation de la robustesse du TRNG à base de TAS MRAM contre des attaques est primordiale avant une utilisation dans des systèmes cryptographiques. Il serait par exemple intéressant d'étudier le comportement du générateur soumis à un champs magnétique externe et à des variations de température. Cela pourrait amener à utiliser des détecteurs d'attaques comme contre-mesure consistant à monitorer la sortie du générateur en temps réel et à déclencher une alarme à la suite d'un certain nombre de bits de sortie identiques.

Dans le cas du PUF, une évaluation dans les mêmes conditions décrites précédemment est aussi à effectuer. Cela permet d'avoir des analyses précises de la robustesse contre les attaques physiques.

Une étude des technologies SOT MRAM dans le cadre de leur utilisation pour la génération d'aléa (TRNG et PUF) est de notre point de vue intéressante à mener dans la mesure où la SOT MRAM présente de nombreux avantages en termes de consommation énergétique et de fiabilité.

Dans un contexte de sécurité globale, la MRAM étant amené à être intégré dans des systèmes sécurisés, Il serait intéressant d'étudier sa robustesse contre les attaques par canaux cachés. La société CROCUS Technologie a déjà proposé une fonction sécurisée innovante appelée « match-in-place » (MIP) [27] basé sur la technologie MLU MRAM qui est elle-même dérivée de la TAS-MRAM. Une évaluation de cette technologie contre les attaques par canaux cachés a été effectuée dans [152]. Une extension de cette étude sur les autres technologies notamment la STT et SOT serait intéressante.

Publications et communications

Publications

Arash Nejat, Frederic Ouattara, Mohammad Mohammadinodoushan, Bertrand Cambou, Ken Mackay and Lionel Torres **"Practical Experiments to Evaluate Quality Metrics of MRAM-based Physical Unclonable Functions**" in IEEE Access, vol. 8, pp. 176042-176049, 2020, doi: 10.1109/ACCESS.2020.3024598.

Frederic Ouattara, Arash Nejat, Lionel Torres and Ken Mackay, "**Practical Experiments on Fabricated TAS-MRAM Dies to Evaluate the Stochastic Behavior of Voltage-Controlled TRNGs**," in IEEE Access, vol. 7, pp. 59271-59277, 2019, doi: 10.1109/ACCESS.2019.2907186.

Sophiane Senni, Frederic Ouattara, Jad Modad, Kaan Sevin, Guillaume Patrigeon, Pascal Benoit, Pascal Nouet, Lionel Torres, François Duhem, Gregory Di Pendina, Guillaume Prenat, **"From Spintronic Devices to Hybrid CMOS/Magnetic System On Chip,"** 2018 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC), Verona, Italy, 2018, pp. 188-191, doi: 10.1109/VLSI-SoC.2018.8644875.

M. Tahoori ; S. M. Nair ; R. Bishnoi ; S. Senni ; J. Mohdad ; F. Mailly ; L. Torres ; P. Benoit ; A. Gamatie ; P. Nouet ; F. Ouattara ; G. Sassatelli ; K. Jabeur ; P. Vanhauwaert ; A. Atitoaie ; I. Firastrau ; G. Di Pendina ; G. Prenat, "**Using multifunctional standardized stack as universal spintronic technology for IoT**," 2018 Design, Automation & Test in Europe Conference & Exhibition (DATE), Dresden, 2018, pp. 931-936, doi: 10.23919/DATE.2018.8342143.

Posters

Frederic Ouattara, Lionel Torres, Ken Mackay. "**Thermally Assisted Switching (TAS) MRAM based TRNG**". Colloque du GDR SoC-SiP, Jun 2018, Paris, France. 12ème Colloque National du GDRSoC-SiP, 2018. lirmm-02079681

Frederic Ouattara, Lionel Torres. "**Magnetic Memory based secure devices**". Colloque du GDR SoC-SiP,Jun 2017, Bordeaux, France. 11ème Colloque National du GDR SoC-SiP, 2017. lirmm-02079609

Références

[1] “MRAM Technology and Business 2019 - i-Micronews.” .

[2] S. A. Wolf *et al.*, “Spintronics: a spin-based electronics vision for the future.,” *Science*, vol. 294, no. 5546, pp. 1488–95, Nov. 2001, doi: 10.1126/science.1065389.

[3] “(PDF) Spintronics-A Vision for Future in Electronics and Computers.” .

[4] Dieny, Speriosu, Parkin, Gurney, Wilhoit, and Mauri, “Giant magnetoresistive in soft ferromagnetic multilayers.,” *Phys. Rev. B. Condens. Matter*, vol. 43, no. 1, pp. 1297–1300, Jan. 1991, doi: 10.1103/physrevb.43.1297.

[5] “International Physics Masterclasses.” .

[6] G. Di Pendina, “Conception innovante et développement d’outils de conception d’ASIC pour Technologie Hybride CMOS / Magnétique,” Oct. 2012.

[7] “The 2007 Nobel Prize in Physics - Popular information.” .

[8] “The electrical conductivity of transition metals,” *Proc. R. Soc. London. Ser. A - Math. Phys. Sci.*, vol. 153, no. 880, pp. 699–717, Feb. 1936, doi: 10.1098/rspa.1936.0031.

[9] A. Fert and I. A. Campbell, “Two-Current Conduction in Nickel,” *Phys. Rev. Lett.*, vol. 21, no. 16, pp. 1190–1192, Oct. 1968, doi: 10.1103/PhysRevLett.21.1190.

[10] B. Loegel and F. Gautier, “Origine de la resistivite dans le cobalt et ses alliages dilues,” *J. Phys. Chem. Solids*, vol. 32, no. 12, pp. 2723–2735, Jan. 1971, doi: 10.1016/S0022-3697(71)80364-5.

[11] M. N. Baibich *et al.*, “Giant Magnetoresistance of (001)Fe/(001)Cr Magnetic Superlattices,” *Phys. Rev. Lett.*, vol. 61, no. 21, pp. 2472–2475, Nov. 1988, doi: 10.1103/PhysRevLett.61.2472.

[12] G. Binasch, P. Grünberg, F. Saurenbach, and W. Zinn, “Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange,” *Phys. Rev. B*, vol. 39, no. 7, pp. 4828–4830, Mar. 1989, doi: 10.1103/PhysRevB.39.4828.

[13] I. Giaever, “Electron tunneling and superconductivity.,” *Science*, vol. 183, no. 4131, pp. 1253–8, Mar. 1974, doi: 10.1126/science.183.4131.1253.

[14] I. Giaever, “Energy Gap in Superconductors Measured by Electron Tunneling,” *Phys. Rev. Lett.*, vol. 5, no. 4, pp. 147–148, Aug. 1960, doi: 10.1103/PhysRevLett.5.147.

[15] “The Nobel Prize in Physics 1973.” .

[16] P. M. Tedrow and R. Meservey, “Spin-Dependent Tunneling into Ferromagnetic Nickel,” *Phys. Rev. Lett.*, vol. 26, no. 4, pp. 192–195, Jan. 1971, doi: 10.1103/PhysRevLett.26.192.

[17] M. Julliere, “Tunneling between ferromagnetic films,” *Phys. Lett. A*, vol. 54, no. 3, pp. 225–226, Sep. 1975, doi: 10.1016/0375-9601(75)90174-7.

[18] 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, no. 16, pp. 3273–3276, Apr. 1995, doi: 10.1103/PhysRevLett.74.3273.

[19] D. Wang *et al.*, "70% TMR at Room Temperature for SDT Sandwich Junctions With CoFeB as Free and Reference Layers," *IEEE Trans. Magn.*, vol. 40, no. 4, pp. 2269–2271, Jul. 2004, doi: 10.1109/TMAG.2004.830219.

[20] S. S. P. Parkin *et al.*, "Giant tunnelling magnetoresistance at room temperature with MgO (100) tunnel barriers," *Nat. Mater.*, vol. 3, no. 12, pp. 862–867, Dec. 2004, doi: 10.1038/nmat1256.

[21] S. Ikeda *et al.*, "Tunnel magnetoresistance of 604% at 300K by suppression of Ta diffusion in CoFeB/MgO/CoFeB pseudo-spin-valves annealed at high temperature," *Appl. Phys. Lett.*, vol. 93, no. 8, p. 082508, Aug. 2008, doi: 10.1063/1.2976435.

[22] M. Niset and P. Kuhn, "EB618/D: Typcial Data Retention for Nonvolatile Memory."

[23] J. M. Slaughter *et al.*, "High speed toggle MRAM with mgO-based tunnel junctions," in *IEEE InternationalElectron Devices Meeting, 2005. IEDM Technical Digest.*, pp. 873–876, doi: 10.1109/IEDM.2005.1609496.

[24] "Method of writing to scalable magnetoresistance random access memory element," Oct. 2001.

[25] "Self-referenced magnetic random access memory cell," Jul. 2010.

[26] 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, no. 7, p. 074002, Feb. 2013, doi: 10.1088/0022-3727/46/7/074002.

[27] "Match-In-Place™. A Novel Way to Perform - Crocus Technology." .

[28] J. C. Slonczewski, "Current-driven excitation of magnetic multilayers," *J. Magn. Magn. Mater.*, vol. 159, no. 1–2, pp. L1–L7, Jun. 1996, doi: 10.1016/0304-8853(96)00062-5.

[29] L. Berger, "Emission of spin waves by a magnetic multilayer traversed by a current," *Phys. Rev. B*, vol. 54, no. 13, pp. 9353–9358, Oct. 1996, doi: 10.1103/PhysRevB.54.9353.

[30] S. Lee, K. Kang, and C. M. Kyung, "Runtime thermal management for 3-D chip-multiprocessors with hybrid SRAM/MRAM L2 cache," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 23, no. 3, pp. 520–533, Mar. 2015, doi: 10.1109/TVLSI.2014.2311798.

[31] S. Sakhare *et al.*, "Enablement of STT-MRAM as last level cache for the high performance computing domain at the 5nm node," in *Technical Digest - International Electron Devices Meeting, IEDM*, Jan. 2019, vol. 2018-December, pp. 18.3.1-18.3.4, doi: 10.1109/IEDM.2018.8614637.

[32] O. Coi, G. Patrigeon, S. Senni, L. Torres, and P. Benoit, "A novel SRAM-STT-MRAM hybrid cache implementation improving cache performance," in *Proceedings of the IEEE/ACM International Symposium on Nanoscale Architectures, NANOARCH 2017*, Sep. 2017, pp. 39–44, doi: 10.1109/NANOARCH.2017.8053704.

[33] N. N. Mojumder, S. K. Gupta, S. H. Choday, D. E. Nikonov, and K. Roy, "A Three-Terminal Dual-Pillar STT-MRAM for High-Performance Robust Memory Applications," *IEEE Trans. Electron Devices*, vol. 58, no. 5, pp. 1508–1516, May 2011, doi:

10.1109/TED.2011.2116024.

[34] S. H. Kang and K. Lee, “Emerging materials and devices in spintronic integrated circuits for energy-smart mobile computing and connectivity,” *Acta Mater.*, vol. 61, no. 3, pp. 952–973, Feb. 2013, doi: 10.1016/J.ACTAMAT.2012.10.036.

[35] X. Fong and K. Roy, “Complimentary polarizers STT-MRAM (CPSTT) for on-chip caches,” *IEEE Electron Device Lett.*, vol. 34, no. 2, pp. 232–234, 2013, doi: 10.1109/LED.2012.2234079.

[36] “(13) Sttram scaling and retention failure | Request PDF.” .

[37] “Voltage-controlled MRAM: Status, challenges and prospects | EE Times.” .

[38] J. G. Alzate *et al.*, “Voltage-induced switching of nanoscale magnetic tunnel junctions,” in *2012 International Electron Devices Meeting*, Dec. 2012, pp. 29.5.1-29.5.4, doi: 10.1109/IEDM.2012.6479130.

[39] Y. Shiota *et al.*, “Pulse voltage-induced dynamic magnetization switching in magnetic tunneling junctions with high resistance-area product,” *Appl. Phys. Lett.*, vol. 101, no. 10, p. 102406, Sep. 2012, doi: 10.1063/1.4751035.

[40] 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, no. 1, pp. 39–43, Jan. 2012, doi: 10.1038/nmat3172.

[41] P. K. Amiri *et al.*, “Electric-Field-Controlled Magnetoelectric RAM: Progress, Challenges, and Scaling,” *IEEE Trans. Magn.*, vol. 51, no. 11, Nov. 2015, doi: 10.1109/TMAG.2015.2443124.

[42] S. Senni, L. Torres, G. Sassatelli, A. Bukto, and B. Mussard, “Exploration of magnetic RAM based memory hierarchy for multicore architecture,” in *Proceedings of IEEE Computer Society Annual Symposium on VLSI, ISVLSI*, Sep. 2014, pp. 248–251, doi: 10.1109/ISVLSI.2014.29.

[43] A. Chernyshov, M. Overby, X. Liu, J. K. Furdyna, Y. Lyanda-Geller, and L. P. Rokhinson, “Evidence for reversible control of magnetization in a ferromagnetic material by means of spin-orbit magnetic field,” *Nat. Phys.*, vol. 5, no. 9, pp. 656–659, Aug. 2009, doi: 10.1038/nphys1362.

[44] I. Mihai Miron *et al.*, “Current-driven spin torque induced by the Rashba effect in a ferromagnetic metal layer,” *Nat. Mater.*, vol. 9, no. 3, pp. 230–234, Mar. 2010, doi: 10.1038/nmat2613.

[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 (80-.).*, vol. 336, no. 6081, pp. 555–558, May 2012, doi: 10.1126/science.1218197.

[46] F. Oboril, R. Bishnoi, M. Ebrahimi, and M. B. Tahoori, “Evaluation of hybrid memory technologies using SOT-MRAM for on-chip cache hierarchy,” *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 34, no. 3, pp. 367–380, Mar. 2015, doi: 10.1109/TCAD.2015.2391254.

[47] T. W. Andre *et al.*, “A 4-Mb 0.18-/spl mu/m 1T1MTJ toggle MRAM with balanced three

input sensing scheme and locally mirrored unidirectional write drivers," *IEEE J. Solid-State Circuits*, vol. 40, no. 1, pp. 301–309, Jan. 2005, doi: 10.1109/JSSC.2004.837962.

[48] T. Sugibayashi *et al.*, "A 16-Mb toggle MRAM with burst modes," in *IEEE Journal of Solid-State Circuits*, Nov. 2007, vol. 42, no. 11, pp. 2378–2385, doi: 10.1109/JSSC.2007.906195.

[49] "Everspin | The MRAM Company." .

[50] "(PDF) Thermally assisted MRAM." .

[51] "TMR Technology | Magnetic Logic Unit (MLU) | Crocus Technology." .

[52] "Basic principles of STT-MRAM cell operation in memory arrays - IOPscience." .

[53] R. Dorrance *et al.*, "Diode-MTJ crossbar memory cell using voltage-induced unipolar switching for high-density MRAM," *IEEE Electron Device Lett.*, vol. 34, no. 6, pp. 753–755, 2013, doi: 10.1109/LED.2013.2255096.

[54] "Low-power non-volatile spintronic memory: STT-RAM and beyond - IOPscience." .

[55] R. Bishnoi, M. Ebrahimi, F. Oboril, and M. B. Tahoori, "Architectural aspects in design and analysis of SOT-based memories," in *Proceedings of the Asia and South Pacific Design Automation Conference, ASP-DAC*, 2014, pp. 700–707, doi: 10.1109/ASPDAC.2014.6742972.

[56] "'Study of Two Writing Schemes for a Magnetic Tunnel Junction Based On Spin Orbit Torque | Request PDF." .

[57] R. Alhalabi, E. Nowak, I. L. Prejbeanu, and G. Di Pendina, "High density SOT-MRAM memory array based on a single transistor," Jan. 2019, doi: 10.1109/NVMTS.2018.8603114.

[58] W. Killmann and W. Schindler, "A proposal for: Functionality classes for random number generators A proposal for: Functionality classes for random number generators 1," 2011.

[59] K. J. Kuhn *et al.*, "Process technology variation," *IEEE Trans. Electron Devices*, vol. 58, no. 8, pp. 2197–2208, Aug. 2011, doi: 10.1109/TED.2011.2121913.

[60] A. Jaiswal, X. Fong, and K. Roy, "Comprehensive Scaling Analysis of Current Induced Switching in Magnetic Memories Based on In-Plane and Perpendicular Anisotropies," *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 6, no. 2, pp. 120–133, Jun. 2016, doi: 10.1109/JETCAS.2016.2547698.

[61] S. W. Golomb, "Shift register sequences - A retrospective account," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2006, vol. 4086 LNCS, pp. 1–4, doi: 10.1007/11863854_1.

[62] R. A. Rueppel, "Stream Ciphers," in *Analysis and Design of Stream Ciphers*, Berlin, Heidelberg: Springer Berlin Heidelberg, 1986, pp. 5–16.

[63] J. L. Massey, "Shift-Register Synthesis and BCH Decoding," *IEEE Trans. Inf. Theory*, vol. 15, no. 1, pp. 122–127, 1969, doi: 10.1109/TIT.1969.1054260.

[64] D. Coppersmith, H. Krawczyk, and Y. Mansour, “The shrinking generator,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1994, vol. 773 LNCS, pp. 22–39, doi: 10.1007/3-540-48329-2_3.

[65] P. Caballero-Gil, A. Fúster-Sabater, and M. E. Pazo-Robles, “New Attack Strategy for the Shrinking Generator,” May 2010.

[66] W. Meier and O. Staffelbach, “The self-shrinking generator,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1995, vol. 950, pp. 205–214, doi: 10.1007/bfb0053436.

[67] E. Zenner, M. Krause, and S. Lucks, “Improved cryptanalysis of the self-shrinking generator,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2001, vol. 2119, pp. 21–35, doi: 10.1007/3-540-47719-5_4.

[68] “Intel Random Number Generator - Rambus.” .

[69] S. hua Zhou, W. Zhang, and N. J. Wu, “An ultra-low power CMOS random number generator,” *Solid. State. Electron.*, vol. 52, no. 2, pp. 233–238, Feb. 2008, doi: 10.1016/j.sse.2007.08.008.

[70] V. Bagini and M. Bucci, “A design of reliable true random number generator for cryptographic applications,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1999, vol. 1717, pp. 204–218, doi: 10.1007/3-540-48059-5_18.

[71] W. Killmann and W. Schindler, “A design for a physical RNG with robust entropy estimators,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2008, vol. 5154 LNCS, pp. 146–163, doi: 10.1007/978-3-540-85053-3_10.

[72] R. C. Fairfield, R. L. Mortenson, and K. B. Coulthart, “An LSI Random Number Generator (RNG),” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1985, vol. 196 LNCS, pp. 203–230, doi: 10.1007/3-540-39568-7_18.

[73] V. Fischer, F. Bernard, N. Bochard, and M. Varchola, “Enhancing security of ring oscillator-based TRNG implemented in FPGA,” in *Proceedings - 2008 International Conference on Field Programmable Logic and Applications, FPL*, 2008, pp. 245–250, doi: 10.1109/FPL.2008.4629939.

[74] T. J. Chaney and C. E. Molnar, “Anomalous Behavior of Synchronizer and Arbiter Circuits,” *IEEE Trans. Comput.*, vol. C–22, no. 4, pp. 421–422, 1973, doi: 10.1109/T-C.1973.223730.

[75] I. Vasyltsov, E. Hambardzumyan, Y. S. Kim, and B. Karpinskyy, “Fast digital TRNG based on metastable ring oscillator,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2008, vol. 5154 LNCS, pp. 164–180, doi: 10.1007/978-3-540-85053-3_11.

[76] M. Epstein, L. Hars, R. Krasinski, M. Rosner, and H. Zheng, “Design and implementation of a true random number generator based on digital circuit artifacts,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2779, pp. 152–165, 2003, doi: 10.1007/978-3-540-45238-6_13.

[77] “Design of an on-chip random number generator using metastability - IEEE Conference Publication.” .

[78] C. Tokunaga, D. Blaauw, and T. Mudge, “True Random Number Generator With a Metastability-Based Quality Control,” *IEEE J. Solid-State Circuits*, vol. 43, no. 1, pp. 78–85, 2008, doi: 10.1109/JSSC.2007.910965.

[79] V. B. Suresh and W. P. Burleson, “Entropy extraction in metastability-based TRNG,” in *Proceedings of the 2010 IEEE International Symposium on Hardware-Oriented Security and Trust, HOST 2010*, 2010, pp. 135–140, doi: 10.1109/HST.2010.5513099.

[80] M. S. Turan *et al.*, “Recommendation for the Entropy Sources Used for Random Bit Generation NIST Special Publication 800-90B Recommendation for the Entropy Sources Used for Random Bit Generation,” *NIST Spec. Publ. 800-90B*, pp. 800–90, 2018, doi: 10.6028/NIST.SP.800-90B.

[81] J. Von Neumann, *Various techniques used in connection with random digits*, vol. 12. 1951.

[82] D. E. 3rd, S. Crocker, and J. Schiller, “Randomness Recommendations for Security.”

[83] M. Dichtl, “Bad and good ways of post-processing biased physical random numbers,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2007, vol. 4593 LNCS, pp. 137–152, doi: 10.1007/978-3-540-74619-5_9.

[84] “Intel Ivy Bridge Random Number Generator - Rambus.” .

[85] “Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment - Peter H. Westfall, S. Stanley Young - Google Livres.” .

[86] A. Rukhin and all, *A statistical Test Suite For Cryptographic Applications*. 2010.

[87] G. Marsaglia, “The marsaglia random number cd-rom with the diehard battery of tests of randomness.” 1995.

[88] “Robert G. Brown’s General Tools Page.” .

[89] T. Devolder *et al.*, “Single-shot time-resolved measurements of nanosecond-scale spin-transfer induced switching: Stochastic versus deterministic aspects,” *Phys. Rev. Lett.*, vol. 100, no. 5, p. 057206, Feb. 2008, doi: 10.1103/PhysRevLett.100.057206.

[90] J. Z. Sun, “Spin-current interaction with a monodomain magnetic body: A model study,” *Phys. Rev. B - Condens. Matter Mater. Phys.*, vol. 62, no. 1, pp. 570–578, Jul. 2000, doi: 10.1103/physrevb.62.570.

[91] Y. Wang, “Reliability analysis of spintronic device based logic and memory circuits,” Télécom ParisTech, Feb. 2017.

[92] W. S. Zhao *et al.*, “Failure and reliability analysis of STT-MRAM,” *Microelectron. Reliab.*,

vol. 52, no. 9–10, pp. 1848–1852, Sep. 2012, doi: 10.1016/j.microrel.2012.06.035.

[93] A. Fukushima *et al.*, "Spin dice: A scalable truly random number generator based on spintronics," *Appl. Phys. Express*, vol. 7, no. 8, p. 83001, 2014.

[94] Y. Suzuki, A. A. Tulapurkar, and C. Chappert, "Spin-Injection Phenomena and Applications," in *Nanomagnetism and Spintronics*, Elsevier, 2009, pp. 93–153.

[95] H. Zhao *et al.*, "Low writing energy and sub nanosecond spin torque transfer switching of in-plane magnetic tunnel junction for spin torque transfer random access memory," in *Journal of Applied Physics*, Apr. 2011, vol. 109, no. 7, p. 07C720, doi: 10.1063/1.3556784.

[96] Y. Acremann, C. H. Back, M. Buess, D. Pescia, and V. Pokrovsky, "Bifurcation in precessional switching," *Appl. Phys. Lett.*, vol. 79, no. 14, pp. 2228–2230, Oct. 2001, doi: 10.1063/1.1407299.

[97] W. H. Choi *et al.*, "A Magnetic Tunnel Junction based True Random Number Generator with conditional perturb and real-time output probability tracking," in *2014 IEEE International Electron Devices Meeting*, 2014, pp. 12.5.1-12.5.4, doi: 10.1109/IEDM.2014.7047039.

[98] S. Oosawa, T. Konishi, N. Onizawa, and T. Hanyu, "Design of an STT-MTJ based true random number generator using digitally controlled probability-locked loop," Aug. 2015, doi: 10.1109/NEWCAS.2015.7182089.

[99] Y. Wang, H. Cai, L. A. B. Naviner, J. O. Klein, J. Yang, and W. Zhao, "A novel circuit design of true random number generator using magnetic tunnel junction," in *2016 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH)*, 2016, pp. 123–128, doi: 10.1145/2950067.2950108.

[100] E. I. Vatajelu, G. Di Natale, and P. Prinetto, "STT-MTJ-based TRNG with on-the-fly temperature/current variation compensation," in *2016 IEEE 22nd International Symposium on On-Line Testing and Robust System Design, IOLTS 2016*, Oct. 2016, pp. 179–184, doi: 10.1109/IOLTS.2016.7604694.

[101] M. Maymandi-Nejad and M. Sachdev, "A Digitally Programmable Delay Element: Design and Analysis," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 11, no. 5, pp. 871–878, Oct. 2003, doi: 10.1109/TVLSI.2003.810787.

[102] E. I. Vatajelu and G. Di Natale, "High-entropy stt-MTJ-based TRNG," *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 27, no. 2, pp. 491–495, Feb. 2019, doi: 10.1109/TVLSI.2018.2879439.

[103] Y. Qu, J. Han, B. F. Cockburn, W. Pedrycz, Y. Zhang, and W. Zhao, "A true random number generator based on parallel STT-MTJs," in *Proceedings of the 2017 Design, Automation and Test in Europe, DATE 2017*, May 2017, pp. 606–609, doi: 10.23919/DATE.2017.7927058.

[104] Y. Qu *et al.*, "Variation-Resilient True Random Number Generators Based on Multiple STT-MTJs," *IEEE Trans. Nanotechnol.*, vol. 17, no. 6, pp. 1270–1281, Nov. 2018, doi: 10.1109/TNANO.2018.2873970.

[105] K. Yang *et al.*, “A 28NM Integrated True Random Number Generator Harvesting Entropy from MRAM,” in *IEEE Symposium on VLSI Circuits, Digest of Technical Papers*, Oct. 2018, vol. 2018-June, pp. 171–172, doi: 10.1109/VLSIC.2018.8502431.

[106] R. Carboni *et al.*, “Random number generation by differential read of stochastic switching in spin-transfer torque memory,” *IEEE Electron Device Lett.*, vol. 39, no. 7, pp. 951–954, Jul. 2018, doi: 10.1109/LED.2018.2833543.

[107] Y. Kim, X. Fong, and K. Roy, “Spin-Orbit-Torque-Based Spin-Dice: A True Random-Number Generator,” *IEEE Magn. Lett.*, vol. 6, 2015, doi: 10.1109/LMAG.2015.2496548.

[108] Y. Liu, Z. Wang, Z. Li, X. Wang, and W. Zhao, “A spin orbit torque based true random number generator with real-time optimization,” in *Proceedings of the IEEE Conference on Nanotechnology*, Jan. 2019, vol. 2018-July, doi: 10.1109/NANO.2018.8626347.

[109] “Spintec - Spin in electronics Research – Grenoble – France.” .

[110] W. C. Black and B. Das, “Programmable logic using giant-magnetoresistance and spin-dependent tunneling devices (invited),” *J. Appl. Phys.*, vol. 87, no. 9, pp. 6674–6679, May 2000, doi: 10.1063/1.372806.

[111] K. J. Hass, K. J. Hass, and G. W. Donohoe, “SEU-Resistant Magnetic Flip Flops.”

[112] W. Zhao, C. Chappert, V. Javerliac, and J. P. Noziére, “High speed, high stability and low power sensing amplifier for MTJ/CMOS hybrid logic circuits,” in *IEEE Transactions on Magnetics*, Oct. 2009, vol. 45, no. 10, pp. 3784–3787, doi: 10.1109/TMAG.2009.2024325.

[113] Y. Guillemenet, L. Torres, G. Sassatelli, and N. Bruchon, “On the Use of Magnetic RAMs in Field Programmable Gate Arrays,” *Int. J. Reconfigurable Comput.*, vol. Q3, p. ID 723950, 2008, doi: 10.1155/2008/723950.

[114] R. M. Brum, “Conception hybride CMOS et mémoires magnétiques : applications aux architectures programmables,” *http://www.theses.fr*, Dec. 2014.

[115] F. Ouattara, A. Nejat, L. Torres, and K. Mackay, “Practical Experiments on Fabricated TAS-MRAM Dies to Evaluate the Stochastic Behavior of Voltage-Controlled TRNGs,” *IEEE Access*, vol. 7, pp. 59271–59277, 2019, doi: 10.1109/ACCESS.2019.2907186.

[116] K. Lofstrom, W. R. Daasch, and D. Taylor, “IC identification circuit using device mismatch,” *Dig. Tech. Pap. - IEEE Int. Solid-State Circuits Conf.*, pp. 372–373, 2000, doi: 10.1109/ISSCC.2000.839821.

[117] “Silicon physical random functions | Proceedings of the 9th ACM conference on Computer and communications security.” .

[118] R. Maes and R. Maes, “Physically Unclonable Functions: Concept and Constructions,” in *Physically Unclonable Functions*, Springer Berlin Heidelberg, 2013, pp. 11–48.

[119] H. Handschuh, G. J. Schrijen, and P. Tuyls, “Hardware intrinsic security from physically unclonable functions,” in *Information Security and Cryptography*, no. 9783642143120, Springer International Publishing, 2010, pp. 39–53.

[120] M. D. Yu and S. Devadas, “Secure and robust error correction for physical unclonable

functions," *IEEE Des. Test Comput.*, vol. 27, no. 1, pp. 48–65, Jan. 2010, doi: 10.1109/MDT.2010.25.

[121] T. McGrath, I. E. Bagci, Z. M. Wang, U. Roedig, and R. J. Young, "A PUF taxonomy," *Applied Physics Reviews*, vol. 6, no. 1. American Institute of Physics Inc., p. 011303, Mar. 2019, doi: 10.1063/1.5079407.

[122] T. Fournel and M. Hébert, "Towards weak optical PUFs by random spectral mixing," Nov. 2016, doi: 10.1109/WIO.2016.7745572.

[123] S. Dolev, Ł. Krzywiecki, N. Panwar, and M. Segal, "Optical PUF for non forwardable vehicle authentication," in *Proceedings - 2015 IEEE 14th International Symposium on Network Computing and Applications, NCA 2015*, Jan. 2016, pp. 204–207, doi: 10.1109/NCA.2015.25.

[124] D. R. Reising, M. A. Temple, and J. A. Jackson, "Authorized and Rogue Device Discrimination Using Dimensionally Reduced RF-DNA Fingerprints," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 6, pp. 1180–1192, Jun. 2015, doi: 10.1109/TIFS.2015.2400426.

[125] M. W. Lukacs, A. J. Zeqolari, P. J. Collins, and M. A. Temple, "'RF-DNA' Fingerprinting for Antenna Classification," *IEEE Antennas Wirel. Propag. Lett.*, vol. 14, pp. 1455–1458, Mar. 2015, doi: 10.1109/lawp.2015.2411608.

[126] J. D. R. Buchanan *et al.*, "Forgery: 'Fingerprinting' documents and packaging," *Nature*, vol. 436, no. 7050, p. 475, Jul. 2005, doi: 10.1038/436475a.

[127] P. Bulens, F. X. Standaert, and J. J. Quisquater, "How to strongly link data and its medium: The paper case," *IET Inf. Secur.*, vol. 4, no. 3, pp. 125–136, Sep. 2010, doi: 10.1049/iet-ifs.2009.0032.

[128] P. Tuyls, G. J. Schrijen, B. Škorić, J. Van Geloven, N. Verhaegh, and R. Wolters, "Read-proof hardware from protective coatings," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Oct. 2006, vol. 4249 LNCS, pp. 369–383, doi: 10.1007/11894063_29.

[129] B. Škorić, S. Maubach, T. Kevenaar, and P. Tuyls, "Information-theoretic analysis of capacitive physical unclonable functions," *J. Appl. Phys.*, vol. 100, no. 2, pp. 024902–1/11, 2006, doi: 10.1063/1.2209532.

[130] Y. Lao and K. K. Parhi, "Reconfigurable architectures for silicon Physical Unclonable Functions," 2011, doi: 10.1109/EIT.2011.5978614.

[131] J. W. Lee, D. Lim, B. Gassend, G. E. Suh, M. Van Dijk, and S. Devadas, "A technique to build a secret key in integrated circuits for identification and authentication applications," in *IEEE Symposium on VLSI Circuits, Digest of Technical Papers*, 2004, no. CIRCUITS SYMP., pp. 176–179, doi: 10.1109/vlsic.2004.1346548.

[132] U. Ruhrmair and J. Solter, "PUF modeling attacks: An introduction and overview," Apr. 2014, pp. 1–6, doi: 10.7873/date.2014.361.

[133] "Physical Unclonable Functions for Device Authentication and Secret Key Generation - IEEE Conference Publication." .

[134] J. Guajardo, S. S. Kumar, G. J. Schrijen, and P. Tuyls, “FPGA intrinsic PUFs and their use for IP protection,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2007, vol. 4727 LNCS, pp. 63–80, doi: 10.1007/978-3-540-74735-2_5.

[135] S. S. Kumar, J. Guajardo, R. Maes, G. J. Schrijen, and P. Tuyls, “The Butterfly PUF protecting IP on every FPGA,” in *2008 IEEE International Workshop on Hardware-Oriented Security and Trust, HOST*, 2008, pp. 67–70, doi: 10.1109/HST.2008.4559053.

[136] U. Rührmair, H. Busch, and S. Katzenbeisser, “Strong PUFs: Models, constructions, and security proofs,” in *Information Security and Cryptography*, no. 9783642143120, Springer International Publishing, 2010, pp. 79–96.

[137] D. Lim, J. W. Lee, B. Gassend, G. E. Suh, M. Van Dijk, and S. Devadas, “Extracting secret keys from integrated circuits,” *IEEE Trans. Very Large Scale Integr. Syst.*, vol. 13, no. 10, pp. 1200–1205, Oct. 2005, doi: 10.1109/TVLSI.2005.859470.

[138] M. Majzoobi, F. Koushanfar, and M. Potkonjak, “Testing techniques for hardware security,” 2008, doi: 10.1109/TEST.2008.4700636.

[139] A. Maiti, J. Casarona, L. McHale, and P. Schaumont, “A large scale characterization of RO-PUF,” in *Proceedings of the 2010 IEEE International Symposium on Hardware-Oriented Security and Trust, HOST 2010*, 2010, pp. 94–99, doi: 10.1109/HST.2010.5513108.

[140] Y. Hori, T. Yoshida, T. Katashita, and A. Satoh, “Quantitative and statistical performance evaluation of arbiter Physical Unclonable Functions on FPGAs,” in *Proceedings - 2010 International Conference on Reconfigurable Computing and FPGAs, ReConFig 2010*, 2010, pp. 298–303, doi: 10.1109/ReConFig.2010.24.

[141] J. Das, K. Scott, D. Burgett, S. Rajaram, and S. Bhanja, “A novel geometry based MRAM PUF,” in *Proceedings of the IEEE Conference on Nanotechnology*, Nov. 2014, pp. 859–863, doi: 10.1109/NANO.2014.6968027.

[142] T. Marukame, T. Tanamoto, and Y. Mitani, “Extracting physically unclonable function from spin transfer switching characteristics in magnetic tunnel junctions,” *IEEE Trans. Magn.*, vol. 50, no. 11, Nov. 2014, doi: 10.1109/TMAG.2014.2325646.

[143] A. Kumar, S. Sahay, and M. Suri, “Switching-Time Dependent PUF Using STT-MRAM,” in *Proceedings of the IEEE International Conference on VLSI Design*, Mar. 2018, vol. 2018-January, pp. 434–438, doi: 10.1109/VLSID.2018.103.

[144] L. Zhang, X. Fong, C. H. Chang, Z. H. Kong, and K. Roy, “Highly reliable memory-based Physical Unclonable Function using Spin-Transfer Torque MRAM,” in *Proceedings - IEEE International Symposium on Circuits and Systems*, 2014, pp. 2169–2172, doi: 10.1109/ISCAS.2014.6865598.

[145] E. I. Vatajelu, G. Di Natale, L. Torres, and P. Prinetto, “STT-MRAM-based strong PUF architecture,” in *Proceedings of IEEE Computer Society Annual Symposium on VLSI, ISVLSI*, Oct. 2015, vol. 07-10-July-2015, pp. 467–472, doi: 10.1109/ISVLSI.2015.128.

[146] E. I. Vatajelu, G. Di Natale, and P. Prinetto, “Zero bit-error-rate weak PUF based on Spin-

Transfer-Torque MRAM memories," in *2017 2nd International Verification and Security Workshop, IVSW 2017*, Sep. 2017, pp. 128–133, doi: 10.1109/IVSW.2017.8031552.

[147] S. Khaleghi, P. Vinella, S. Banerjee, and W. Rao, "An STT-MRAM based strong PUF," in *Proceedings of the 2016 IEEE/ACM International Symposium on Nanoscale Architectures, NANOARCH 2016*, Sep. 2016, pp. 129–134, doi: 10.1145/2950067.2950080.

[148] L. Zhang, X. Fong, C. H. Chang, Z. H. Kong, and K. Roy, "Optimizating Emerging Nonvolatile Memories for Dual-Mode Applications: Data Storage and Key Generator," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 34, no. 7, pp. 1176–1187, Jul. 2015, doi: 10.1109/TCAD.2015.2427251.

[149] E. I. Vatajelu, G. Di Natale, M. Indaco, and P. Prinetto, "STT MRAM-based PUFs," in *Proceedings -Design, Automation and Test in Europe, DATE*, Apr. 2015, vol. 2015-April, pp. 872–875, doi: 10.7873/date.2015.0505.

[150] Y. S. Chen *et al.*, "On the Hardware Implementation of MRAM Physically Unclonable Function," *IEEE Trans. Electron Devices*, vol. 64, no. 11, pp. 4492–4495, Nov. 2017, doi: 10.1109/TED.2017.2755867.

[151] A. Nejat, F. Ouattara, M. Mohammadinodoushan, B. Cambou, K. Mackay, and L. Torres, "Practical experiments to evaluate quality metrics of MRAM-Based physical unclonable functions," *Am. J. Int. Law*, vol. 8, pp. 176042–176049, 2020, doi: 10.1109/ACCESS.2020.3024598.

[152] J. Clement, B. Mussard, D. Naccache, and L. Torres, "Implementation of AES using NVM memories based on comparison function," in *Proceedings of IEEE Computer Society Annual Symposium on VLSI, ISVLSI*, Oct. 2015, vol. 07-10-July-2015, pp. 356–361, doi: 10.1109/ISVLSI.2015.37.

Les mémoires magnétiques (MRAM) font partie des technologies de mémoires non volatiles émergentes ayant connu un développement rapide cette dernière décennie. Un des avantages de cette technologie réside dans les domaines d'applications variées dans lesquelles elle peut intervenir. En plus de sa fonction principale de stockage d'information, la MRAM est utilisée de nos jours dans des applications de type capteurs, récepteur RF et sécurité matérielle. Dans cette thèse, nous nous sommes intéressés à l'utilisation des MRAM dans la conception des primitives de sécurité matérielle élémentaires. Dans un premier temps, une exploration dans la conception de TRNG (True Random Number Generator) basée sur des mémoires de type STT-MRAM (Spin Transfert Torque MRAM) a été menée dans le but de réaliser un démonstrateur et de prouver son efficacité pour les applications sécurisées. Les méthodes d'extraction d'aléa dans les mémoires STT et TAS (Thermally Assisted Switching) sont présentées. Nous avons ainsi évalué ces mémoires magnétiques dans le cadre des TRNG mais également pour la génération de PUF (Physically Unclonable Functions) sur des dispositifs physiques.

Mots-clés : MRAM, TRNG, PUF, Sécurité matérielle

Magnetic memories (MRAM) are one of the emerging non-volatile memory technologies that have experienced rapid development over the past decade. One of the advantages of this technology lies in the varied fields of application in which it can be used. In addition to its primary function of storing information, MRAM is nowadays used in applications such as sensors, RF receivers and hardware security. In this thesis, we are interested in the use of MRAMs in the design of elementary hardware security primitives. Initially, an exploration in the design of TRNG (True Random Number Generator) based on STT-MRAM (Spin Transfert Torque MRAM) type memories was carried out with the aim of producing a demonstrator and proving its effectiveness for secure applications. Random extraction methods in STT and TAS (Thermally Assisted Switching) memories are presented. We have thus evaluated these magnetic memories within the framework of TRNGs but also for the generation of PUFs (Physically Unclonable Functions) on physical devices.

Keywords: MRAM, TRNG, PUF, Hardware security