Magnetic Tunnel Junctions for Next Generation of Conventional and Unconventional Computing Schemes

A DISSERTATION

SUBMITTED TO THE FACULTY OF

THE UNIVERSITY OF MINNESOTA

BY

Brandon Zink

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Advisor: Prof. Jian-Ping Wang

September 2022

Acknowledgements

My dissertation could not have been completed without the tremendous support from my mentors, colleagues, family, and friends. It is through their support that this work has been possible.

I am very honored to have Prof. Jian-Ping Wang as my research advisor and I would like to thank him for providing me the opportunity to conduct my research in a collaborative, productive, and supporting environment. I am grateful for his continuous guidance, support, and encouragement during my time at the University of Minnesota.

I am incredibly grateful to have Prof. Chris Kim, Prof. Steven Koester and Prof. Paul Crowel as my final examination committee members and am very thankful for their helpful feedback and suggestions. I would also like to thank Prof. Sachin Sapatnekar, Prof. Ulya Karpuzcu, Dr. Masoud Zabihi, Dr. Husrev Cilasun, and Prof. Marc Reidel for their support and discussion on my research topics.

The work presented in my dissertation are completed through extensive collaborations. I'm truly appreciative of all the collaborations and discussions from all my external collaborators, Dr. Jean Yang-Scharlotta and Dr. K. Michael Han from Jet Propulsion Laboratory, Dr. Frederick Mancoff from Everspin Technologies, and Prof. Weigang Wang from Arizona University. I am also very appreciative of all the support I've received from my colleagues within Prof. Wang's group. Their engaging discussion and collaboration provide additional support throughout my PhD work. In particular, I would like to thank Dr. Yang Lv, Dr. Delin Zhang, Dr. Will Echtenkamp, Dr. Bin Ma, Dr. Xiaohui Chao, Dr. Hongshi Li, Dr. Kai Wu, Dr. Thomas Peterson, Dr. Protyush Sahu, Deyuan Lyu,

Yihong Fan, Onri Jay Benally, and Denis Tonini. It is my great pleasure to work with these talented scholars.

Finally, my deepest gratitude goes to my family and wonderful friends, who gave me unconditional love and support.

Abstract

Magnetic tunnel junctions (MTJs) have several novel features that make them promising devices for next generation memory and computing technologies. These features include multifunctionality, tunable stochasticity, and capability of being tuned by multiple forces. In this dissertation, I demonstrate how these properties make MTJs promising solutions in conventional memory and logic applications as well as unconventional computing applications such as probabilistic bits with increased information capacity and stochastic computing units.

Devices that generate random asynchronous, or telegraphic, switching signals have been proposed as probabilistic bits (p-bits) in new paradigms of probabilistic computing schemes for advanced computation. For the first part of my dissertation, I demonstrate that tunable telegraphic switching signals can be generated from MTJs through the combination of an external magnetic field and a DC bias voltage. Previous studies show that tunable telegraphic signals can be generated on MTJs with low thermal stability using only a single bias current or bias voltage. However, my results show that this 'dual-biasing' method has a unique capability called two-degrees of tunability, which gives this method two key advantages over the single biased method. One is that it can overcome the challenges imposed by the effects of device variations in large-scale networks and the second is that the signals generated have two times more information capacity than those generated by single-biased MTJs.

In the 2nd part of my dissertation, I explore the interplay between the effects of the voltage-controlled exchange coupling (VCEC) and spin-orbit torque (SOT) switching mechanisms. Previous experimental work from our group has demonstrated that VCEC

switching can be achieved in perpendicularly magnetized MTJs (p-MTJs) at switching current densities nearly one order of magnitude lower than those for spin transfer torque (STT) and SOT switching. In this dissertation, I show that by combining the SOT and VCEC effects, the VCEC switching current density can be reduced even further, thus providing a pathway to optimize the performance of future magnetoresistive random access memory (MRAM) technologies based on VCEC switching.

In the 3rd part of my dissertation, I describe a method that was invented by our research group that performs stochastic computing within the hardware for computational random-access memory (CRAM), which is called SC-CRAM. Stochastic computing (SC) has several attractive capabilities such as performing complex arithmetic functions with a small number of logic gates, noise resilience, and error tolerance. However, there are significant costs in circuit area and energy consumption for the hardware required to generate stochastic bit-streams. The method described in my dissertation overcomes these costs by embedding the bit-stream generation and computation steps within the same CRAM cells. Furthermore, SC-CRAM shows significant reductions in circuit area for certain neuromorphic computing tasks when compared to conventional computing methods in CRAM.

Finally, I study the prospects of MTJs for future applications involving high radiation environments, such as space exploration. Ionizing radiation levels beyond 10 krad have detrimental effects on modern CMOS technology. However, my results demonstrate that MTJs can be exposed to ionizing radiation levels as high as 1 Mrad without significantly influencing the properties key to their performance in magnetic random access memory

(MRAM) cells. This resilience to ionizing radiation makes MTJs strong candidates for future ‘rad-hard’ devices.

Table of Contents

Acknowledgments i
Abstract iii
List of Tables ix
List of Figures x
Chapter 1 Introduction 1
1.1 Magnetic tunnel junctions (MTJs) 1
1.1.1 Tunneling Magnetoresistance 2
1.1.2 Anisotropy 3
1.2 Writing mechanisms 6
1.2.1 Spin transfer torque (STT) 8
1.2.2 Spin-orbit torque (SOT) 9
1.2.3 Voltage controlled exchange coupling (VCEC) 11
1.3 MTJs for non-conventional computing applications 13
1.3.1 Neuromorphic computing background 13
1.3.2 Computational Random Access memory (CRAM) 19
1.3.3 MTJs for stochastic computing 21
1.3.4 MTJs in neural spiking networks 25
1.3.5 MTJs as probabilistic bits in Boltzmann machines 27
1.4 Dissertation overview 34
Chapter 2 Fabrication and characterization of MTJs 37
2.1 Fabrication process 37
2.1.1 Bottom electrode definition 38
2.1.2 Pillar definition 39
2.1.3 Top electrode definition 39
2.1.4 Magnetic thermal annealing 40
2.2 Device characterization 41
2.2.1 Field switching measurements 42
2.2.2 Current/voltage switching measurements 44
2.2.3 Synchronous switching probability measurements 45
2.2.4 Asynchronous switching probability measurements 49

Chapter 3 Tunable random telegraphic signal generation via dual-biasing.....53
3.1 Motivation and background.....53
3.2 Experimental set-up.....55
3.3 Single MTJ results.....58
3.4 Bipolar signal generation.....60
3.5 Discussions.....70
3.5.1 Comparison with Stoner-Wolfarth model.....71
3.5.2 Physical mechanism for the two-degrees of tunability capability.....74
3.5.3 Influence of intrinsic thermal stability.....82
3.6 Summary.....91
Chapter 4 Interplay between SOT and VCEC.....94
4.1 Motivation and background.....94
4.2 Experimental set-up.....96
4.3 DC switching results.....100
4.3.1 Effect of VCEC voltage.....100
4.3.2 Effects of VCEC voltage and SOT current.....101
4.3.3 SOT-assisted VCEC switching.....103
4.4 Discussions.....106
4.4.1 Factors influencing coercivity.....106
4.4.2 Factors influencing stray field.....110
4.4.3 Influence of shift field on VCEC switching.....112
4.4.4 Influence of SOT current on VCEC switching.....115
4.5 SOT and VCEC for field-free dual-biasing.....117
4.6 Summary.....119
Chapter 5 Stochastic Computational Random-Access Memory (SC-CRAM) 121
5.1 Motivation and background.....121
5.2 SC-CRAM overview.....123
5.2.1 Task scheduling.....124
5.2.2 Arithmetic functions in SC-CRAM.....126
5.2.3 Advantages of SC-CRAM.....130
5.3 Examples of SC-CRAM applications.....132

5.3.1 Local image thresholding 133
5.3.2 Bayesian inference system 135
5.3.2 Bayesian belief network 137
5.4 Accuracy analysis and discussions 138
5.4.1 Accuracy analysis 139
5.4.2 Performance evaluation 145
5.5 Summary 148
Chapter 6 MTJs for radiation-resilient memory 150
6.1 Motivation 150
6.2 Experimental set-up 153
6.3 Results 155
6.3.1 DC switching results 155
6.3.2 Pulsed switching results 158
6.4 Discussion 166
6.5 Summary 169
Chapter 7 Thesis Conclusion 171
Bibliography 174
Appendix A. Matlab codes for SC-CRAM functions 186
Appendix B. Matlab codes for SC-CRAM applications 215
Appendix C. Publication List 231

List of Tables

Table 1.1. Parameters for I_{OUT} and preset state of output MTJ for CRAM Boolean logic functions 21

Table 3.1. Overview of bipolar and binary synchronization ranges 67

Table 3.2. Intrinsic properties and characterization of all MTJs tested 83

Table 5.1. Number of MTJs and computation steps for various arithmetic functions in SC-CRAM (relative to conv-CRAM). 131

Table 5.2. MTJ specifications for SC-CRAM performance analysis 138

Table 5.3. Performance evaluation of SC-CRAM vs conv-CRAM. The values shown are relative to conv-CRAM.....145

Table 6.1. List of devices tested for analysis of the effects of total ionizing dose on p-MTJs 153

List of Figures

Figure 1.1. **(a)** General schematic of a typical MTJ stack and magnetization orientation of the free (FL) and reference (RL) layers for **(b)** in-plane and **(c)** out-of-plane MTJs. 2

Figure 1.2. Illustration of the spin dependent tunneling effect in the parallel (left) and anti-parallel (right) states. 3

Figure 1.3. Schematic of MRAM cells based on **(a)** field switching, **(b)** STT switching, **(c)** SOT switching, and **(d)** VCEC switching. 7

Figure 1.4. Illustration of spin transfer torque switching via the spin filtering effect for **(a)** AP-to-P switching and **(b)** P-to-AP switching. 9

Figure 1.5. Illustration of **(a)** the spin Hall effect and **(b)** the Rashba-Edelstein effect. 10

Figure 1.6. **(a)** Overall stack structure of p-MTJ with SAF free layer and illustration of reflectivities at both positive and negative E-field. **(b)** Illustration of difference in phases in electron wave function between spin-up and spin-down states. **(c)** Illustration of the influence of E-field on the penetration depth of the electron wave function. **(d)** Dependence of ΔE (difference between energy in FM and AFM coupling configurations) on applied E-field. 11

Figure 1.7. Generalized configuration of various neural networks including **(a)** feed-forward networks, **(b)** recurrent networks, **(c)** Hopfield networks, and **(d)** Boltzmann machines. 14

Figure 1.8. Implementation of neural networks for feature extraction (images extracted from [19]). 15

Figure 1.9. Data encoding with temporal sequences of electrical spikes via (a) rate coding and (b) temporal coding. Images extracted from [26] and [27], respectively 18

Figure 1.10. General schematic of **(a)** STT-CRAM and **(b)** SOT-CRAM. 19

Figure 1.11. Current flow during **(a)** memory write operation, **(b)** memory read operation, and **(c)** logic mode (image extracted from [30]).....20

Figure 1.12. Logic gates used for efficient implementation of various arithmetic functions including **(a)** Y = 1-A (NOT gate), **(b)** unipolar multiplication (AND gate), **(c)** bipolar multiplication (XNOR gate), **(d)** addition scaled by 1/2 (MAJ gate), **(e)** addition scaled by a factor of S (2x1 MUX), and **(f)** Y = A+B-AB (OR gate) 22

Figure 1.13. Comparison between conventional and stochastic implementation for image processing at **(a)** no noise, **(b)** 1%, **(c)** 2%, **(d)** 5%, **(e)** 10%, and **(f)** 15% noise injection (image extracted from Ref. [38]).....23

Figure 1.14. Examples of MTJ-based approaches for generating random numbers centered at 0.5 including **(a)** conditional perturb and real-time output probability tracking (image extracted from Ref. [46]) and **(b,c)** digitally controlled probability-locked loop (images extracted from Ref. [47]) 24

Figure 1.15. Examples of MTJ-based analog-to-stochastic bit-stream conversion circuits for **(a)** vision chips (image extracted from Ref. [52]) and **(b)** Bayesian inference networks (image extracted from Ref. [53]).....25

Figure 1.16. Proposed MTJ-based designs of **(a)** a stochastic neuron and **(b)** a stochastic synapse in spiking neural networks. **(c)** Architecture for a spintronic-based core computing kernel. Images were extracted from [62].....26

Figure 1.17. **(a)** Generic structure of the p-bit, **(b)** ideal transfer curve of the p-bit, which is based on the time-averaged magnetization state of the nano-magnet ($\langle m_i \rangle$), and **(c)** time-domain plots showing $\langle m_i \rangle$ fluctuations at various inputs (normalized units). (Images were extracted from [64]).....29

Figure 1.18. **(a)** Circuit diagram of an invertible AND gate with a network of 3 interconnected p-bits. **(b)** Output frequency for all combinations of [A B C], which shows successful implementation of an AND gate since the dominant cases follow proper AND logic criteria. Demonstration of invertible AND

logic with output bit C clamped to **(c)** logic state '0' and **(d)** logic state '1'. (Images were extracted from [64])....30

Figure 1.19. **(a)** An example of a 5-city travelling salesman problem implemented using a network of 16 p-bits. **(b)** An example of a Bayesian inference network to solve genetic relatedness problems implemented using p-bit circuits. (Images extracted from [67])....32

Figure 1.20. Structure of a p-bit cell using **(a)** SOT and **(b)** STT switching mechanisms. (Images extracted from [67]) 33

Figure 1.21. **(a)** Diagram of information flow and circuit diagram of invertible AND gate **(b)** Waveforms of a ramp input test signal and output A, B, and C of all three p-bit block. **(c)** Relative histogram of [ABC] with various combinations of A, B, and C clamped. These experimental results confirm the simulation results shown in Fig. 1.18(b) – 1.18(d) (Images extracted from [80])..... 34

Figure 2.1. General overview of MTJ fabrication process. **(a)** Original MTJ stack consisting of the ferromagnetic free layer, the MgO tunneling barrier, the ferromagnetic fixed layer, the metallic seed layers, and the Si substrate (top-to-bottom). Definition of the **(b)** bottom electrode, **(c)** MTJ nanopillars and contact Vias, and **(d)** top electrodes. Note that SiO_2 is deposited after patterning the MTJ nanopillars, however, this step is not shown in this image 37

Figure 2.2. Illustration of Photolithography process....38

Figure 2.3. Image of the post-annealing set-up 41

Figure 2.4. **(a)** Illustration of ground-signal-ground probe layout. **(b)** Image of 3-D projected magnet stage 42

Figure 2.5. Example of a resistance vs field plot where the data is obtained from an in-plane MTJ patterned into a 110 nm x 45 nm elliptical nanopillar 44

Figure 2.6. Example of a **(a)** resistance vs current and **(b)** the corresponding voltage vs current data from an in-plane MTJ patterned into a 110 nm x 45 nm elliptical nanopillar 44

Figure 2.7. **(a)** Examples of P_{SW} vs V_P curves for **(a)** P to AP and **(b)** AP to P switching. **(c)** Illustration of the three STT switching modes and their dependence on pulse width (image extracted from [83]) 47

Figure 2.8. **(a)** Task scheduling for synchronous probabilistic switching measurements. **(b)** Diagram of testing set-up. 48

Figure 2.9. **(a)** Illustration of the energy barrier between the AP- and P-states for MTJs with thermal stability factors below 20 k_bT. **(b)** Example of telegraphic switching signals generated measured from low-Δ MTJs. **(c)** Example of the average magnetization curve with respect to bias voltage (image extracted from [90]). **(d)** Example of an I_C intercept vs H_{bias} plot used to calculate H_{stray}. **(e,f)** AP and P-state components of the effective thermal stability factor with respect to the bias current 49

Figure 2.10. Illustration of the testing set-up for asynchronous probabilistic switching measurements for **(a)** using a bias current where the MTJ and oscilloscope are connected in parallel and **(b)** using a bias voltage where the MTJ and oscilloscope are connected in series 52

Figure 3.1. **(a)** Illustration of the evolution of the energy barrier between the AP and P states during the dual-biasing process. **(b)** Schematic of the dual-biasing experiment 55

Figure 3.2. **(a)** Field switching minor loop of the MTJ tested and illustration of the orientation of H_{bias} and V_{bias} of the dual-biasing method (inset). **(b)** Time-domain data showing the effects of H_{bias} and V_{bias} on both the average resistance state and average switching rate 56

Figure 3.3. **(a)** AP-rate curves vs V_{bias} at all H_{bias} values tested. **(b,c)** V_{bias} dependence of the AP- and P-state components of the effective thermal stability factors at fixed H_{bias} values. **(d,e)** H_{bias} dependence of the AP- and P-state components

of the effective thermal stability factors at fixed V_{bias} values. Images were extracted from [78].....59

Figure 3.4. ***(*a)** Experimental set-up for generating bipolar random numbers. ***(*b)** Field switching plots for both top and bottom MTJs tested. Images extracted from [94] 63

Figure 3.5. Time domain data at a bias field of 15 Oe and bias voltage of 1060 mV of **(a)** output voltages measured from oscilloscope, **(b)** bipolar representation of the output signal, and **(c)** binary representation of the output signal. Images extracted from [94].....64

Figure 3.6. Average dwell times for both low-state and high-state pulses from bipolar signals at bias fields of **(a)** 5 Oe, **(b)** 10 Oe, **(c)** 15 Oe, and **(d)** 20 Oe. At each bias field, the synchronization region (SR), where $T_H \approx T_L$, is indicated. Average bipolar state at bias fields of **(e)** 5 Oe, **(f)** 10 Oe, **(g)** 15 Oe, and **(h)** 20 Oe. Each bias field shows the range of data points which passed the three NIST STS tests used in our analysis, which is labelled "Bipolar RNG". Images extracted from [94].....65

Figure 3.7. Average dwell times for both low-state and high-state pulses from binary signals at bias fields of **(a)** 5 Oe, **(b)** 10 Oe, **(c)** 15 Oe, and **(d)** 20 Oe. At each bias field, the synchronization region (SR), where $T_1 \approx T_0$, is indicated. Average bipolar state at bias fields of **(e)** 5 Oe, **(f)** 10 Oe, **(g)** 15 Oe, and **(h)** 20 Oe. Images extracted from [94].....69

Figure 3.8. Illustration of the energy barrier between AP- and P- states at **a.)** equilibrium (no external biases), **b.)** $H_{bias} > 0$ (favoring P-state orientation, **c.)** $V_{bias} > 0$ (favoring AP-state orientation, and **d.)** both H_{bias} and $V_{bias} > 0$.....72

Figure 3.9. Comparison between experimental data (also presented in Fig. 4.3) and calculations using the Stoner-Wohlfarth model for **(a-b)** Δ_{AP} measured and calculated **(c-d)** Δ_P measured and calculated.....74

Figure 3.10. ***(a)*** AP-rate versus V_{bias} at several fixed bias fields, ***(b)*** fitting parameters V_{fit} and V_0 with respect to H_{bias} and ***(c)*** contour plot of the AP-rate over all the bias fields and bias voltages tested. Images extracted from [78].....75

Figure 3.11. Calculations for the fitting coefficients ***(a)*** C_1, ***(b)*** C_2, and ***(c)*** C_3 in Eq. 4.9 vs V_{bias}. For each plot, the black and red dotted lines represent the maximum and minimum error bounds (95% confidence) and the solid blue line is the average value within the error bounds 76

Figure 3.12. **(a)** $H_K^{\uparrow}$ vs V_{bias} plots determined from linear fits of H_{bias} vs $(\Delta_{AP})^{1/2}$ plots. **(b)** Plots of the intrinsic critical switching voltage with respect to voltage bias. These values were calculated using the raw Δ_{AP} and Δ_P measurements as well as the H_K calculations.....78

Figure 3.13. Percent changes in ν, h, and $\Delta_{AP/P}$ with respect to V_{bias}, where the ***(a)*** AP-state and ***(b)*** P-state components are calculated at fixed fields of 122, 127, 132, 142, and 152 Oe, and with respect to H_{bias}, where the ***(c)*** AP-state and ***(d)*** P-state components are calculated at fixed bias voltages of 220, 240, 260, 280, and 300 mV. In each case, the range of V_{bias} or H_{bias} values in which these percent changes are being calculated are set in order to produce AP-rates between -0.8 to +0.8, which are determined using our fitting parameters of the AP-rate data.....80

Figure 3.14. Resistance vs field plots for **a.)** MTJ 6 and **b.)** MTJ 2 (see table 3.2). Images were extracted from [79] 84

Figure 3.15. **(a)** AP-rate versus bias voltage curves at H_{min} and H_{max} for a semi-stable MTJ and an sMTJ (MTJs 6 and 2 from table 3.2). H_{min} = 60 Oe and 30.5 Oe and H_{max} = 80 Oe and 55.5 Oe for MTJ 6 and 2, respectively. **(b)** Line plots showing the center (V_0) and width (V_R) of the voltage dependent transfer curves for each MTJ category at H_{min} and H_{max}. Images were extracted from [79] 85

Figure 3.16. **(a)** Average switching rate vs bias voltage at H_{max} and H_{min} for MTJs 6 and 2 (see table 3.2) and **(b)** Maximum switching rate (f_{max}) achieved for each MTJ category. Images were extracted from [79].....86

Figure 3.17. Effective thermal stability factor versus bias voltage at H_{max}, H_{min}, and H_{mid} for ***(*a)** AP-state and ***(b)*** P-state components for MTJs 6 and 2 (see table *3.2*). Note that H_{mid} = 70 Oe for MTJ 6 and 45.5 Oe for MTJ 2.....88

Figure 3.18. Degree of separation between Δ_{AP} and Δ_P tunability with voltage and field (D_V and D_H, respectively) for each MTJ category. D_V represents the order of magnitude decrease in τ_P for every 1 order of magnitude increase in τ_{AP}, and D_H represents the order of magnitude decrease in τ_{AP} for every 1 order of magnitude increase in τ_P. Image extracted from [79].....89

Figure 3.19. Slopes Δ_{AP} and Δ_P with respect to **(a)** V_{bias} and **(b)** H_{bias} for each MTJ category. Images extracted from [79].....90

Figure 4.1. **(a)** Diagram of probe connection with respect to the SAF p-MTJ device with top view of electrode pattern for all 5 terminals, and **(b)** the out-of-plane minor loop of resistance (R_{MTJ}) vs. external magnetic field (H_{ext}) curve of 150-nm p-MTJ devices with a current of 2 nA (V_{VCEC} ~ 2.5 mV). Images were extracted from the Supplementary Information in [123] 97

Figure 4.2. **(a)** R_{MTJ} vs. H_{ext} curves at V_{VCEC} = ±0.75 V. Measurements and linear fits of **(b)** H_C vs. V_{VCEC} and **(c)** H_{stray} vs. V_{VCEC}. **(d)** Magnetic anisotropy (K_{eff}) vs. V_{VCEC} obtained from H_C measurements fitted through the Kurkijärvi-Fulton-Dunkelberger equation. Image was extracted from the Supplementary information in [123] 100

Figure 4.3. **(a, b)** External magnetic field (H_{ext}) vs. normalized TMR loops of 150-nm SAF p-MTJ pillars with a [Co (0.3 nm)/Pd (0.7 nm)]$_3$/Co (0.3 nm)/Ru (0.6 nm)/Ta (0.3 nm)/CoFeB (1.0 nm) SAF free layer and the 250-nm Ta (5.0 nm)/Pd (2.0 nm) bi-layered spin Hall channel at I_{SOT} = 0, 0.25 mA, 0.75 mA, and 1.25 mA (J_{SOT} = 0, 1.4×10^7 A/cm^2, 4.3×10^7 A/cm^2, and 7.1×10^7 A/cm^2) with V_{VCEC} of +0.5 V and -0.5 V. **(c, d)** H_C and H_{stray} as a function of V_{VCEC}

at I_{SOT} = 0, 0.5 mA, and 1.0 mA (J_{SOT} = 0, 2.9×10^7 A/cm^2, and 5.7×10^7 A/cm^2), where each set shows the slope of H_C and H_{stray} with V_{VCEC}, which were obtained through linear fits. **(e, f)** H_C and H_{stray} as a function of I_{SOT} at V_{VCEC} = -0.5 V, 0.1 V, and +0.5 V, where both H_C and H_{stray} show non-linear behavior with I_{SOT}. Image was extracted from [123] 102

Figure 4.4. **(a)** Schematic of p-MTJ devices with SAF free layers switched through SOT + VCEC and the 150-nm MTJ pillar. The SAF free layer has a stack of [Co (0.3 nm)/Pd (0.7 nm)]$_3$/Co (0.3 nm)/Ru (0.6 nm)/Ta (0.3 nm)/CoFeB (1.0 nm) and the bi-layered spin Hall channel has a stack of Ta (5.0 nm)/Pd (2.0 nm) and a width of 250-nm. **(b)** R_{MTJ} vs. V_{VCEC} loop of the 150-nm p-MTJ pillar, where H_{shift} = -245 Oe and J_{SOT} = 6.8×10^7 A/cm^2 (I_{SOT} = 1.2 mA) are utilized to assist VCEC switching. Note that H_{shift} is used to center the R_{MTJ} vs V_{VCEC} curve around V_{VCEC} ~ 0. Image was extracted from [123]..... 104

Figure 4.5. **(a)** The measured and fitted H_C vs. I_{SOT}, illustrating that H_C is slightly smaller at positive I_{SOT}. **(b)** H_C changes as a function of I_{SOT}, following the origin from H_{SOT}^{IP}, E_{SOT}, JH, H_{SOT}^{Z}, and the combined influence of all four factors. **(c)** The measured and calculated H_C change (δH_C) as a function of I_{SOT}. Image extracted from the Supplementary information in [123] 106

Figure 4.6. **(a)** H_{stray} vs. I_{SOT}, illustrating that H_{stray} is larger for positive I_{SOT}. **(b)** H_{stray} as a function of I_{SOT}, following the origin from H_{SOT}^{Z}, E_{SOT}, JH, and the combined influence of all factors. **(c)** The measured and calculated H_{stray} change (δH_{stray}) as a function of I_{SOT}. Image extracted from the Supplementary information in [123] 110

Figure 4.7. **(a)** VCEC switching curves at I_{SOT} = 0.8 mA with H_{shift} = -240 Oe, -230 Oe, and -220 Oe and **(b)** the corresponding binary state of the p-MTJs, where AP-state denotes a high binary state. **(c)** VCEC switching curves at I_{SOT} = 0.9 mA with H_{shift} = -240 Oe and -230 Oe and **d)** the corresponding binary state of the p-MTJs. Curves for H_{shift} = -240 Oe and -230 Oe are shifted along

the vertical axis in each plot for easy comparison between these curves. Image extracted from the Supplementary information in [123] 113

Figure 4.8. ***(a)*** VCEC switching curves at H_{shift} = -240 Oe with I_{SOT} = 0.7 mA, 0.8 mA, and 0.9 mA and ***(b)*** the corresponding binary state of the p-MTJs, where AP-state denotes a high binary state. ***(c)*** VCEC switching curves at H_{shift} = -230 Oe when I_{SOT} = 0.7 mA, 0.8 mA, and 0.9 mA and ***(d)*** the corresponding binary state of the p-MTJs. Curves for I_{SOT} = 0.7 mA and 0.8 mA are shifted along the vertical axis. Image extracted from the Supplementary information in [123] 116

Figure 4.9. **(a)** R_{MTJ} vs I_{VCEC} plot showing stochastic VCEC switching at H_Z = -245 Oe and J_{SOT} = 71.4 MA/cm^2 (I_{SOT} = 1.25 mA) and **(b)** binary representation of R_{MTJ} measurements, where the AP-state corresponds to binary state '1' (Black line with circles) and a moving average of the binary value (red line) 118

Figure 5.1. Current paths for **(a)** reset, **(b)** perturb, **(c)** logic, and **(d)** read steps for multiplication of bit-streams A and B using AND logic in SC-CRAM..... 124

Figure 5.2. **(a)** Circuit diagram of a 2-to-1 MUX and (**b)** scheduling operations in SC-CRAM. For simplicity, reset steps and final read steps are not shown..... 126

Figure 5.3. **(a)** Circuit diagram of an XOR gate and (**b)** scheduling operations in SC-CRAM. For simplicity, reset steps and final read steps are not shown..... 127

Figure 5.4. **(a)** Circuit diagram the square root function in SC-CRAM where $Y \approx \sqrt{X}$, X_1 and X_2 are independent bit-streams with probabilities equal to X, and C_1 and C_2 are bit-streams with fixed probabilities of 0.67 and 0.18, respectively. (**b)** scheduling operations in SC-CRAM. For simplicity, reset steps and final read steps are not shown.....129

Figure 5.5. **(a)** Circuit diagram for a JK flip-flop to compute the function $Y \approx A/(A+B)$ and (**b)** scheduling operations in SC-CRAM. For simplicity, reset steps and final read steps are not shown.....130

Figure 5.6. Illustrations of the applications analyzed. **(a)** Effect of local binarization on a degraded image (extracted from [41]), (**b)** plot of data fusion results for object location (extracted from [53]), and **(c)** Bayesian belief network for heart disaster prediction (extracted from [53]) 133

Figure 5.7. Block diagram for calculating local thresholds in SC-CRAM using input bit-streams, A (based on pixel intensities). Note that bit-streams for $A_{i,1}$ and $A_{i,2}$ are independent but have equal probabilities.....133

Figure 5.8. Block diagram of a Bayesian inference system in SC-CRAM for object location on a 64 x 64 2D grid.....136

Figure 5.9. Block diagram of a Bayesian belief network in SC-CRAM for heart disaster prediction 138

Figure 5.10. Accuracy of **(a)** AND multiplication **(b)** MAJ3 scaled addition, **(c)** MUX scaled addition, **(d)** XOR absolute value subtraction **(e)** square root with AND-OR-OR logic, and **(f)** scaled division with a JK flip flop in SC-CRAM 140

Figure 5.11. Illustrations of **(a)** degraded gray-scale image for 'SC-CRAM' sign used to test the accuracy of local image thresholding in SC-CRAM. **(b-d)** Binarized image at output of local image thresholding circuit for **(b)** 0% noise injection, **(c)** 10% noise injection, and **(d)** 30% noise injection 142

Figure 5.12. The map of 2-D plane for object location via a Bayesian inference system with **(a)** no noise **(b)** 10% noise injection, and **(c)** 30% noise injection ...143

Figure 5.13. Output accuracy vs noise injection rate for **(a)** local image thresholding **(b)** object location via Bayesian Inference, and **(c)** heart disaster prediction via a Bayesian belief network 144

Figure 5.14. Breakdown of energy consumption in SC-CRAM and conv-CRAM for **(a-b)** local image thresholding **(c-d)** object location via Bayesian Inference, and **(e-f)** heart disaster prediction via a Bayesian belief network 147

Figure 6.1. Sample data acquired for device 12 (d_{MTJ} = 80nm). ***(*a)** STT hysteresis plot showing DC switching properties. ***(*b)** Switching probability distribution plot measured in the P to AP state switching direction for pulse widths of 100ns. Both plots display the curves obtained before TID exposure, after 300 krad, and after 1 Mrad. Image was extracted from [152].....154

Figure 6.2. Changes of ***(*a)** zero-bias TMR, ***(*b)** V_H, ***(*c)** $I_{P\rightarrow AP}$, and ***(*d)** $I_{AP\rightarrow P}$ before TID exposure, after 300 krad, and after 1 Mrad for every device tested. Images were extracted from [152] 156

Figure 6.3. Change in ***(*a-b)** thermal stability, ***(*c-d)** intrinsic critical switching voltage, and ***(*e-f)** write energy at pulse widths of 500ns in the ***(*a, c, e)** P to AP and ***(*b, d, f)** AP to P switching directions. Images were extracted from [152]..... 160

Figure 6.4. Distribution of $d\Delta$ in the ***(*a-c)** P to AP and ***(*d-f)** AP to P switching directions at pulse widths of ***(*a, d)** 500ns, ***(*b, e)** 100ns, and ***(*c, f)** 50ns. Images were extracted from [152].....162

Figure 6.5. Distribution of dV_{C0} in the **a-c.)** P to AP and **d-f.)** AP to P switching directions at pulse widths of **a, d.)** 500ns, **b, e.)** 100ns, **c, f.)** 50ns. Images were extracted from [152] 163

Figure 6.6. dE_{WRITE} distribution in the ***(*a-c)** P to AP and ***(*d-f)** AP to P switching direction for pulse widths of ***(*a, d)** 500ns, ***(*b, e)** 100ns, and ***(*c, f)** 50ns. Images were extracted from [152].....165

Figure 6.7. Calculations for the mean and standard deviation of ***(*a-b)** $d\Delta$, ***(*c-d)** dV_{C0}, and ***(*e-f)** dE_{WRITE} with respect to pulse width. Images were extracted from [152] 167

Chapter 1: Introduction

1.1 Magnetic Tunnel Junctions

The elementary storage component in state-of-the-art magnetoresistive random access memory (MRAM) is the magnetic tunnel junction (MTJ). The MTJ is a two terminal nanomagnetic device consisting of several thin-film magnetic device which consists of several ultra-thin layers (~1 – 10nm) forming a nanopillar. Schematic of a typical MTJ is shown in Fig. 1.1(a). There are four key layers that determine the properties of the MTJ which are the pinning layer, the fixed layer, the tunnel barrier, and the free layer. The fixed and the free layer are both ferromagnetic layers where the magnetization has either in-plane orientation or an out-of-plane orientation. MTJs with in-plane magnetization patterned are typically patterned into elliptical nano-pillars with the magnetization pointing along the long-axis of the ellipse and MTJs with out-of-plane magnetization are patterned into circular nano-pillars. The magnetization of the fixed layer is pinned in a pre-determined direction, which is set through an exchange interaction with the pinning layer. The pinning layer consists of an anti-ferromagnetic material, typically PtMn or IrMn. The magnetization of the free layer on the other hand can be oriented in the same direction (parallel) or in the opposite direction (anti-parallel) to the magnetization of the fixed layer. For in-plane MTJs, these two directions are determined by the orientation of the long-axis of the elliptical cross section as illustrated in Fig. 1.1(b). For out-of-plane MTJs, the magnetization either points towards or away from the interface of the free layer and tunnel barrier as illustrated in Fig. 1.1(c).

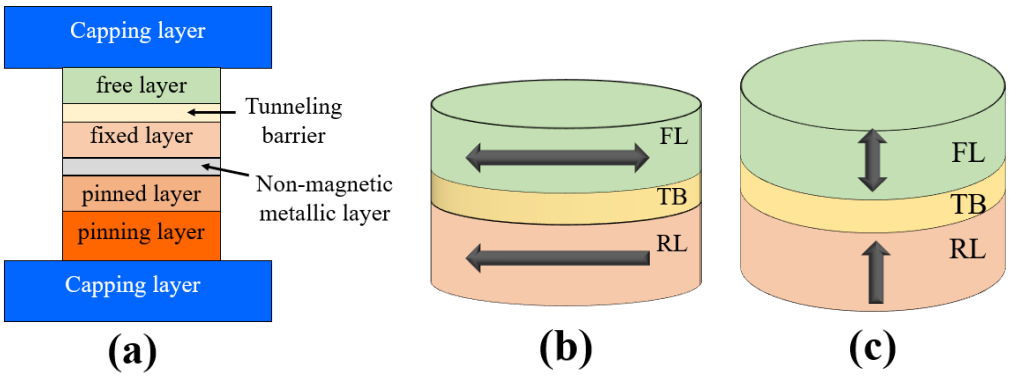


Fig. 1.1. (a) General schematic of a typical MTJ stack and magnetization orientation of the free (FL) and reference (RL) layers for **(b)** in-plane and **(c)** out-of-plane MTJs.

1.1.1.) Tunneling Magnetoresistance

An MTJ can exist in one of two binary states determined by the magnetization direction of the free layer ($\hat{m}_f$) relative to the magnetization direction of the fixed layer ($\hat{m}_p$). In an MTJ, $\hat{m}_f$ can either be parallel (P) or anti-parallel (AP) to $\hat{m}_p$. The magnetization state can be determined through the resistance of the MTJ (R_{MTJ}) where R_{MTJ} is larger in the AP-state than in the P-state, which will be referred to as R_{AP} and R_P, respectively. The difference in R_{AP} and R_P is described by the tunneling magnetoresistance (TMR) ratio, where TMR = $([R_{AP} - R_P]/R_P)*(100\%)$.

TMR arises from the differences in tunneling probabilities for spin up electrons between the AP and P-states (T_{AP} and T_P, respectively), as illustrated in Fig. 1.2. During the tunneling process, electrons preserve their spin state, meaning they will only tunnel to sub-bands with the same spin orientation. When the MTJ switches from the P-state to the AP-state, the density of states at the fermi energy level for spin up electrons decreases, indicating that $T_{AP} < T_P$. Spin polarization is defined as the ratio of the difference in tunneling probabilities between spin up and spin down electrons to the total tunneling

probability, as seen in Eq. 1.1 [1-2]. The TMR ratio can be expressed in terms of spin polarization of the free and fixed layers (P_f and P_p, respectively) according to the Julliere model [3], as seen in Eq 1.1. Note that most modern MTJs use CoFeB for both the free and fixed layers, meaning $P_f = P_p$.

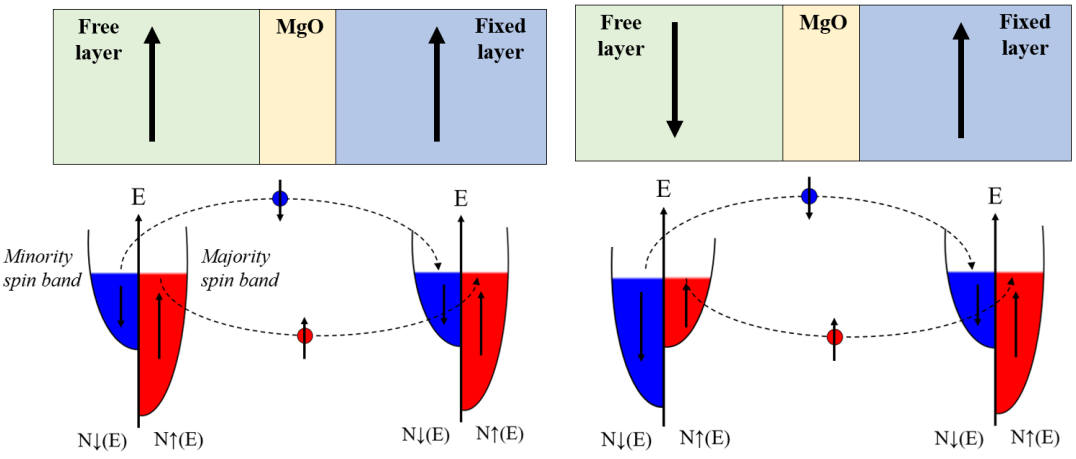


Fig. 1.2. Illustration of the spin dependent tunneling effect in the parallel (left) and anti-parallel (right) states.

$$TMR = \frac{2P_f P_p}{1 - P_f P_p} \quad where \quad P_{f(p)} = \frac{T_{f(p)}^{\uparrow} - T_{f(p)}^{\downarrow}}{T_{f(p)}^{\uparrow} + T_{f(p)}^{\downarrow}} \tag{1.1}$$

1.1.2.) Magnetic Anisotropy

The reason that $\hat{m}_f$ can only point along AP and P-state directions is due to magnetic anisotropy, which describes the preference for the magnetization of a ferromagnetic sample to lie in a particular direction. Anisotropy is quantified as the energy per unit volume required to change the magnetization from the lowest energy direction (easy axis) to its highest energy direction (hard axis). There are three kinds of anisotropy: magnetocrystalline anisotropy, shape anisotropy, and interfacial anisotropy. Magnetocrystalline anisotropy (E_{MA}) is an intrinsic property that refers to the

magnetization direction preference caused by the atomic lattice structure of the sample. In magnetic thin films, E_{MA} depends on the angle between the magnetization and the easy axis as well as the bulk anisotropy constant, K_B, as shown in Eq. 1.2. Shape anisotropy (E_{MS}) is caused by the demagnetization fields (H_d) at the edges of the sample. In thin films, E_{MS} is entirely dependent on the direction of $\hat{m}_f$ and saturation magnetization of the sample (M_S), since H_d is a function of M_S, as shown in Eq. 1.3. For elliptical MTJs, H_d is stronger when $\hat{m}_f$ is along the short axis than the long axis, therefore, the preferred direction of $\hat{m}_f$ will be along the long axis. Interfacial anisotropy (E_I) is induced at the interface between magnetic and non-magnetic layers in thin films. E_I is dependent on the thickness of the magnetic layer (t_f) and the interfacial anisotropy constant, K_I and defined as K_I/t_f. In MTJs, the effective anisotropy, K_U, is the combined influence of E_{MA}, E_{MS}, and E_I as shown in Eq. 1.4. The stack of modern MTJs are based CoFeB free layers and MgO free layers, where the relative contribution of E_{MA} is considered to be negligible [4], therefore, K_U is completely dependent on the contributions of E_{MS} and E_I. When t_f is sufficiently large, the MTJ E_{MS} will dominate, therefore, $\hat{m}_f$ will favor in-plane orientation. For t_f sufficiently small so that $K_U > 0$, E_I will dominate and therefore, $\hat{m}_f$ will favor out-of-plane direction. Experiments have shown that for CoFeB-MgO based MTJs, $\hat{m}_f$ favors out-of-plane orientation when $t_f \leq 1.3$ nm [4].

$$E_{MA} = K_{B,0} + K_{B,1}\sin^2(\theta) \tag{1.2}$$

$$E_{MS} = -\frac{1}{2}\overrightarrow{H_d}\cdot\overrightarrow{M_S} \quad \rightarrow \quad |\overrightarrow{H_d}| = -4\pi M_S \quad \rightarrow \quad E_{MS} = 2\pi M_S^2 \tag{1.3}$$

$$K_U = K_B - 2\pi M_S^2 + \frac{K_I}{t_f} \tag{1.4}$$

A key metric used to determine the MTJ's performance, besides TMR ratio, is the thermal stability factor (Δ), which is defined as the ratio of the energy barrier required for $\hat{m}_f$ to switch states (E_B) to the thermal energy (k_bT, where k_b in Boltzmann's constant and T is the temperature). The retention time (τ) represents the expected amount of time the MTJ will retain its state until thermal fluctuations cause $\hat{m}_f$ to switch and is expressed using the Néel-Arrhenius equation, $\tau = \tau_0 * \exp(\Delta)$, where τ_0 is the inverse attempt frequency, which is often assumed to be 1 ns [5]. For modern MRAM applications, the desired retention time is > 10 years, therefore Δ needs to be larger than 60. In the absence of external influences on $\hat{m}_f$, E_B is given by the anisotropy energy, therefore, Δ is defined using the expression in Eq. 1.5, where V is the volume of the MTJ's free layer. Equation 1.5 also shows the expression for Δ in terms of the anisotropy field (H_K), where $H_K = 2K_U/M_S$, which represents the hypothetical field needed to align $\hat{m}_f$ perpendicular to the easy-axis direction. Since MTJs with in-plane (IP) and perpendicular (PMA) orientation have different sources of magnetic anisotropy, H_K and Δ have separate expressions for IP and PMA MTJs. Equation 1.6 shows how the dimensions of in-plane MTJs influence H_K and Δ, where AR is the aspect ratio of the MTJ (long-to-short axis dimensions) and w is the cell width [6]. These expressions show that $H_K^{IP} \propto 1 - AR^{-1}$ meaning that Δ_{IP} increases as the eccentricity of the MTJ's elliptical shape increases (AR increases) and $\Delta_{IP} = 0$ for circular MTJs (AR = 1). On the other hand, Eq. 1.7 shows that H_K^{PMA} is not influenced by AR, therefore Δ_{PMA} is also not influenced by AR. For PMA MTJs, the parameter that has the largest influence on H_K^{PMA} and Δ_{PMA} is t_f, where Δ_{PMA} increases linearly as t_f decreases.

$$\Delta = \frac{K_U V}{k_b T} = \frac{H_K M_S V}{2 k_b T} \tag{1.5}$$

$$In-plane: H_K^{IP} = 2\frac{4\pi M_S t_f (AR-1)}{wAR}, \qquad \Delta_{IP} = \frac{\pi^2 \left(M_S t_f\right)^2 w(AR-1)}{k_b T} \tag{1.6}$$

$$perpendicular: \; H_K^{PMA} = \frac{2K_I}{M_S t_f} - 4\pi M_S, \qquad \Delta_{PMA} = \left(K_I - 2\pi M_S t_f\right) \tag{1.7}$$

1.2 MTJ switching mechanisms

In this dissertation, I will focus on four mechanisms used for switching resistance states in MTJs. These are magnetic field, spin-transfer torque (STT), spin-orbit torque (SOT), and voltage-controlled exchange coupling (VCEC). The oldest method of switching MTJ states in MRAM cells is through an external magnetic field generated by current carrying wires, which is illustrated in Fig. 1.3(a). The field generated needs to exceed the anisotropy field of the MTJ's free layer to reliably switch states. It should be noted that this method of switching MTJ states is no longer pursued in modern MRAM designs since the dimensions of a field switching MRAM cell cannot be scaled to sizes below 90nm due to the drastic increase in current required to generate fields with decreasing wire size [7-8]. However, from an experimental perspective, field switching measurements are still useful since it provides key intrinsic properties of the MTJ being tested (see section 2.2.1 for details of these measurements).

The scaling limitations of MRAM cells can be overcome through STT switching, which offers a much more practical writing method for MRAM cells than field switching. Figure 1.3(b) shows a general schematic of an STT-MRAM cell, where the MTJ state is manipulated with a voltage or current applied across the MTJ (V_{MTJ} and I_{MTJ}, respectively). A major short-coming of STT-MRAM is that the MTJs are susceptible to dielectric breakdown after multiple write cycles [9]. In recent years, researchers have attempted to develop MRAM cells which utilize SOT as the switching mechanism [10-15]. SOT devices

also use spin polarized currents to switch the MTJ, however, unlike an STT switching, SOT switching does not require the current to be applied directly across the MTJs tunneling barrier for switching. Instead, the current is passed through a non-magnetic heavy metal (HM) with strong spin-orbit coupling (typically Ta or W). An SOT-MRAM cell consists of an MTJ fabricated on top of an HM-based SOT channel which is adjacent to the MTJ's free layer, as illustrated in Fig. 1.3(c).

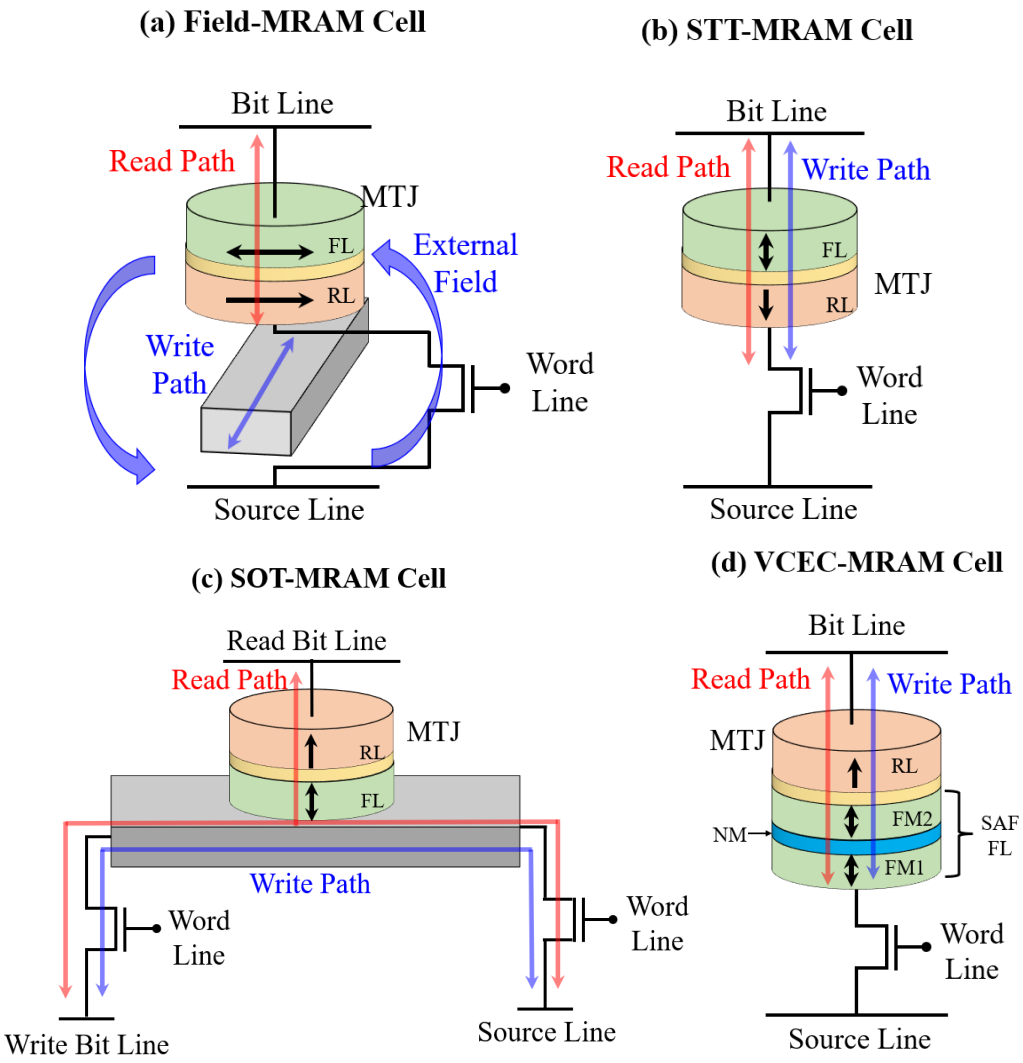


Fig. 1.3. Schematic of MRAM cells based on **(a)** field switching, **(b)** STT switching, **(c)** SOT switching, and **(d)** VCEC switching.

One key challenge for both STT and SOT-MRAM is the large switching current densities required for switching (J_C) where the best reported values for J_C are greater than 10^6 A/cm^2 for both STT and SOT [16-17]. VCEC is a promising solution to reduce the switching current densities in MRAM cells since VCEC switching has been accomplished with current densities as low as 10^5 A/cm^2 [18]. Fig. 1.3(d) shows a general schematic of a VCEC-MRAM cell, where the MTJ stack is designed so that the free layer is a synthetic antiferromagnetic (SAF) structure consisting of two FM layers separated by a non-magnetic metal (NM). As with STT switching, VCEC switching is accomplished by applying V_{MTJ} directly across the MTJ. However, VCEC switching relies on modulation of the interlayer exchange coupling (IEC) in the SAF free layer rather than the STT effect (see section 1.2.3 for more details).

1.2.1 Spin Transfer Torque

The mechanism that allows for STT switching is the spin filtering effect, which creates a spin polarized current through the MTJ. Figure 1.4(a) illustrates how the spin filtering effect causes switching in the AP-to-P direction when the polarization of V_{MTJ} is set so that the electrons flow from the fixed layer to the free layer. The density of states for electrons with spin states that align with the direction of $\hat{m}_p$ (↑ electrons) is larger than for electrons with opposing spin states (↓ electrons). This means that more ↑ electrons will tunnel to the free layer than ↓ electrons, thus creating a spin-polarized current. This spin polarization generates a torque on the free layer magnetization and eventually causes the MTJ to switch to the P-state.

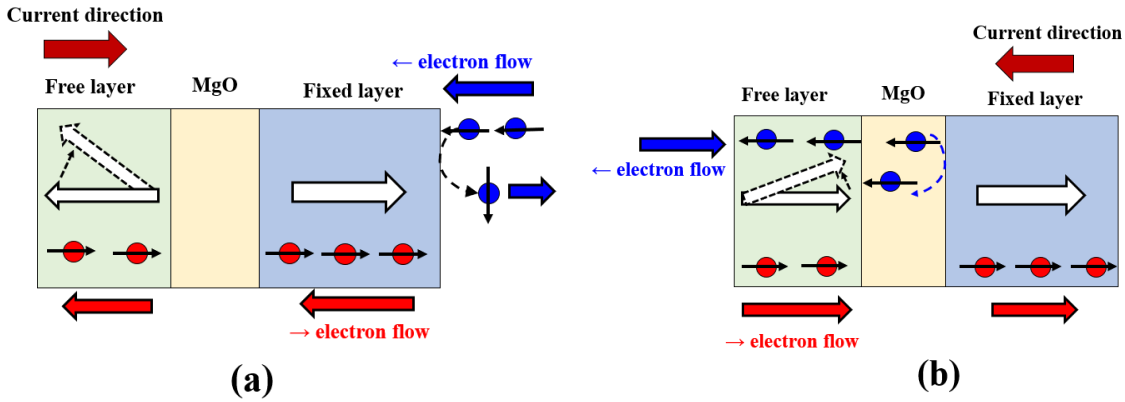


Fig. 1.4. Illustration of spin transfer torque switching via the spin filtering effect for **(a)** AP-to-P switching and **(b)** P-to-AP switching.

The STT effect can also cause switching in the P-to-AP direction when the polarization of V_{MTJ} is set so that the electrons flow from the free layer to the fixed layer, as illustrated in Fig. 1.4(b). Note that the current will still be spin polarized, however, the polarization will favor ↑ electrons, meaning that the spin polarization of the current at the free layer alone will not cause either layer to switch. However, since the magnetization of the fixed layer is larger than the free layer, the density of states for ↓ electrons is lower in the fixed layer than in the free layer. This means that some of the ↓ electrons will reflect at the interface of the fixed layer and tunnel barrier back to the free layer. These reflections will cause a shift in the density of states to favor ↓ electrons in the free layer and eventually cause the MTJ to switch to the AP-state.

1.2.2. Spin Orbit Torque

SOT switching is caused by two spin-orbit coupling (SOC) phenomena which are the spin hall effect (SHE) and the interfacial Rashba-Edelstein effect. Both effects initiate spin accumulation at the HM/FM interface, then the spin polarized electrons diffuse into the FM layer, thus exerting a torque on the FM magnetization. SHE is an SOC effect where a

charge current in the longitudinal direction is converted into a spin current in the transverse direction, as shown in Fig. 1.5(a). The transverse spin current is by spin resolved SOC scattering of the charge current in the HM layer where the ↑ and ↓ electrons are deflected in opposite directions. The polarization direction of the spin current and the charge-to-spin conversion efficiency is represented in the spin hall angle (θ_{SH}), which is an intrinsic property of the HM layer. The spin current (J_S) generated from charge current (J_C) is expressed in Eq 1.8, where σ is the polarization of the spin current, ħ is Planck's constant and e is the charge of an electron.

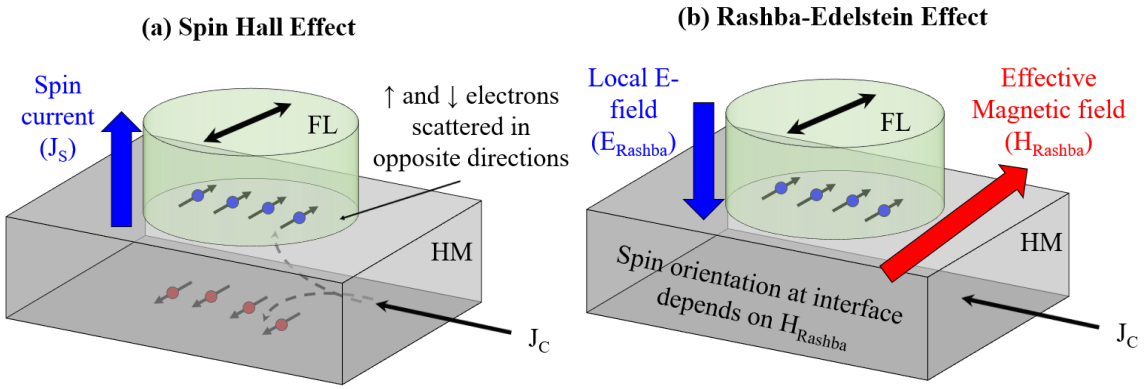


Fig. 1.5. Illustration of **(a)** the spin Hall effect and **(b)** the Rashba-Edelstein effect.

$$\overrightarrow{J_S} = \frac{\hbar}{2e}\theta_{SH}\left(\overrightarrow{J_C}\ x\ \vec{\sigma}\right) \tag{1.8}$$

The interfacial Rashba-Edelstein effect arises in structures with broken inversion symmetry. For HM/FM structures, an internal electric field (Rashba field), E_{Rashba}, is built up perpendicular to the interface of these layers, as illustrated in Fig. 1.5(b). An electron with momentum, p, passing through E_{Rashba} experience an effective magnetic field in the direction $\vec{E}_{Rashba}\ x\ \vec{p}$, thus polarizing the electrons and initiating spin accumulation. Therefore, there is a relation between p and spin polarization indicating that the magnitude

of the spin accumulation at the HM/FM interface is controlled by magnitude of the charge current.

1.2.3. Voltage – controlled exchange coupling

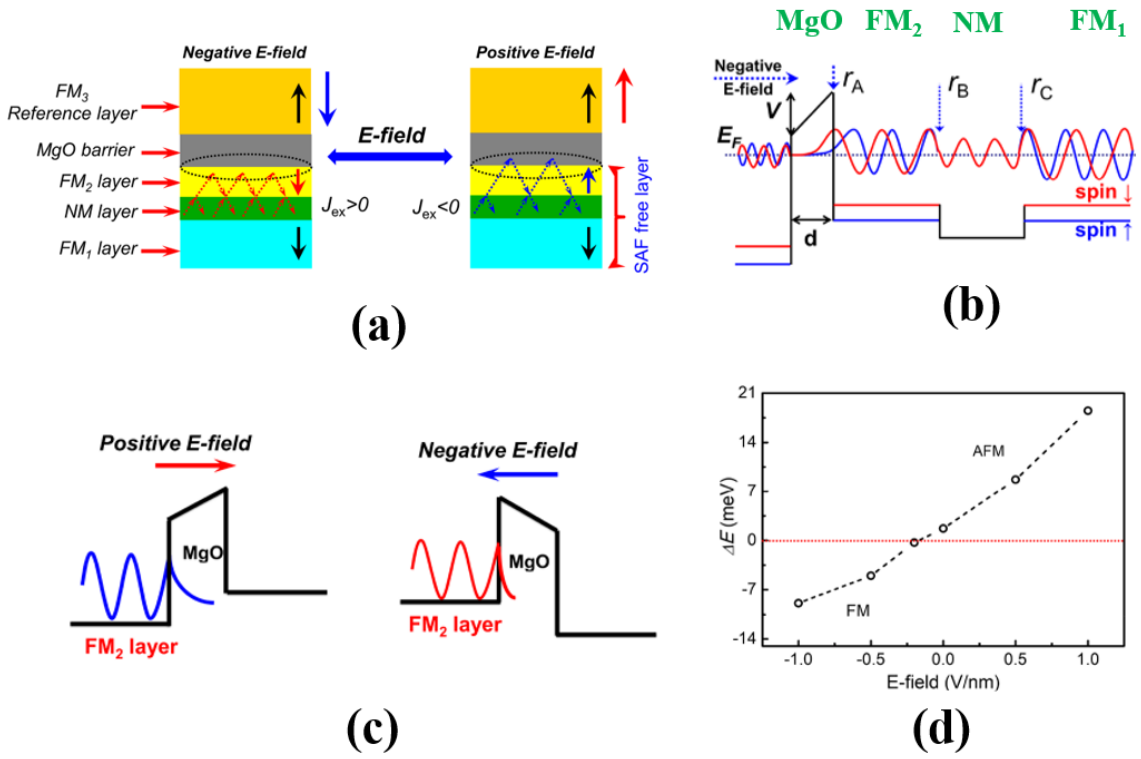


Fig. 1.6. (a) Overall stack structure of p-MTJ with SAF free layer and illustration of reflectivities at both positive and negative E-field. **(b)** Illustration of difference in phases in electron wave function between spin-up and spin-down states. **(c)** Illustration of the influence of E-field on the penetration depth of the electron wave function. **(d)** Dependence of ΔE (difference between energy in FM and AFM coupling configurations) on applied E-field. Images extracted from [18].

VCEC switching is accomplished through E-field modulation of IEC in MTJs with perpendicular anisotropy (p-MTJs) with SAF free layers. It should be noted that E-field effects on p-MTJs have been extensively studied, however, most studies investigate E-field effects relying on voltage-controlled magnetic anisotropy (VCMA). While VCMA can be used as a means of decreasing the MTJ's switching energy due to a reduction in interfacial

perpendicular anisotropy, it is not a writing mechanism by itself since the E-field is unipolar. Conversely, VCEC can achieve bidirectional switching since the sign of interlayer exchange coupling (IEC) between the two FM layers within the SAF-free layer structure can be tuned with the E-field due to modulation of the spin reflectivity at the interfaces. Therefore, unlike VCMA, VCEC can be used as an independent writing mechanism.

Figure 1.6(a) illustrates the overall stack structure of p-MTJs with a SAF free layer. In this example, FM_3 is the reference layer and $FM_2/NM/FM_1$ is the SAF free structure where FM_1 and FM_2 are coupled through the NM layer via the RKKY exchange interaction. When $V_{MTJ} = 0$, the coupling configuration and IEC strength of FM_1 and FM_2 are determined by the thickness of the NM layer. Note that in the example in Fig. 1.6(a), FM_1 and FM_2 are antiferromagnetically (AFM) coupled at $V_{MTJ} = 0$, meaning that the magnetization of FM_2 and FM_3 are oriented in opposite directions, thus, the MTJ is in the AP-state. Bruno's theory states that the energy preference for FM and AFM coupling configurations are dependent on the spin reflectivity at each interface in the SAF layers. The penetration length of the electron wave function is shifted when V_{MTJ} is applied, which changes the reflection phases, thus modulating IEC (both magnitude and sign).

Figures 1.6(b) shows how the penetration depth of the electron wave function differs between the spin-up and spin-down electrons and (c) shows how V_{MTJ} modulates the penetration depth of the electron wave functions. This change in penetration depth effectively modulates the difference in reflection phases between spin-up and spin-down electrons. Since the exchange coupling energy is dependent on reflectivity, V_{MTJ} can induce transition between FM and AFM coupling in the SAF free layer, as illustrated in Fig. 1.6(d).

From the data reported in [18], FM_1 and FM_2 favors FM coupling when the E-field generated from V_{MTJ} is less than -0.25 V/nm, otherwise AFM coupling is favored.

1.3 Unconventional Computing Schemes

In the past two decades, several unconventional computing schemes and/or architectures have been investigated to overcome the challenges in performing neuromorphic functions. Many intriguing solutions have been studied such as oscillator-based logic, reservoir computing, and Hopfield networks of weakly coupled oscillators. In my dissertation, the three methods I focus on are stochastic computing, stochastic oscillator-based computing, and logic-in-memory computation.

The rest of this section is organized as follows. In section 1.3.1, I will provide an overview of neuromorphic computing in order to illustrate the motivations for studying unconventional computing schemes for future applications. In section 1.3.2, I will explain how logic operations can be performed directly within the memory cell in the computational random access memory (CRAM) architecture. The last three sections explain different types of methods for probabilistic computation and how MTJs can be utilized for each scheme.

1.3.1 Neuromorphic Computing

Successful realization of artificial intelligence (AI) requires real-time analysis of large data sets. However, modern computers based on the Von-Neumann architecture suffer from large circuit area, energy consumption, and speed costs when processing data sets at the scale needed for AI. A promising solution to meet the demands of AI is neuromorphic (NM) computing. NM computing takes inspiration from biological nervous systems to

perform a variety of cognitive tasks such as classification, decision making, and optimization.

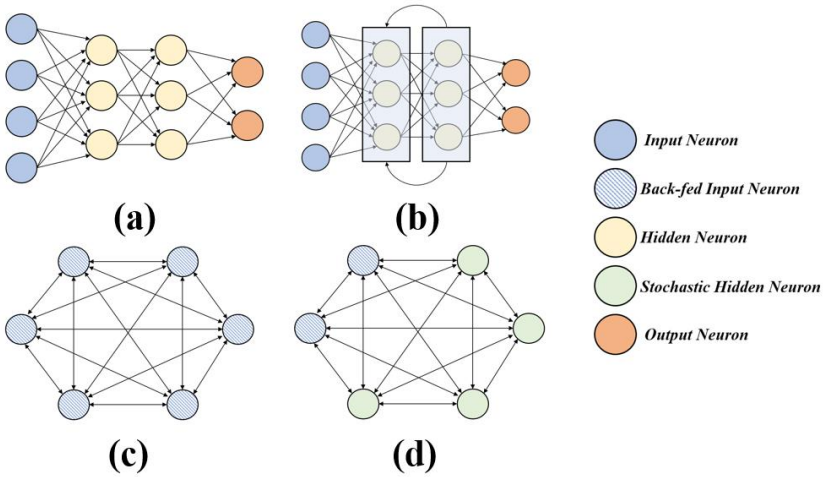


Fig. 1.7. Generalized configuration of various neural networks including **(a)** feed-forward networks, **(b)** recurrent networks, **(c)** Hopfield networks, and **(d)** Boltzmann machines.

The most common algorithm in NM computing is the neural network (NN), examples of which are shown in Fig. 1.7. NNs consist of layers of non-linear processing units called neurons interconnected by adaptive memory elements called synapses. The connection strength between two neurons is established by their synaptic weight (W). Individual tuning of W between each pair of connected neurons in a NN allows for machine learning. Supervised learning refers to learning processes where weights are tuned by an external operator who know the desired output for a given input whereas unsupervised learning refers to learning processes where the weights evolve autonomously. Fig. 1.7(a) is an example of a feed-forward NN, however, it should be noted that many other types of NNs have been proposed such as recurrent NNs, Hopfield networks, and Boltzmann machines, which are shown in Fig. 1.7(b) – (d).

The most common applications of NNs are those involving recognition, classification, or feature extraction of images. Figure 1.8 illustrates how feed-forward NNs are implemented for these tasks. The values applied to the input neurons represent the pixel intensities within the image. The output of each neuron in the input layer is passed to each neuron in the hidden layer, which are then passed to each neuron in the output layer. The dimensions of each layer of neurons from the input to the output layers are progressively reduced to transform large amounts of clustered data into high level information. In the example shown in Fig. 1.8, a data set containing information on every single image pixel is reduced to a dog. The non-linearity feature of the neurons along with proper tuning of synaptic weights through learning allows for extraneous information from the input data to be filtered.

Fig. 1.8. Implementation of neural networks for feature extraction (images extracted from [19])

Unfortunately, the performance capabilities NM systems have several constraints when mapped onto existing processors, primarily due to two limiting factors. The first factor is network size. The human brain contains ~10^{11} neurons and ~10^{15} synapses [20]. The size and complexity of the networks needed to emulate the human brain entirely is so immense that not even the largest supercomputers can accurately simulate them [20]. Modern CMOS devices can be used to build small-scale NNs to perform simple NM tasks. However, a large number of transistors are required to imitate each neuronal and synaptic function, thus limiting the prospects of large-scale CMOS-based NNs. To reduce are costs from

synaptic functions, learning algorithms are currently run through software. However, results in NM systems sacrificing some of their best qualities such as speed, low energy consumption and defect tolerance as well as restricting the learning capabilities to supervised algorithms. Therefore, the performance capabilities of NM systems will be maximized through on-chip learning methods. For these NM systems to be realized, alternatives to CMOS devices should be explored which are capable of emulating both neurons and synapses at the nano-scale.

The second limiting factor can be broadly described as computation models, which includes computer architecture, data encoding schemes, and device functionality. In each of these categories, modern computing models are not compatible with ideal NM models. This is because NM models take inspiration from biological nervous systems, which operate differently to modern computers in many fundamental ways. One is that modern computers store information directly in a memory array, which is physically separated from the processing circuits. Data is continuously fetched between the memory and processing units during the computation process which leads to significant power loss and loss of computation speed due to bottlenecks in task scheduling. Alternatively, memory in NM systems is stored in the synaptic weights, which directly contribute to data processing between neurons, thus making the processing and memory units completely embedded.

Another fundamental difference is in the physical representation of data transmitted between neurons and the encoding technique for interpreting the data. Modern encoding schemes are based on arrays of binary bits, which are processed sequentially, then interpreted through a compiler. Binary encoding in large-scale NNs is very inefficient since each arithmetic operation at each neuron requires multiple layers of Boolean logic gates

and the number of these gates (and therefore computation steps) increase with both network size and resolution of data. On the other hand, the human brain encodes information through a temporal sequence of electrical pulses, or spiking signals transmitted between neurons, as illustrated in Fig. 1.9(a) – (b). These spiking signals are decoded using a method called neural coding, which can implement either rate coding or temporal coding schemes. In rate coding (Fig. 1.9(a)), the spiking signals are assumed to be probabilistic, therefore, signals are averaged over a given measurement time, whereas temporal coding (Fig. 1.9(b)) considers the precise timing between the emitted spikes. In some respects, rate coding more accurately emulates the human brain since neuroscience data indicates that neural signals operate at the thermal limit [21], and therefore, consume very little energy, but are inherently stochastic. Despite the probabilistic nature of these neuron signals, computations in human brains are reliable. A common interpretation of this is that the brain compensates for variability of the individual components through redundancy. However, neuroscience studies also showed that groups of neurons can be modelled as non-linear oscillators whose frequency is dependent on the incoming signal [22-23]. From this perspective, information is encoded through synchronization states of oscillator neural assemblies [24-25], meaning that temporal information between neuron signals is also an important factor for decoding and encoding data. Therefore, the encoding scheme in the human brain is likely some combination of temporal and rate coding schemes.

The third key difference between modern computers and biological nervous systems is device functionality. Based on the data encoding method used for biological nervous systems (described in the previous paragraph), each neuronal unit should (1) be able to operate at low power, (2) have non-linear transfer properties, (3) have intrinsic stochasticity

which is tunable, and (4) contain some degree of signal coupling between neurons (for oscillator neural assemblies). Synaptic units should (1) store analog information, (2) have easy and efficient methods for tuning analog information, and (3) ability to control coupling strength between signals. Some of these neuronal and synaptic properties can be emulated using CMOS devices, however, most of them are very inefficient to implement since they require excessive number of transistors. For example, CMOS based circuits can implement non-linear activation functions and produce stochasticity through true random number generation. However, there are significant costs in both circuit area and power consumption to perform both of these tasks. The short-comings that are present in CMOS-based NM systems means that both post-CMOS technology and novel computation models (including architecture, data encoding schemes, and device functionality) need to be developed for the full potential of NM computing to be realized.

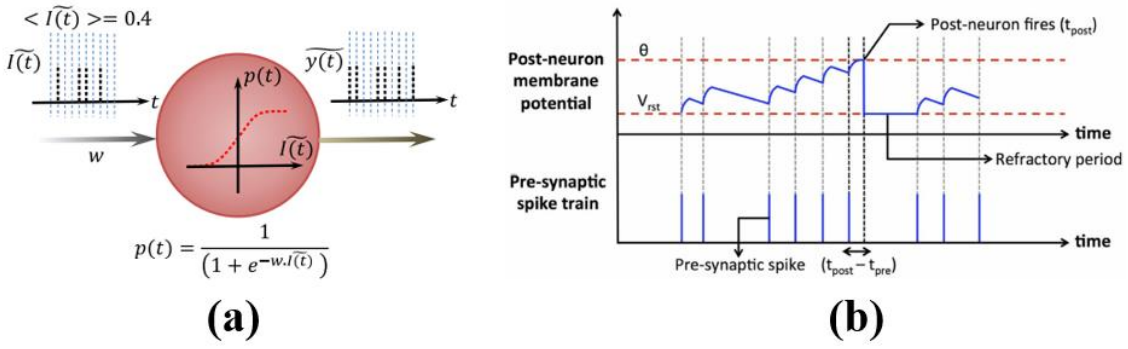


Fig. 1.9. Data encoding with temporal sequences of electrical spikes via (a) rate coding and (b) temporal coding. Images extracted from [26] and [27], respectively.

1.3.2. Computational Random Access Memory

The separation between memory and computing units in modern general purpose processors has detrimental effects on their performance in emerging neuromorphic applications in terms of speed, energy efficiency, and area. To overcome these challenges, in-memory computing methods have been proposed, where computation tasks are performed entirely within the memory array. In recent years, an MTJ-based solution has been proposed called computational random access memory (CRAM) [28-32].

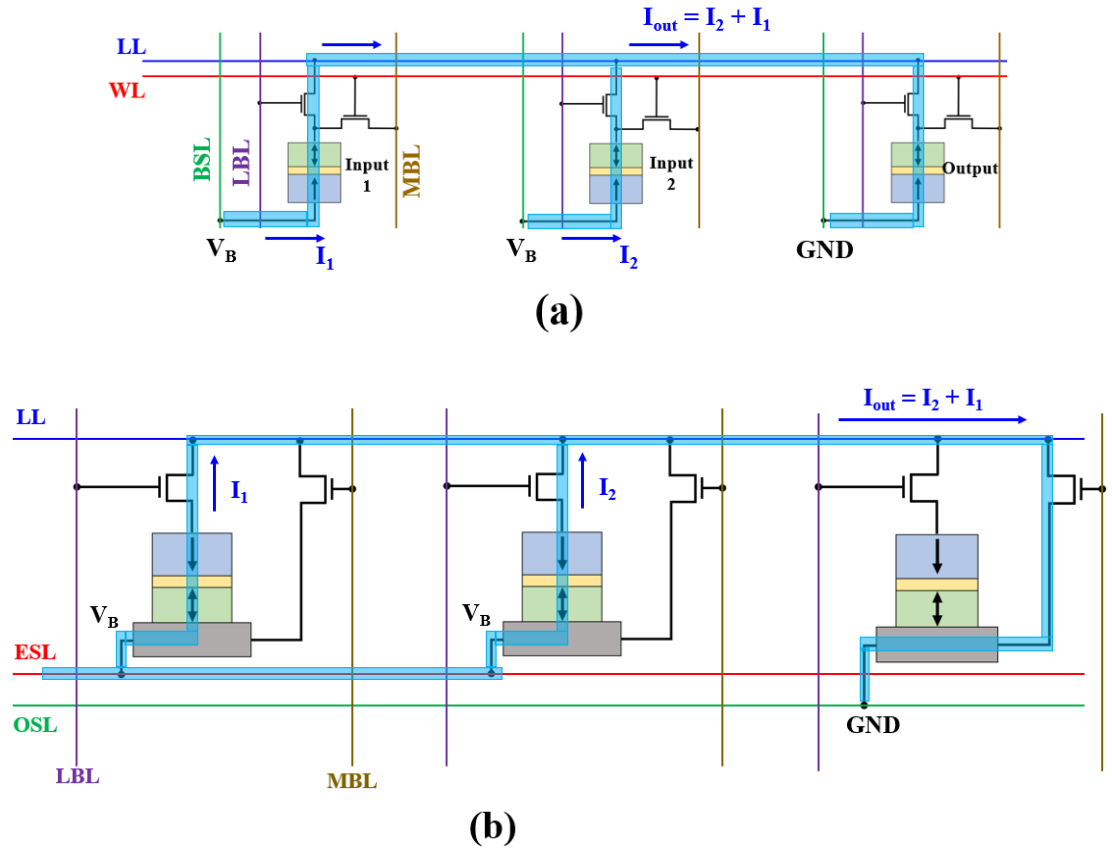


Fig. 1.10. General schematic of **(a)** STT-CRAM and **(b)** SOT-CRAM.

General CRAM structures for STT switching and SOT switching designs are shown in Fig. 1.10(a) and 1.10(b), respectively. Both use a 2T1MTJ bit-cell with separate memory and logic lines. As with conventional STT-MRAM or SOT-MRAM, each bit can be read

during the memory mode. To understand the logic mode, we must consider a system with 3 MTJs which is shown in Fig. 1.11(a) – (c). Here, two of the MTJs are memory elements function as input MTJs and the third MTJ provides the output value of the logic function. During logic mode, a voltage pulse, V_{BSL}, is applied to the input MTJs which results in a current (I_{OUT}) applied to the output MTJ through the logic line whose magnitude is dependent on the resistance states of the two input MTJs. If $I_{OUT} > I_C$ (critical threshold current required to switch the output MTJ), then the output MTJ will switch states. Any logic function can be performed with this process if V_{BSL} and the preset state of the output MTJ are set properly, the parameters of which are shown in Table 1.1.

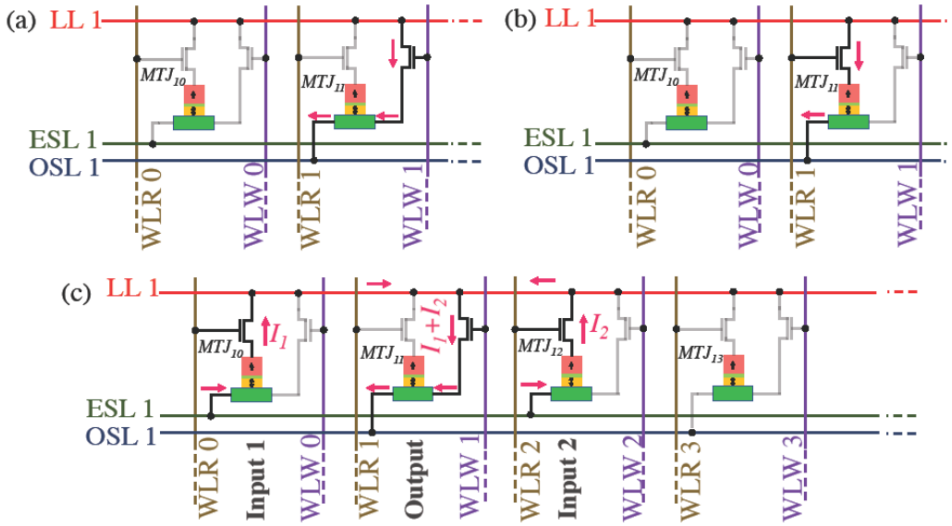


Fig. 1.11. Current flow during **(a)** memory write operation, **(b)** memory read operation, and **(c)** logic mode (image extracted from [30]).

NM computing systems ideally have completely embedded processing and memory units; however, this has not been accomplished with conventional CMOS technology. Therefore, the true in-memory computation capability of CRAM is a promising feature for its prospects in NM computing. M. Zabihi et al showed how the CRAM architecture can implement a NM inference engine for handwritten digit recognition [29]. When compared

to a near-memory processing (NMP) system [33], CRAM achieved the digit recognition task with 1400X faster execution speeds and 40X less energy. These advantages of CRAM over NMP were attributed to high locality of operations and large amounts of parallelism in CRAM.

Table 1.1: Parameters for I_{OUT} and preset state of output MTJ for CRAM Boolean logic functions

Gate	Output Preset	Parameters for I_{OUT} (determines V_{BSL})
Not	0	$I_{OUT} > I_C$ when R_{IN} is low
Buffer	1	
AND	1	$I_{OUT} > I_C$ when either $R_{IN,1}$ or $R_{IN,2}$ (or both) is low
NAND	0	
OR	1	$I_{OUT} > I_C$ when both $R_{IN,1}$ and $R_{IN,2}$ are low
NOR	0	
MAJ3	1	$I_{OUT} > I_C$ when 2 out of 3 R_{IN} values are low
$\overline{MAJ3}$	0	
MAJ5	1	$I_{OUT} > I_C$ when 3 out of 5 R_{IN} values are low
$\overline{MAJ5}$	0	

1.3.3. Stochastic bit-stream generation in stochastic computing

Stochastic computing (SC) is a scheme where numbers are represented as a stream of random binary bits [34-37]. The numbers are interpreted as a probability of individual bits being '1'. In other words, given a bit-stream s with N stochastic bits, the number interpreted by that bit-stream (S) is represented as the percentage of '1' bits in s. Note that even though each element within the bit-stream is binary, the numbers interpreted in the bit-stream are analog, which we will refer to as stochastic values.

There are two key advantages to SC over traditional deterministic models. The first is that the stochastic values are not dependent on position of elements in the bit-stream, therefore provides a high degree of error tolerance [38-42]. The second is the low-cost implementation of arithmetic operations [43-44]. For example, multiplication of stochastic bit-streams A and B can be done using a single AND gate for unipolar representation (between 0 and 1) or a single XNOR gate for bipolar representation (between -1 and +1). Figures 1.12(a) – (f) shows several examples of arithmetic functions being implemented by single logic gates in SC. Note that scaled addition of bit-streams A and B can be done using either a MAJ3 gate or a 2x1 MUX by introducing an additional input bit-stream S, which has a constant and pre-determined probability. When using MAJ3 logic, S needs to be set to 0.5, however, when using a 2x1 MUX, S can be any value.

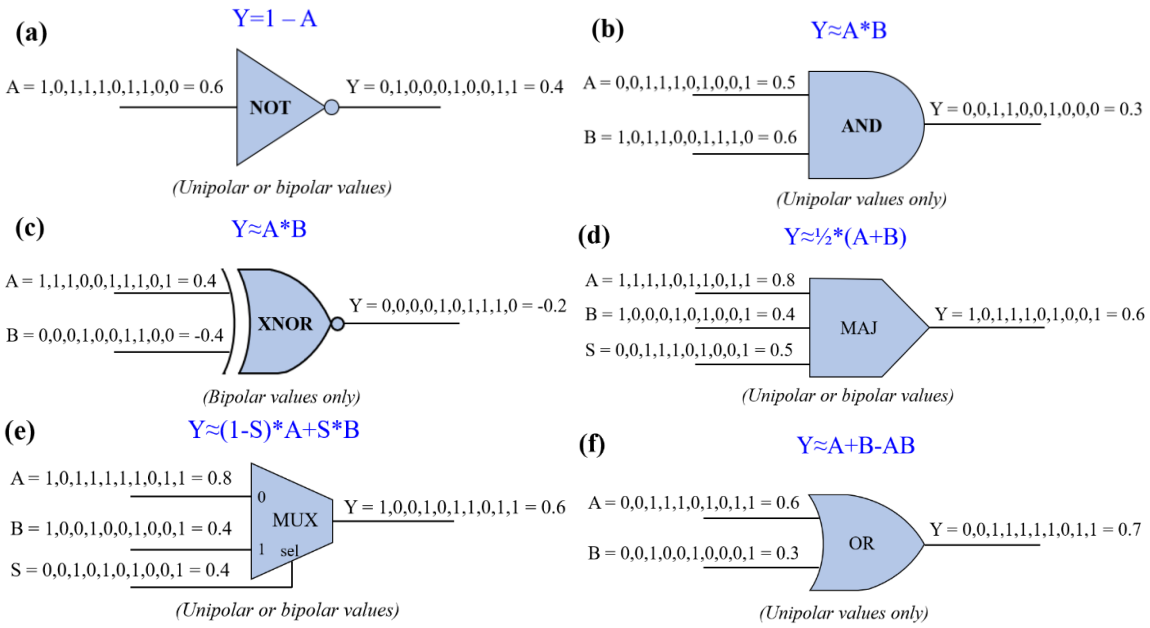


Fig. 1.12. Logic gates used for efficient implementation of various arithmetic functions including **(a)** Y = 1-A (NOT gate), **(b)** unipolar multiplication (AND gate), **(c)** bipolar multiplication (XNOR gate), **(d)** addition scaled by 1/2 (MAJ gate), **(e)** addition scaled by a factor of S (2x1 MUX), and **(f)** Y = A+B-AB (OR gate).

The major drawback to SC is the reduction of precision in numeric calculations. This can only be avoided by increasing the length of the bit-streams; however, this comes at the cost of reduced computation speed. Therefore, for tasks which require solving complex numeric problems, stochastic computing not a very attractive solution. However, for many neuromorphic computing tasks, high mathematical precision is not necessary [42].

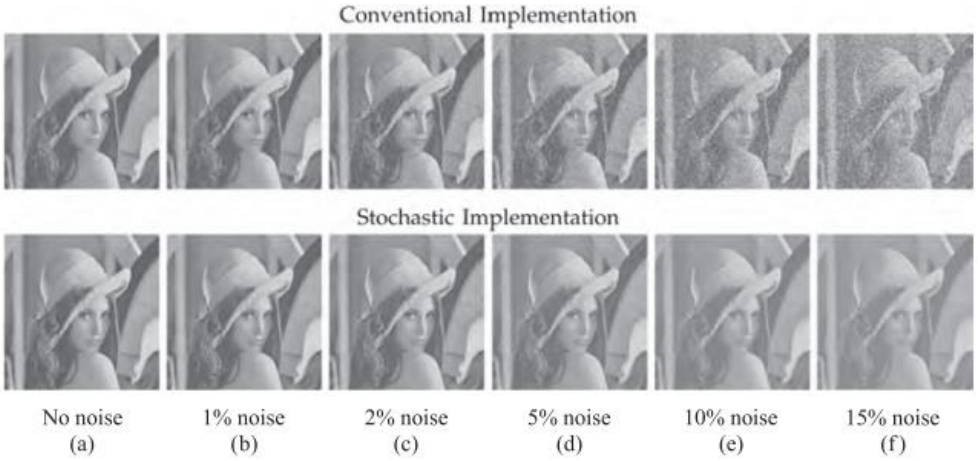


Fig. 1.13. Comparison between conventional and stochastic implementation for image processing at **(a)** no noise, **(b)** 1%, **(c)** 2%, **(d)** 5%, **(e)** 10%, and **(f)** 15% noise injection (image extracted from Ref. [38]).

The error resilience and fault tolerance inherent in SC makes it a strong candidate for tasks such as recognition, edge detection, and classification [38-39, 41, 44-45]. To illustrate this point, consider the example shown in Fig. 1.13 (extracted from [38]). SC-based NN were able to effectively process the image at 15% noise injection whereas the performance of the conventional computing based NN decreased noticeably at just 5% noise injection. This is a very important feature for classification tasks since the classification accuracy will not be influenced by slight modifications to objects in input images or imperfections. In the example provided, along with several other NM applications, approximations are

used to filter non-critical information from large datasets, therefore, SC is a strong solution for these types of tasks.

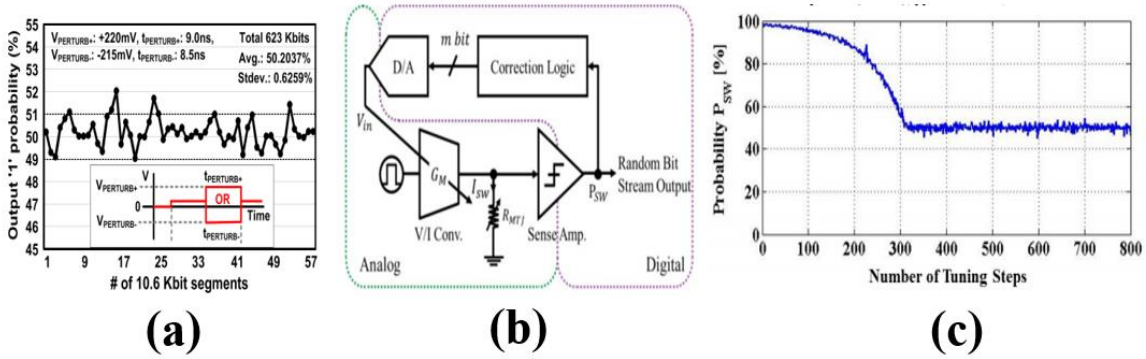


Fig. 1.14. Examples of MTJ-based approaches for generating random numbers centered at 0.5 including **(a)** conditional perturb and real-time output probability tracking (image extracted from Ref. [46]) and **(b,c)** digitally controlled probability-locked loop (images extracted from Ref. [47]).

Several studies have explored the potential of generating true random numbers using a single MTJ, as illustrated in the examples shown in Fig. 1.14(a) – (c) [46-47]. By using a synchronous approach, true random numbers can be produced by simply applying a series of voltage or currents pulses to the MTJ and setting the amplitude and pulse width to ensure a 50% switching probability. This capability solves one of the key short-comings of CMOS in probabilistic computing which is the large area cost in random number generation. This method of generating random numbers has been proposed as a method of generating stochastic bit-streams and optimizing learning algorithms in SC-based artificial NNs [48]. However, the prospects of MTJs in SC can be expanded further since their switching probabilities can be tuned between 0 to 100%. By utilizing their full tuning range, a single MTJ can function as an analog-to-stochastic bit converter [49-53] for several applications such as vision chips [52] and SC-based Bayesian networks [53], shown in Fig. 1.15(a) –

(b). Furthermore, when considering the influence of multiple biases simultaneously, a single MTJ could perform both multiplication and addition in stochastic computing [54].

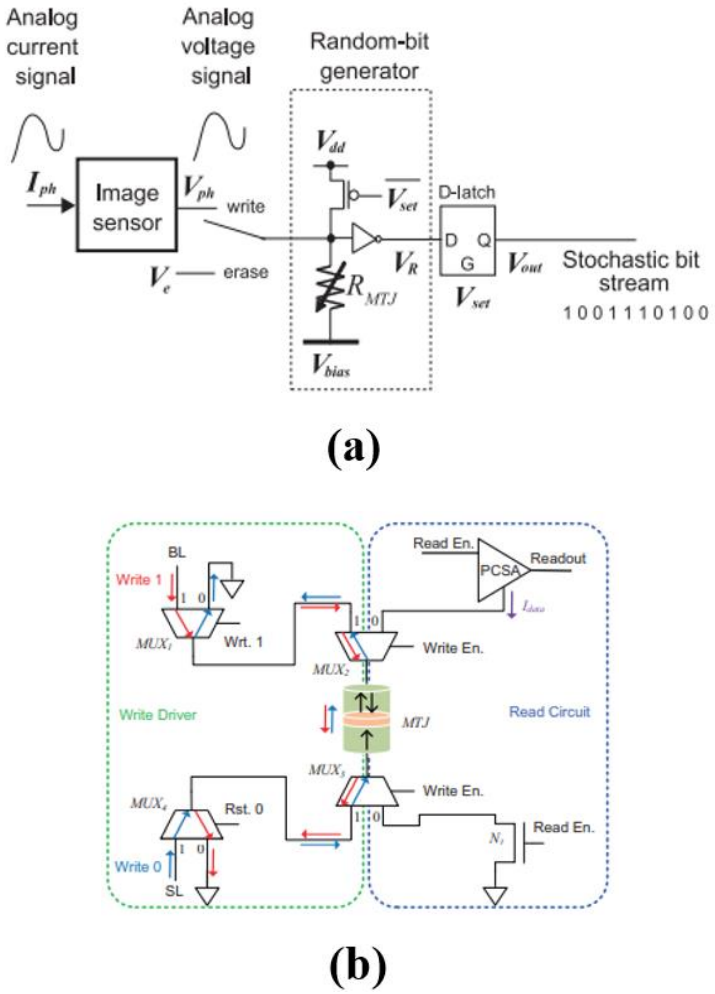


Fig. 1.15. Examples of MTJ-based analog-to-stochastic bit-stream conversion circuits for **(a)** vision chips (image extracted from Ref. [52]) and **(b)** Bayesian inference networks (image extracted from Ref. [53]).

1.3.4. MTJs in neural spiking networks

Tunable stochasticity of MTJs also has numerous prospects in spiking NNs (SNN) based on probabilistic neural coding models. One of these prospects is to use MTJs to build stochastic spiking neurons for hardware realization of neural rate coding schemes, an

example is shown in Fig. 1.16(a) – (c). Several studies have implemented synchronous methods to generate stochastic neural spiking signals [26, 55-57], however, they can also be generated using asynchronous methods [57-59].

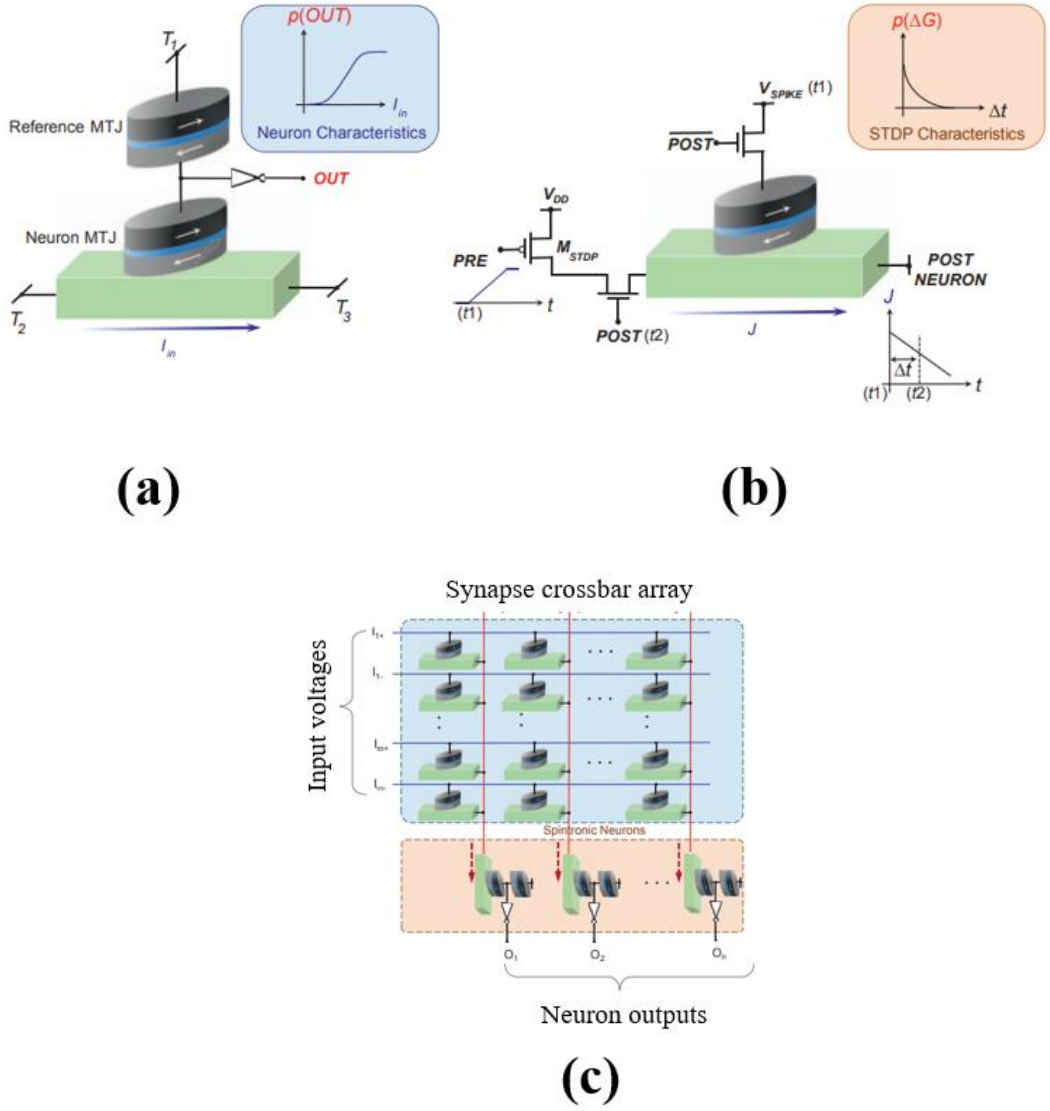


Fig. 1.16. Proposed MTJ-based designs of **(a)** a stochastic neuron and **(b)** a stochastic synapse in spiking neural networks. **(c)** Architecture for a spintronic-based core computing kernel. Images were extracted from [62].

Furthermore, temporal encoding can be realized in stochastic SNNs which use single MTJs as stochastic binary synapses [60-62]. In these circuits, synaptic weights are stored

as the switching probability of the MTJ. Then a feedback system is employed so that the switching probability of the MTJ is based on the temporal correlation between neural spikes at different layers. These stochastic binary synapses can be used to implement on-chip stochastic learning algorithms [27] as well as unsupervised learning and probabilistic inference [62]. Temporal encoding can also be realized using asynchronous random signals from MTJs by applying a dual-biasing scheme [63], which was developed by our group and will be discussed in more detail in chapter 3.

1.3.5. MTJs as probabilistic bits in Boltzmann machines

A Boltzmann machine (BM) is a type of recurrent stochastic NN with binary nodes. The configuration of states for all binary nodes, ν, has energy $E(\nu)$, defined in Eq. 1.9 [64], where s_i and s_j are the binary states of nodes i and j, b_i is the bias, or intrinsic excitability, of node i, and W_{ij} is the connection weight between nodes i and j. The probability of node i being in binary state 1, $P(s_i)$, is defined in Eq. 1.10, where σ is the logistical sigmoid function. BMs will reach a Boltzmann distribution when given enough time, where the probability of being in state configuration ν, $P(\nu)$, is defined in Eq. 1.10, where u can be any possible state of the system. Equation 1.10 shows that the system has the highest probability of being in the state configuration with the lowest energy. A subcategory of BMs is called the restrictive Boltzmann machine (RBM), whose architecture is restricted to two fully-connected, non-recurrent layers [65]. One is called the visible layer, whose outputs are clamped to 0 or 1 when provided salient inputs, and the second layer is called the hidden layer, where associations between input vectors can be learned. By enforcing conditional independence of bother layer, learning process can be enhanced, thus making RBMs a more attractive solution for machine learning applications than unrestricted BMs.

$$E(v) = -\sum_i s_i^v b_i - \sum_{i<j} s_i^v s_j^v W_{ij} \tag{1.9}$$

$$P(s_i) = \sigma\left(b_i + \sum_j s_j W_{ij}\right) \quad \rightarrow \quad P(v) = \frac{e^{-E(v)}}{\sum_u e^{-E(u)}} \tag{1.10}$$

In recent years, the probabilistic bit (p-bit) has been introduced as a potential hardware solution for primitive computing elements in BMs. P-bits are unstable, stochastic units which fluctuate randomly between bipolar states -1 and 1. The generic structure for a p-bit is shown in Fig. 1.17(a), which can be constructed using any 3-terminal random signal generator whose randomness can be tuned. The bipolar output state of p-bit(i), m_i, is controlled by input I_i (generic term with arbitrary units) as defined in Eq. 1.11, where rand(-1,+1) is a random number uniformly distributed between -1 and +1. Note that m_i is stochastic, therefore, I_i does not control the state of mi at any given instant, but rather the time-averaged state of m_i, or $\langle m_i \rangle$, as illustrated in Fig. 1.17(b). The blue line in this figure shows m_i over a single sweep of I_i from -4 to +4, which shows that mi = -1 when $I_i < -3$ and $m_i = +1$ when $I_i > 3$. However, as I_i approaches -2, mi occasionally switches to +1, then m_i fluctuates evenly between +1 and -1 for I_i near 0, and then as I_i approaches +2, m_i = +1 with the occasional switch to -1. This is further illustrated in Fig. 1.17(c), which show the time-domain fluctuations of m_i over 100 computation steps (arbitrary time units). The orange line in Fig. 1.17(b) represents $\langle m_i \rangle$, which is calculated by averaging m_i over 100 time units at each point, shows that $\langle m_i \rangle = \tanh(I_i)$.

$$m_i = sign[\tanh(I_i) + rand(-1,+1)] \quad \rightarrow \quad \langle m_i \rangle = \tanh(I_i) \tag{1.11}$$

$$I_i = I_0\left(b_i + \sum_j W_{ij} m_j\right) \tag{1.12}$$

In order for p-bit circuits to be suitable computation units for BM hardware, the input of each p-bit should be correlated with the output fluctuations of the other p-bits such that I_i follows the relation defined in Eq. 1.12. In this expression, W_{ij} is the connection strength between p-bits i and j, b_i is a localized bias term, and I_0 sets the global scale for the strength of the interactions. By inserting Eq. 1.12 in Eq. 1.11, then $\langle m_i \rangle$ has the same correlation with b_i, W_{ij}, and m_j as $P(s_i)$ has in Eq. 1.10. Note that the only difference is that $P(s_i)$ follows a sigmoidal curve whereas $\langle m_i \rangle$ follows a hyperbolic tangent curve. This is because $P(s_i)$ has unipolar representation (0 to 1 range) whereas $\langle m_i \rangle$ has bipolar representation (-1 to +1 range).

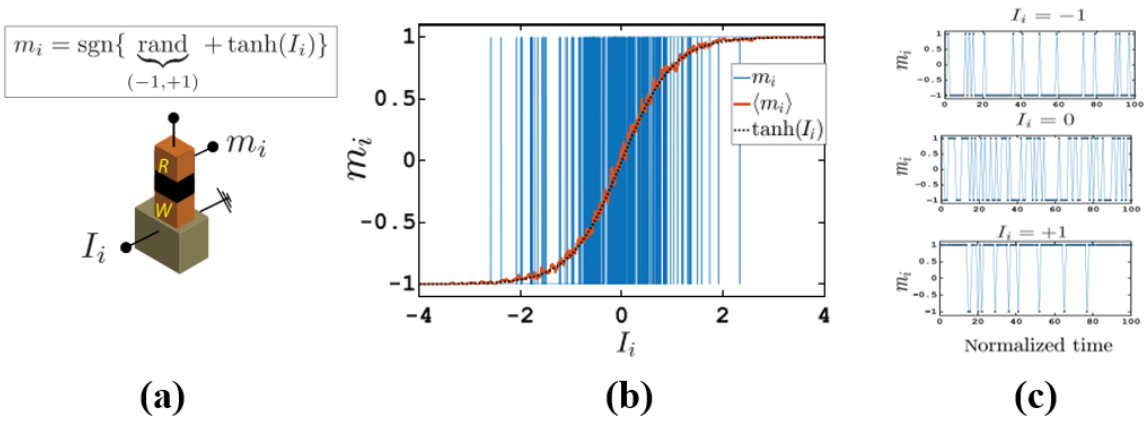


Fig. 1.17. (a) Generic structure of the p-bit, **(b)** ideal transfer curve of the p-bit, which is based on the time-averaged magnetization state of the nano-magnet ($\langle m_i \rangle$), and **(c)** time-domain plots showing $\langle m_i \rangle$ fluctuations at various inputs (normalized units). (Images were extracted from [64]).

Figure 1.18(a) provides an example of a small network of 3 interconnected p-bits for implementation of AND logic. The connection strengths between the two input p-bits (A and B) and between each input p-bit and the output p-bit (C) are W_0 and $2W_0$, respectively. Similarly, the biases on A, B, and C are b0, b_0, and $2b_0$, respectively. These values were

chosen according to interconnection matrix to implement AND logic described in [64]. The results for uncorrelated and correlated p-bits is shown in Fig. 1.18(b).

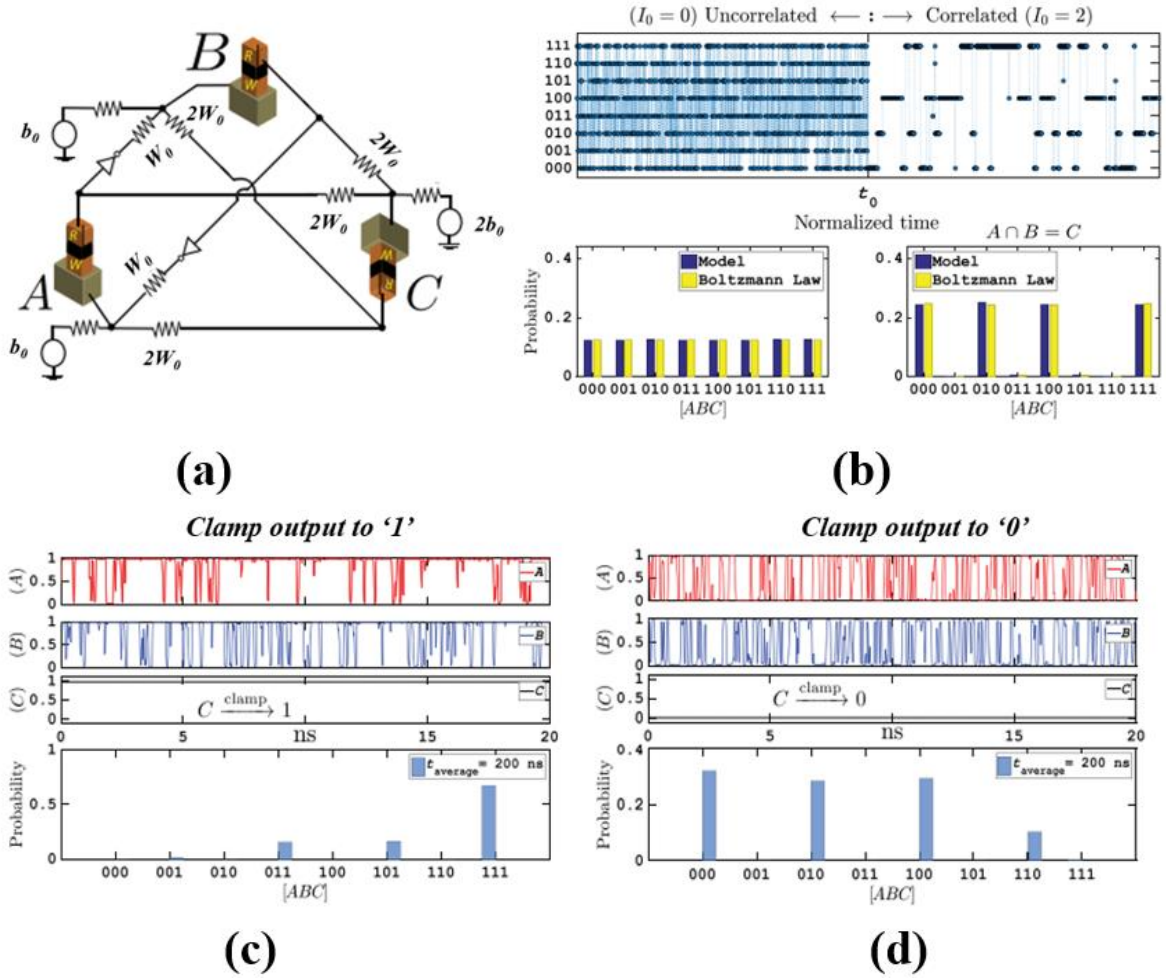


Fig. 1.18. (a) Circuit diagram of an invertible AND gate with a network of 3 interconnected p-bits. **(b)** Output frequency for all combinations of [A B C], which shows successful implementation of an AND gate since the dominant cases follow proper AND logic criteria. Demonstration of invertible AND logic with output bit C clamped to **(c)** logic state '0' and **(d)** logic state '1'. (Images were extracted from [64]).

For uncorrelated p-bits ($I_0 = 0$), the p-bits randomly fluctuate between all 8 possible configurations of [ABC] at equal probabilities. However, for correlated p-bits ($I_0 = 2$), then the fluctuations of the p-bits will only favor configurations that follow AND logic critieria. Therefore, the three p-bits will fluctuate between configurations [000], [100], [010], and

[111] at probabilities of ~25%, and ~0% for all other configurations. These results show that (1) p-bit networks are satisfactory as BM hardware since their signals follow Boltzmann law and (2) p-bit circuits can perform Boolean logic as well as the rare and novel 'invertible' Boolean logic operation. This is illustrated in Fig. 1.18(c) and (d), where the state of C is clamped to either 1 or 0, the the states of the A and B are measured. In the case where C is clamped to 1 (Fig. 1.18(c)), A and B are both 1 for a majority of the measuring time, which satifies AND logic. In the case where C is clamped to 0 (Fig. 1.18(d)), A and B fluctuate between all three cases which satifies AND logic criteria at approximately equal probabilities of ~33%.

It has been shown that combinatorial optimization problems can be mapped onto a network of p-bits with an appropriate problem-specific weight matrix, such that the optimal solution corresponds to the lowest energy solution [66-67]. A well-known example of an optimization problem is the classic N-city traveling salesman problem (TSP). It involves finding the shortest route by which a salesman can visit all cities once starting from a particular one. In recent work, a 5-city TSP was mapped to a network of 16 p-bits and translated into a p-circuit that is simulated using SPICE. Each p-bit has two indices, the first denoting the order in which a city is visited and the second denoting the city. Note that it is assumed that the salesman will be starting and ending the trip at city 0. Through a process called simulated annealing, the correct solution was found, where that the cities should be visited in the order 1-3-2-4, as shown in Fig. 1.19(a). This result was based on four the p-bits with indices (1,1), (2,3), (3,2) and (4,4) having large probabilities of existing in state 1 after simulated annealing. Additionally, RBMs based on p-bit networks have been

proposed for other NM applications such as Bayesian inference, shown in Fig. 1.19(b) [67-68], belief networks [65], and accelerated machine learning [69].

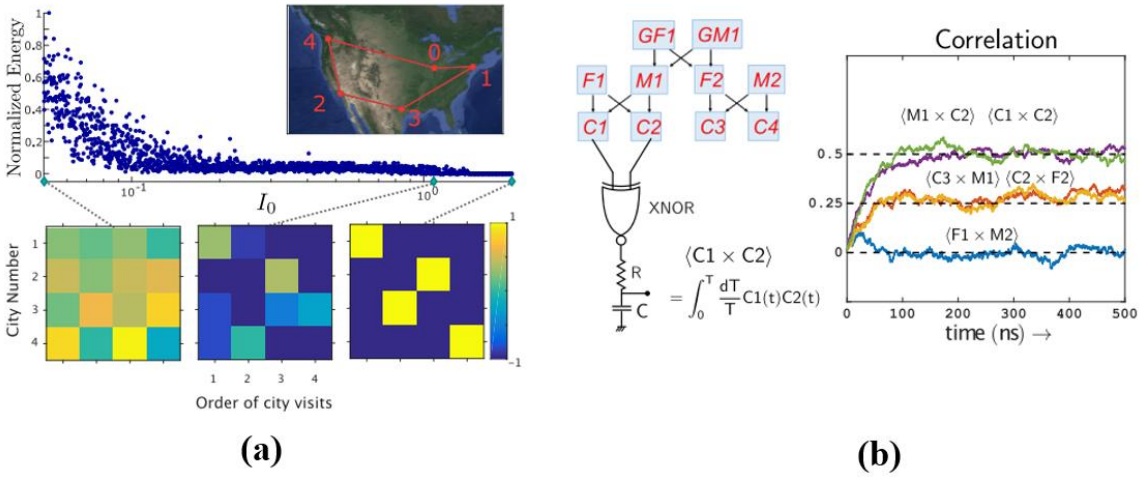


Fig. 1.19. (a) An example of a 5-city travelling salesman problem implemented using a network of 16 p-bits. **(b)** An example of a Bayesian inference network to solve genetic relatedness problems implemented using p-bit circuits. (Images extracted from [67]).

Lastly, MTJs generating tunable asynchronous stochastic signals have promising prospects as elementary components for probabilistic bits, or p-bits [67, 70-75]. The structure of p-bits is similar to STT- or SOT-MRAM cells [70], as shown in Fig. 1.20(a) and Fig. 1.20(b), respectively. The transfer characteristics of p-bits (shown in Fig. 1.17(b)) are defined by the time-averaged output of the p-bit, which can be tuned using either STT [70, 73-74] or SOT biases [70-71].

There are several methods of generating random telegraphic signals in MTJs that are also compatible for p-bit designs. One is to fabricate in-plane MTJs as circular nano-pillars, which would eliminate the influence of shape anisotropy, meaning $\Delta_{IP} \approx 0$. Recent experimental work has demonstrated that MTJs with very low Δ can be realized in circular IP MTJs [70-75]. Other possible methods include (1) using PMA MTJs with free layer

thicknesses engineered so that $H_K^{PMA} \approx 0$, (2) reducing net magnetic moment of the free layer, (3) hard-axis initialization in PMA MTJs, and (4) introducing perpendicular anisotropy to IP MTJs [75-76]. However, these solutions are not as attractive as circular IP MTJs since solution (1) generates much slower switching rates and solutions (2) to (4) require external magnetic fields.

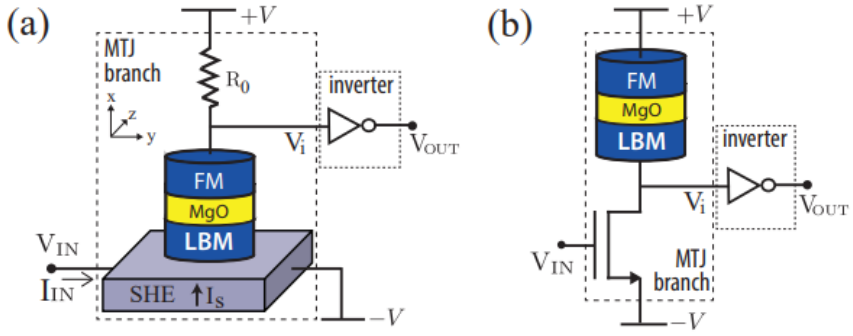


Fig. 1.20. Structure of a p-bit cell using **(a)** SOT and **(b)** STT switching mechanisms. (Images extracted from [67]).

The method that will be investigated in this dissertation generates random telegraphic signals using two external biases which favor opposite states and causes random toggling between the AP- and P-states. This method is referred to as 'dual-biasing' and it offers additional flexibility and tunability compared to single-biasing methods [77-79]. Two novel features in p-bits could possibly be introduced with dual-biasing which are (1) establish a method for on-board corrections to the effects of variations in device dimensions from their nominal ones and (2) design PSL based NM systems composed of p-bits which have neuronal and synaptic weight tuning capabilities embedded in a single device. The application of this method has already been validated by our group when we reported the first experimental demonstration of MTJ-based hardware implementation of PSL [80], as shown in Fig. 1.21(a) – 1.21(c). Dual-biasing is described in further detail in chapter 3.

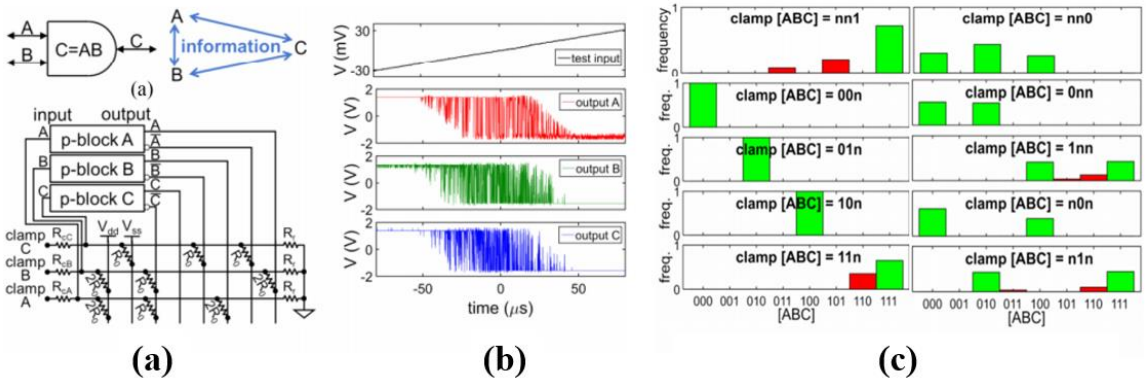


Fig. 1.21. (a) Diagram of information flow and circuit diagram of invertible AND gate **(b)** Waveforms of a ramp input test signal and output A, B, and C of all three p-bit block. **(c)** Relative histogram of [ABC] with various combinations of A, B, and C clamped. These experimental results confirm the simulation results shown in Fig. 1.18(b) – 1.18(d) (Images extracted from [80]).

1.4 Dissertation overview

In this dissertation, I will present my research on magnetic tunnel junctions in new paradigms of novel computing and memory applications. The organization of my dissertation are as follows.

In chapter 2, I will outline the fabrication method used for most of the magnetic tunnel junction devices patterned for my research. Then, I will describe the testing set-up and processes for most of my experiments.

In chapter 3, I will present a method of generating telegraphic switching signals with tunable stochasticity for probabilistic computing applications. This method involves the interplay between the effects of spin transfer torque and an external field. I will also discuss the physical mechanisms behind this dual-biasing method and I will also discuss its effectiveness on devices with varying intrinsic properties.

In chapter 4, I will examine the combination VCEC and SOT switching mechanisms for MRAM cells with ultra-low switching current density. Here, I will discuss the influencing factors of SOT on the coercivity and stray field of p-MTJ devices with SAF free layers as well as the influencing factors of SOT on their VCEC switching properties. Most of this chapter will focus on how the interplay between VCEC and SOT effects to reduce the switching current density for memory applications. However, at the end of this chapter, I will present results which indicate that SOT and VCEC can also be used as a dual-biasing method to generate tunable stochastic switching signals, thus eliminating the need for the external bias field in the method presented in chapter 3.

In chapter 5, I will introduce a method of implementing a stochastic computing process using the CRAM architecture, which I refer to as SC-CRAM. Unlike traditional stochastic computing processes which have separate circuitry for the stochastic bit generators and processing units, SC-CRAM is capable of generating stochastic bits directly within the CRAM cells. Furthermore, I compare the performance of SC-CRAM to conventional CRAM in three neuromorphic applications including local image thresholding, Bayesian inference, and Bayesian belief networks. I demonstrate that SC-CRAM benefits from smaller subarray sizes and similar energy consumption compared to conventional CRAM.

In chapter 6, I demonstrate that magnetic tunnel junctions are exceptionally resilient to the effects of ionizing radiation, making them a promising solution for on-board memory platforms in high radiation applications, such as space exploration. In my experiments, I used synchronous switching probability measurements on 24 magnetic tunnel junctions to measure their thermal stabilities, critical switching voltages, and write energies. I then repeated these measurements after exposure to γ-radiation.

Lastly, my dissertation is concluded in chapter 7.

Chapter 2: Fabrication and Characterization of Magnetic Tunnel Junctions

2.1 Fabrication process

In this section, I will describe the basic process flow of nano-sized MTJ device fabrication. Illustration of the fabrication process of our MTJs is shown in Fig. 2.1(a) – 2.1(d). Our groups process flow consists of four steps: (1) bottom electrode definition, (2) MTJ pillar definition and planarization, (3) top electrode definition, and lastly (4) magnetic thermal annealing.

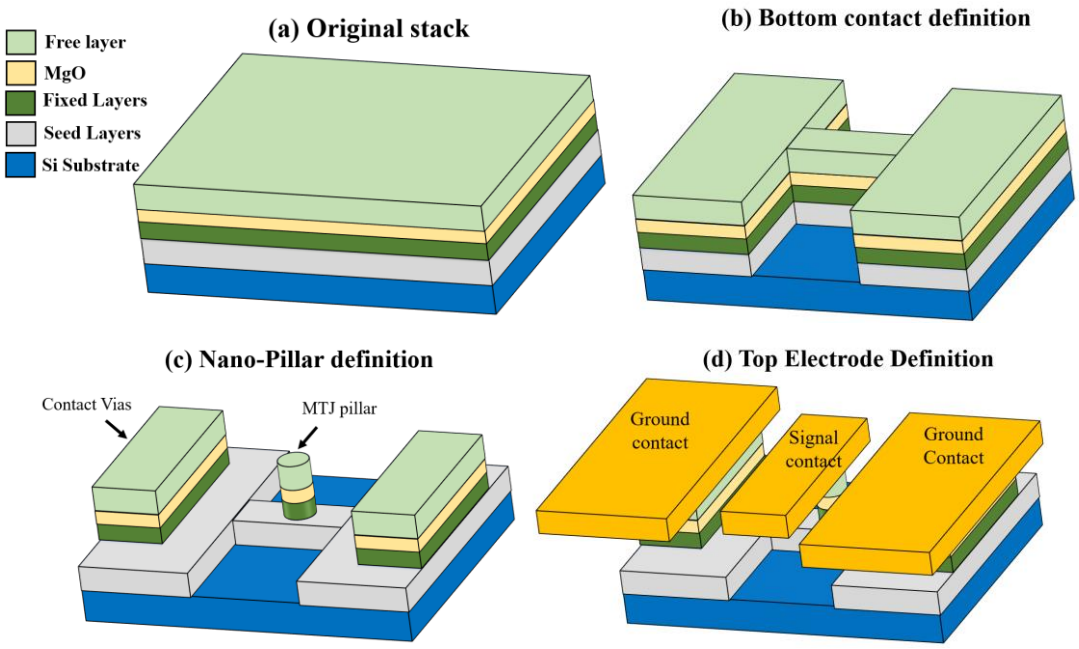


Fig. 2.1. General overview of MTJ fabrication process. **(a)** Original MTJ stack consisting of the ferromagnetic free layer, the MgO tunneling barrier, the ferromagnetic fixed layer, the metallic seed layers, and the Si substrate (top-to-bottom). Definition of the **(b)** bottom electrode, **(c)** MTJ nanopillars and contact Vias, and **(d)** top electrodes. Note that SiO_2 is deposited after patterning the MTJ nanopillars, however, this step is not shown in this image.

2.1.1. Bottom Electrode Definition

Figure 2.1(b) shows the pattern for the bottom electrode for each MTJ fabricated. This pattern is defined using photolithography, the process of which is illustrated in Fig. 2.2. First, an organic, photosensitive polymer called photoresist is applied to the surface of the sample through spin coating. The desired pattern is printed on a photomask prior to exposure, which is aligned with the sample. Then the sample is exposed to intense UV light which causes chemical change in the photoresist in selected regions determined by the pattern on the photomask. Using a special solution called a developer, the photoresist in the affected areas is removed. There are two types of photoresist: positive and negative. When using positive resist, UV exposure causes the resist to become soluble to the developer, therefore, the photoresist in the exposed areas will be removed. However, UV exposure causes negative resist to become insoluble to the developer, therefore, the photoresist in the unexposed areas will be removed.

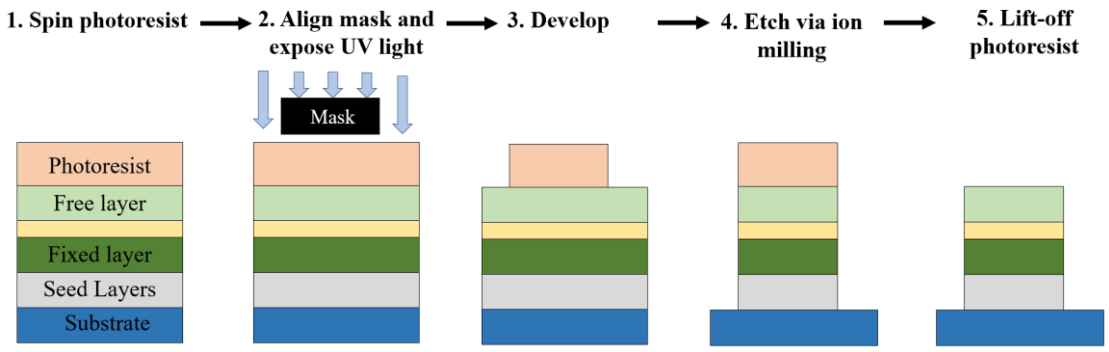


Fig. 2.2. Illustration of the Photolithography process.

For the samples fabricated in this dissertation, positive photoresist was used, meaning that the areas in Fig. 2.1(b) where the MTJ layers and the seed layers are removed were uncovered. More specifically, we used AZ 1512 photoresist, aligned the photomask and

sample using an Ma6 contact aligner and developed using AZ 340 developer diluted in de-ionized water with a ratio of 1 part developer and 5 parts water. After developing, we etched away the exposed regions on the sample all the way to the substrate with ion milling. The remaining photoresist on the sample was lifted off through sonication in a solution called Microposit 1165 remover.

2.1.2. Pillar Definition

The MTJ pillars are defined using electron beam lithography (EBL) which is a process similar to photolithography except the exposure source is an electron beam rather than UV light. The full EBL exposure pattern for our process is shown in Fig. 2.1(c) which includes the elliptical cross section of the MTJ pillar as well as large squares on both sides of the MTJ pillar. The patterns on both sides of the MTJ pillar are called contact Viases and provide an electrical connection from ground probes to the bottom electrode defined in the previous step. In our process, we use Ma-N 2403 e-beam resist (negative resist), exposed using a 100keV Vistec EBL system, and developed using MF-319 developer. The unexposed areas were etched down to the metallic seed layers using ion milling. After etching, SiO_2 is deposited using plasma-enhanced chemical vapor deposition (PECVD). This is done to isolate the bottom electrode and to protect the sidewall of the MTJ pillars from oxidation. The remaining ebeam resist on the sample was lifted off through sonication in a solution called Microposit 1165 remover.

2.1.3. Top Electrode Deposition

Fig. 2.1(d) shows the pattern for the top electrode for each MTJ fabricated, which is defined with photolithography using the same process parameters that were used to define

the bottom electrodes. After developing, the metallic contacts for the top electrodes were deposited using electron beam evaporation. During the electron beam evaporation process, joule heating of a crucible containing the material to be deposited causes the material to evaporate onto the sample. In our process, we deposited 10nm of Ti followed by 100nm of Au. Finally, the remaining photoresist was lifted off through sonication in 1165 remover. Note that the shaded regions in Fig. 2.1(d) are exposed during photolithography, therefore, the only remaining Ti/Au on the sample after lift-off will be in those shaded regions.

2.1.4. Magnetic Thermal Annealing

The last step in the MTJ fabrication process is magnetic thermal annealing. This step is crucial since it maximizes the TMR of the MTJs and allows us to define the easy-axis direction. This process has two components working simultaneously. One is the thermal annealing component where the sample is placed in a high-vacuum chamber (pressure < 10^{-5} mTorr) and heated to a temperature of approximately 325 – 350°C. This allows for the atoms in the devices to diffuse and eliminate many lattice deformities. The second component is the external magnetic field, which is applied before, during, and after thermal annealing. Lattice deformities can create non-uniformities in the overall magnetization of the free and fixed layers. These non-uniformities will weaken the overall saturation magnetization of the MTJ and thus reduce its TMR ratio. By applying an external magnetic field during thermal annealing, this can be avoided since the magnetization direction in the entire sample will align with the field direction. For our samples, the annealing temperatures were between 325 – 350 °C, the magnitude of the external field was 1 Tesla, and the annealing time was 1 hour. A picture of our magnetic thermal annealing stage is shown in Fig. 2.3.

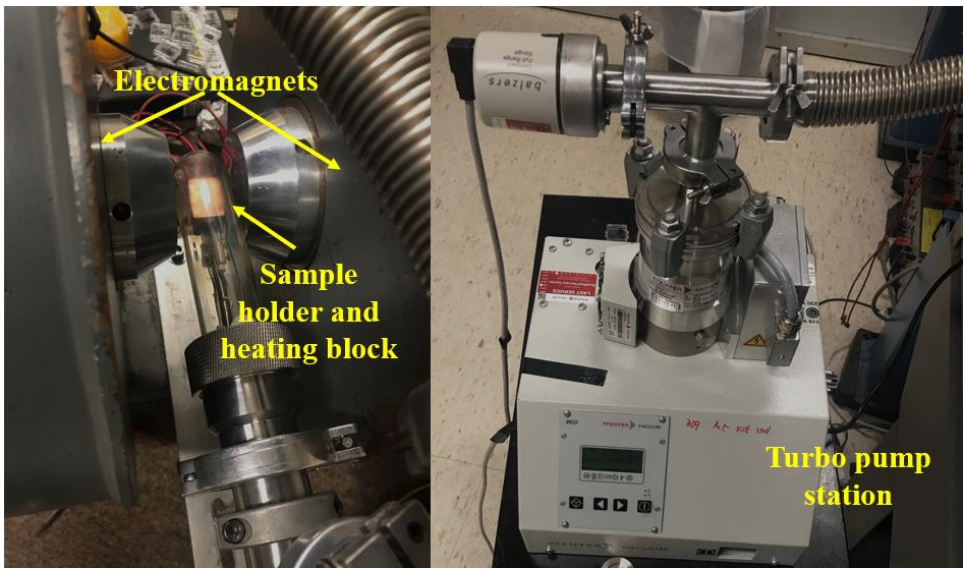


Fig. 2.3. Image of the post-annealing set-up.

The direction of the external field relative to the sample depends on 1) the shape of the MTJs and 2) their desired application. For MTJs with perpendicular anisotropy, only rapid thermal annealing was applied, and magnetic thermal annealing was not used. Our in-plane MTJs were patterned into either elliptical or circular nano-pillars. Note that the influence of shape anisotropy is ideally symmetric, therefore, for circular MTJs, the x-y direction of the external field does not matter as long as there is no z-component in the field direction. For elliptical nano-pillars, the direction of the external field was typically applied along the long-axis of the ellipse to maximize the thermal stability of the MTJs. However, in some of our experiments, MTJs with low thermal stability were desired, therefore, the direction of external field was applied along the short-axis of the ellipse.

2.2 Device Characterization

Most of the measurements carried out can be grouped into one of four categories: Field switching, DC current/voltage switching, synchronous probabilistic switching, and

asynchronous probabilistic switching. For all measurements, we used a Cascade ground-signal-ground (GSG) probe, model ACP40 which has a bandwidth from DC to 40 GHz. The probe layout on the sample is illustrated in Fig. 2.4(a). An external magnetic field was often required for our measurements, which was supplied using a 3-D projected magnet testing stage as shown in Fig. 2.4(b).

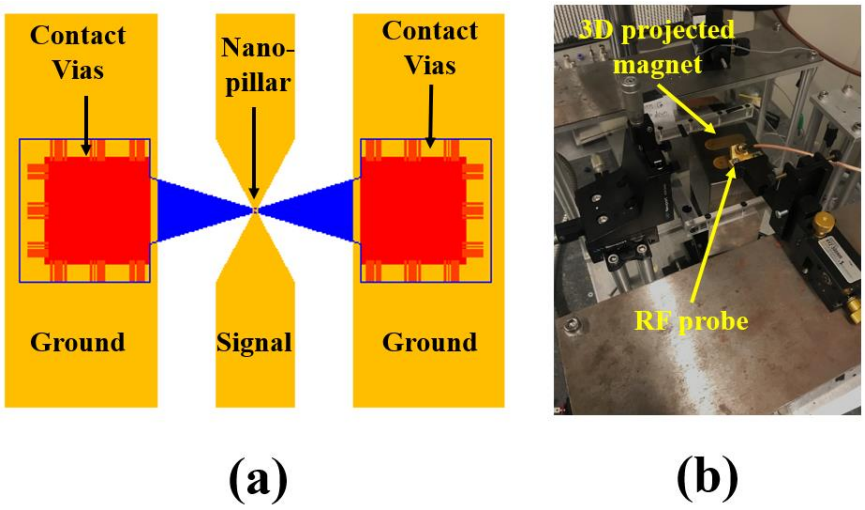


Fig. 2.4. (a) Illustration of ground-signal-ground probe layout. **(b)** Image of 3-D projected magnet stage.

2.2.1 Field switching measurements

Field switching of MTJs states is no longer pursued in modern MRAM designs since the dimensions of a field switching MRAM cell cannot be scaled to sizes below 90nm due to the drastic increase in current required to generate fields with decreasing wire size [7-8]. However, from an experimental perspective, field switching hysteresis (R-H) plots are still useful since they provide key intrinsic properties of the MTJ being tested. These plots are obtained by sweeping an external magnetic field (either in-plane or out-of-plane direction) from negative-to-positive values and then back to negative values (or vice versa)

in field steps of 1 – 5 Oe. At each field step, the resistance of the MTJ is measured using a DC bias current or voltage applied to the MTJ. This DC bias is typically set so that the voltage across the MTJ is ~1 mV, which is sufficiently small enough so that it does not influence the state of the MTJ. However, the DC bias voltage will be higher for experiments where the influence of voltage is studied. These measurements only required two probes connected at the top and bottom terminals of the MTJ. The time interval at each field step was larger than 1 ms, therefore, DC probes can be used in these measurements, however, GSG probes were used when available to avoid the effects of electro-static discharge (ESD) in the MTJs.

An example of an R-H plot on an MTJ with in-plane anisotropy is shown in Fig. 2.5. The MTJ tested in this plot was patterned into an elliptical nanopillar nominal dimensions of 110nm x 45nm and the field was applied along the easy-axis direction. From this R-H plot, we can find P- and AP-state resistances, and therefore TMR ratio, as well as its coercivity and stray field (H_C and H_{STRAY}, respectively). H_{STRAY} represents the field produced by the fixed layer magnetization acting on the free layer and can be measured from the R-H curve by calculating the center field between the P-to-AP and AP-to-P switching fields. In the example in Fig. 2.5, H_{STRAY} = 64 Oe and H_C = 23 Oe. H_C represents the field required for the MTJ to switch states and is measured relative to H_{STRAY}. Measuring H_C not only provides the field required to switch the MTJ, but it also provides information on the switching energy of the MTJ since H_C is proportional to H_K. This is a useful parameter when testing STT and SOT switching.

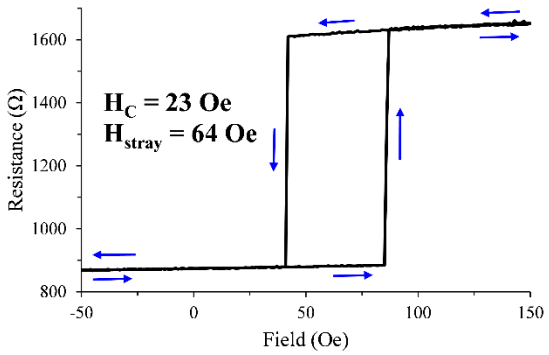


Fig. 2.5. Example of a resistance vs field plot where the data is obtained from an in-plane MTJ patterned into a 110 nm x 45 nm elliptical nanopillar.

2.2.2. Current/voltage switching measurements

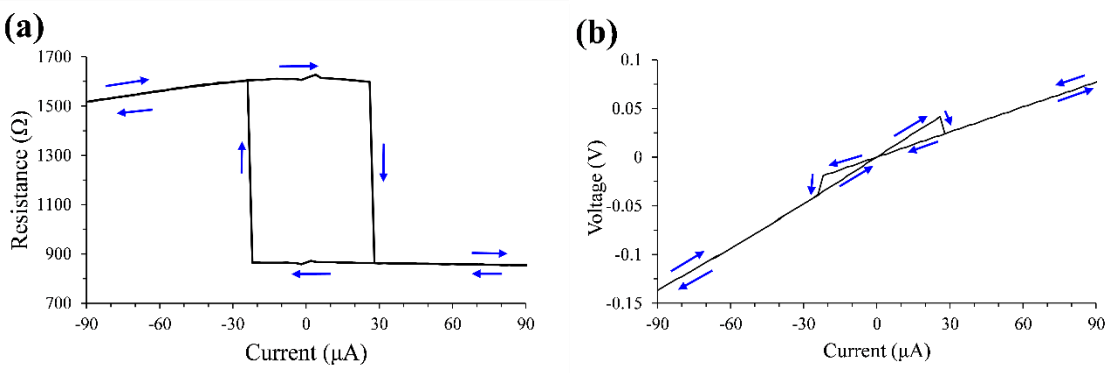


Fig. 2.6. Example of a **(a)** resistance vs current and **(b)** the corresponding voltage vs current data from an in-plane MTJ patterned into a 110 nm x 45 nm elliptical nanopillar.

Current/voltage switching (I-V) measurements are used to demonstrate the STT, SOT, or VCEC switching capabilities of the MTJ being tested. For I-V measurements, a DC current or voltage is swept from negative-to-positive and back. The step size and interval are typically set so that the total measurement time is less than 20 seconds to avoid cumulative effects of voltage breakdown over multiple measurements. Since MTJs have an intrinsic stray field, a constant bias field (either in-plane or out-of-plane) is typically applied throughout the measurement in order to center the I-V switching curve around 0.

Figures 2.6(a) and 2.6(b) are examples of R vs I and V vs I plots, respectively. The data shown in Fig. 2.5 and Fig. 2.6 were obtained from the same device.

2.2.3. Synchronous Probabilistic Switching

Thermal stability of MTJs can be obtained from switching probability distribution measurements through either synchronous or asynchronous methods. Synchronous methods use sequence of voltage (or current) pulses with amplitude V_P and width t_P, then finding the percentage of pulses which caused the MTJ to switch, or P_{SW}, using the process described in [54, 81-82]. For STT switching, voltage (or current) required for STT to exceed the damping torque is intrinsic critical switching voltage, or V_{C0} (or I_{C0} for current pulses). Equation 2.1 expresses V_{C0} and I_{C0} for both IP and PMA MTJs, where α is the damping constant and η is the spin torque efficiency, which is defined in terms of the spin polarization, P_f. H_{eff} is the effective field, expressed in Eq. 2.2, which is defined as the sum of the applied field, H_K, and the demagnetization field. It should be noted that I_{C0} in PMA MTJs is significantly lower than I_{C0} for IP MTJs.

$$V_{C0} = \frac{2e\alpha M_S V R_{MTJ}}{\eta \hbar} H_{eff} \quad \rightarrow \quad \left(I_{C0} = \frac{V_{C0}}{R_{MTJ}} \right) \tag{2.1}$$

$$H_{eff} = \begin{matrix} H + H_K^{IP} + 2\pi M_S & (for\ IP\ MTJs) \\ H + H_K^{PMA} - 4\pi M_S & (for\ PMA\ MTJs) \end{matrix} \tag{2.2}$$

If $V_P \leq 0.8 * V_{C0}$ and $t_P \geq 10$ns, then thermal fluctuations will cause the MTJ to switch with a probability of $P_{SW} = 1 - \exp(-t_P/\tau)$ [82]. Equation 2.3 shows a modified version of the Néel – Arrhenius equation which shows the influence of V_P on τ. The influence of V_P on P_{SW} is shown in switching probability distribution curves, as illustrated in Fig. 2.7(a) and 2.7(b). The critical switching voltage (V_C) is defined as the voltage when $\tau = t_P$ and is

expressed using Eq. 2.4. P_{SW} distribution curves can be used to extrapolate V_C by finding V_P when $P_{SW} = 1 - \exp(-1) \approx 0.63$. By obtaining P_{SW} distribution curves at multiple pulse widths, Δ and V_{C0} can be calculated through by linearly fitting the data for V_C and $\ln(t_p/\tau_0)$, where the y-intercept and slope of the fitted line is equal to V_{C0} and $-V_{C0}/\Delta$, respectively. However, it should be noted that, for PMA MTJs, this method is not accurate when $t_p <$ 300ns since their V_C versus $\ln(t_p/\tau_0)$ plots show significant non-linearities caused by large voltage-controlled magnetic anisotropy (VCMA) effects seen in PMA MTJs [81].

$$\tau = \tau_0 \exp\left(\Delta\left[1 - \frac{V_P}{V_{C0}}\right]\right) \tag{2.3}$$

$$V_C(t_P) = V_{C0}\left(1 - \frac{1}{\Delta}\ln\left(\frac{t_P}{\tau_0}\right)\right) \tag{2.4}$$

Equations 2.3 and 2.4 express the switching time in the thermal agitation regime, which occurs when $t_P \geq 10$ ns. However, there are two other switching regimes which are called the precessional and dynamic reversal regimes. The critical switching voltage for all three switching regimes are illustrated in Fig. 2.7(c). Precessional switching defines magnetization switching when $t_P \leq 1$ ns and in this regime, the magnetization switching is dependent on initial thermal distribution and independent of thermal agitation [83]. In this regime, $\tau \propto (V - V_{C0})^{-1}\ln(\pi/2\theta)$, where θ is the initial angle between the magnetization and the easy axis [82-85]. The switching probability with respect to pulse width is defined in Eq. 2.5 and Eq. 2.6. The dynamic reversal regime occurs when $t_P > 1$ ns and $t_P < 10$ ns. Magnetization switching is dependent on both the initial thermal distribution and thermal agitation. Since dynamic reversal is a combination of the two mechanisms, it is difficult to obtain an explicit formula for the switching probability and switching time [83].

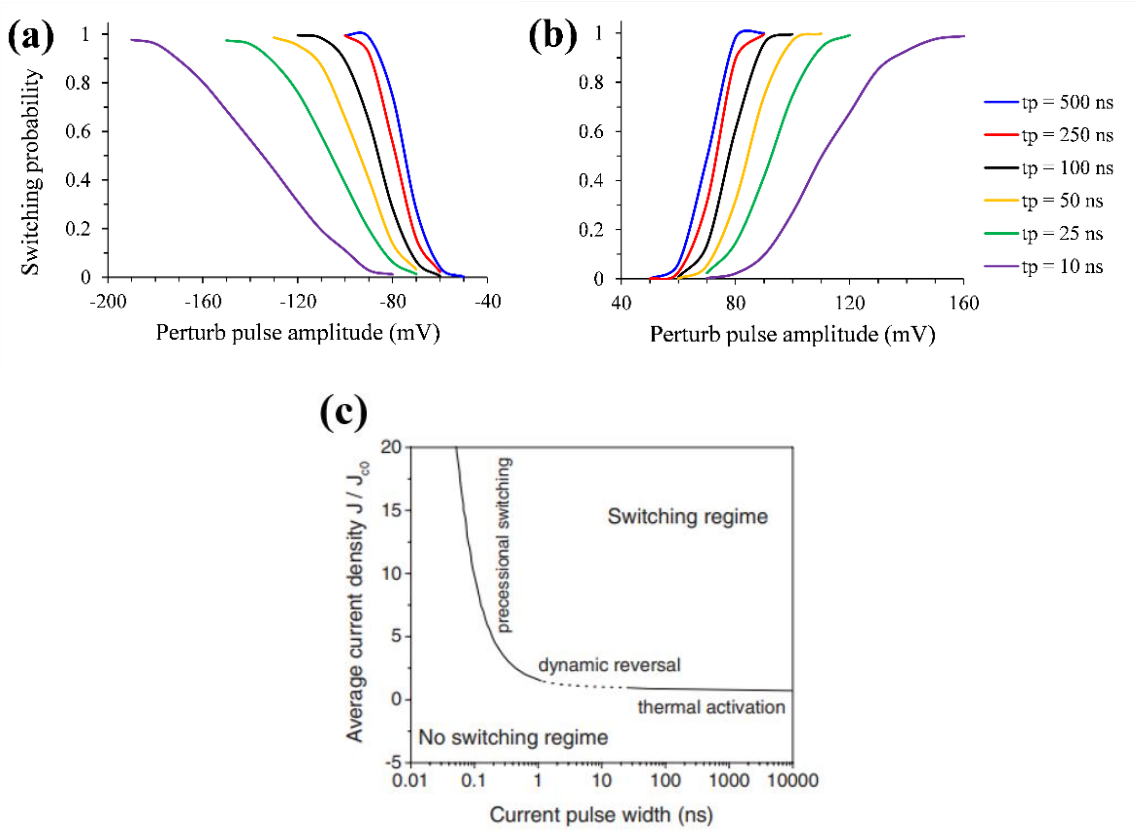


Fig. 2.7. (a) Examples of P_{SW} vs V_P curves for **(a)** P to AP and **(b)** AP to P switching. **(c)** Illustration of the three STT switching modes and their dependence on pulse width (image extracted from [83]).

$$P_{SW}(t_P) \propto \exp\left(\frac{H_K M_S Vol}{2_{k_b T}}(1 - cos^2\,\emptyset)\right)(V - V_{C0})\sin^2\emptyset \tag{2.5}$$

$$\emptyset = \frac{\pi}{2}\exp\left(-\frac{\eta\mu_B}{eM_S t_F}(J - J_{C0})t_P\right) \tag{2.6}$$

Synchronous probabilistic switching measurements were used to determine key MTJ properties such as thermal stability factor (Δ) and intrinsic critical switching voltage (V_{C0}). The experiment process is outlined in Fig. 2.8(a), which consists of a series of synchronized perturb and reset voltage pulses applied to the MTJ being tested, as well as a constant read bias voltage or current (typically 1 mV or 1 μA, respectively). The perturb pulse switches the MTJs resistance state with a probability dependent on the pulse amplitude (V_P) and

pulse width (t_P). After then resistance of the MTJ is measured, a reset pulse is applied to the MTJ whose amplitude and pulse width is set to be large enough to ensure that the MTJ switches back to its original state. The switching probability (P_{SW}) can be calculated for a given V_P and t_P by repeating the cycle N_P times, where P_{SW} = number of times the MTJ switches states/N_P. Full switching probability distribution plots can be obtained with respect to V_P and/or t_P. Note that P_{SW} data can be obtained for both P-to-AP and AP-to-P switching directions by changing the polarity of the perturb and reset pulses.

A schematic for these experiments is shown in Fig. 2.8(b). The perturb pulses were generated using a Picosecond 10,070A pulse generator if t_P < 10ns and an Agilent HP 8100A pulse generator if $t_P \geq$ 10ns. The Agilent HP 8100A also was used to generate the reset pulses. The Keithley 2400 source meter was used to generate the DC read biases and measure the resistance of the MTJ. As with the STT switching measurements, most synchronous probabilistic switching measurements were collected at a fixed bias field to center MTJs STT switching behavior around 0. An RF bias tee (Picosecond model 5342-229) was used to isolate the DC read bias from the perturb and reset pulses.

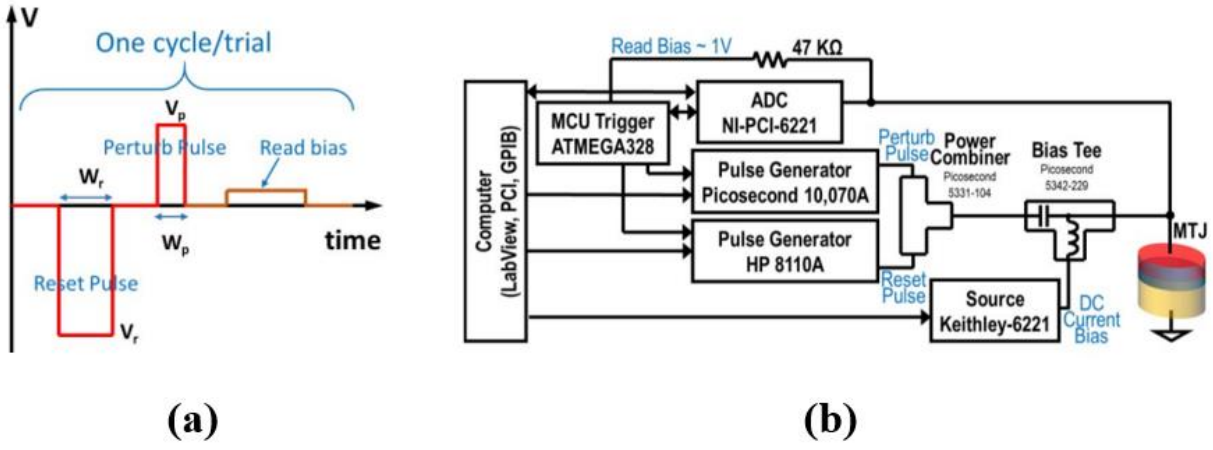


Fig. 2.8. (a) Task scheduling for synchronous probabilistic switching measurements. **(b)** Diagram of testing set-up. Images extracted from [54]

2.2.4. Asynchronous Probabilistic Switching

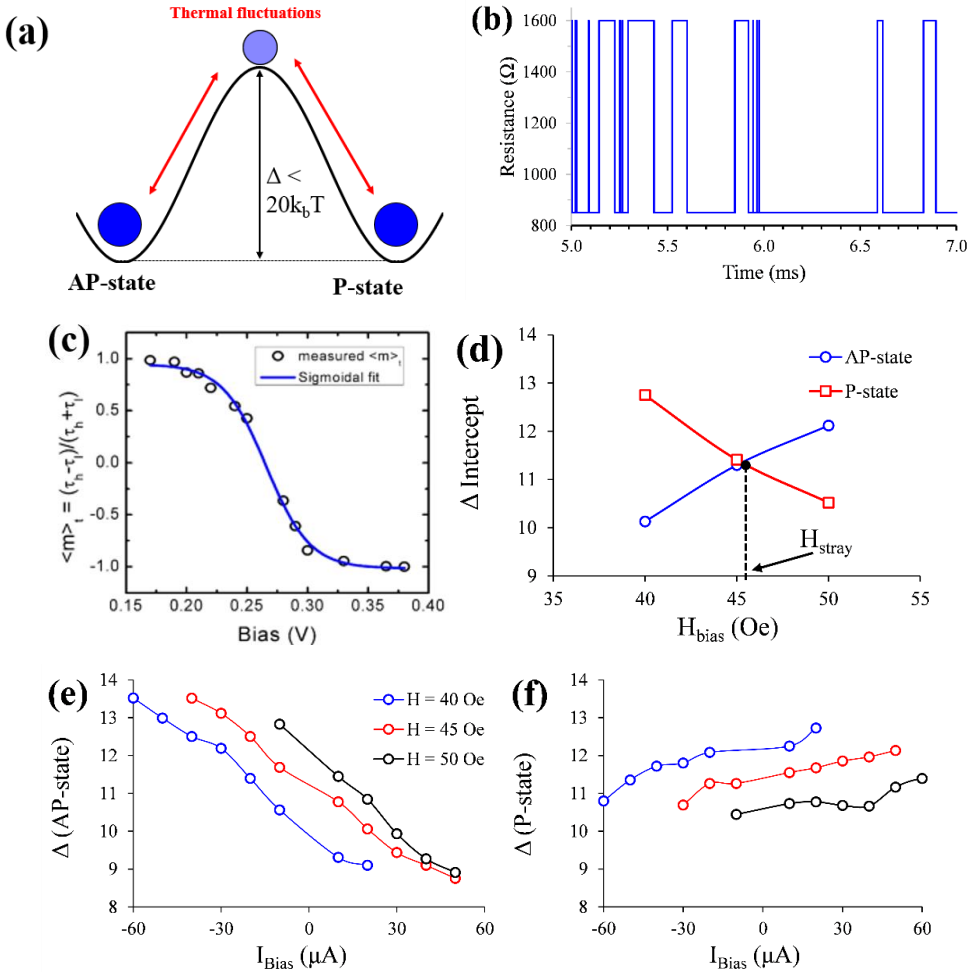


Fig. 2.9. (a) Illustration of the energy barrier between the AP- and P-states for MTJs with thermal stability factors below 20 k_bT. **(b)** Example of telegraphic switching signals generated measured from low-Δ MTJs. **(c)** Example of the average magnetization curve with respect to bias voltage (image extracted from [90]). **(d)** Example of an I_C intercept vs H_{bias} plot used to calculate H_{stray}. **(e,f)** AP and P-state components of the effective thermal stability factor with respect to the bias current.

Thermally unstable MTJs where $\Delta < 20$ will often show random fluctuations near the switching fields, as illustrated in Fig. 2.9(a). For these MTJs, synchronous methods cannot be used to find Δ since these MTJs may switch during or before the reading cycle due to low retention times ($\tau < 0.5$ ms). Therefore, asynchronous methods should be implemented

when calculating Δ for unstable MTJs. In these devices, thermal fluctuations will cause the MTJ to randomly fluctuates between AP and P-states continuously, at switching rates ranging from 1 kHz to above 1 GHz [86-91]. These fluctuations are measured using time-domain data, or telegraphic signals, an example of which is shown in Fig. 2.9(b). For asynchronous measurements, P_{SW} represents the percentage of time the MTJ spends in the AP-state, which can be tuned with a bias current (I_{bias}) as shown in Fig. 2.9(c). Note that I_{bias} is held constant for the duration of the measurement.

To calculate Δ using asynchronous measurements, time-domain should be collected at multiple I_{bias} values under a bias field (H_{bias}) equal to H_{stray}. However, measuring H_{stray} from R-H plots of unstable MTJs is often inaccurate since the MTJ randomly switches states near the switching fields. Therefore, P_{SW} versus I_{bias} curves should be obtained at multiple bias fields to calculate H_{stray}. At each H_{bias}, linear interpolation on the P_{SW} curves for data points between P_{SW} = 0.2 and 0.8 to find I_{bias} where P_{SW} = 0.5, which we will define as I_C. Now, H_{stray} can be obtained from the y-intercept of the I_C versus H_{bias} plot, as shown in Fig. 2.9(d). For asynchronous measurements, Δ is calculated from the average AP- and P-state dwell times (τ_{AP} and τ_P) which are used to calculate the AP- and P-state components of Δ (Δ_{AP} and Δ_P) via the Néel Arrhenius equation [5], as shown in Eq. 2.7. Note that Δ_{AP} and Δ_P are dependent on I_{bias} and H_{bias}, as shown in Fig. 2.9(e) and 2.9(f), and therefore represent effective thermal stabilities rather than the intrinsic Δ of the MTJ. To calculate Δ, the intercepts for the Δ_{AP} and Δ_P versus I_{bias} curves ($\Delta_{AP,0}$ and $\Delta_{P,0}$) were obtained then plotted with respect to H_{bias}. The intrinsic Δ can now be acquired by finding $\Delta_{AP,0}$ and $\Delta_{P,0}$ at $H_{bias} = H_{stray}$. Note that at H_{stray}, $\Delta_{AP,0} = \Delta_{P,0} = \Delta$.

$$\tau_{AP(P)} = \tau_0 \exp\left(\Delta_{AP(P)}\right) \tag{2.7}$$

In this work, we used two different experimental set-ups were used for asynchronous probabilistic switching measurements, which are shown in Fig. 2.10(a) and 2.10(b). In both set-ups, a Keithley 2400 source meter is used supply the bias voltages (or currents) to the MTJ and the time-domain signals are obtained using mixed signal oscilloscopes. The oscilloscopes used in each set-up both have bandwidths of at least 1 GHz.

For measurements where a DC bias current is applied to generate telegraphic switching signals, the set-up in Fig. 2.10(a) was used. In this set-up, the DC bias current is constant, therefore, the random fluctuations between resistance states is captured through voltage measurements. The oscilloscope used in this set-up is a Micsig tBook mini TO1104 portable oscilloscope, which has a maximum sampling rate of 1 GHz and an input resistance of 1 MΩ.

When a DC bias voltage is used, the set-up in Fig. 2.10(b) was used. The oscilloscope in this set-up is a Tektronix DPO 72004C mixed signal oscilloscope, which has a bandwidth of 50 GHz and an input resistance of 50 Ω. Unlike the set-up in Fig. 2.10(a), the MTJ is connected in series with the oscilloscope's input so that the MTJ creates a voltage divider circuit with the 50 Ω input resistance. The voltage measured on the oscilloscope will be lower when the MTJ is in the AP-state than in the P-state. Note nano-sized MTJs typically have resistances on the order of 1 kΩ in both states. Furthermore, the DC bias voltages we used in these experiments ranged from 200 mV and 600 mV. Since the MTJs resistance >> 50 Ω, the voltage measured at the oscilloscope were on the order of 1 mV (in both states), meaning that the voltage across the MTJ is approximately the same in both the AP and P-states.

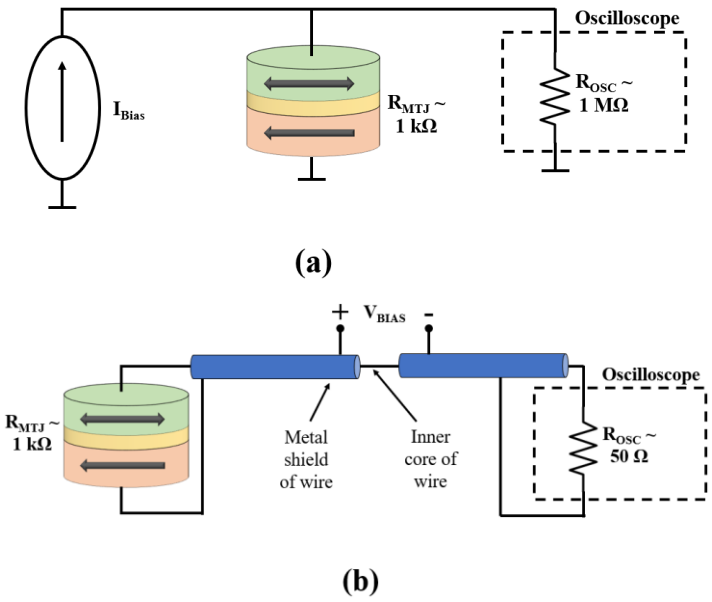


Fig. 2.10. Illustration of the testing set-up for asynchronous probabilistic switching measurements for **(a)** using a bias current where the MTJ and oscilloscope are connected in parallel and **(b)** using a bias voltage where the MTJ and oscilloscope are connected in series.

Chapter 3: Tunable random telegraphic signal generation via dual-biasing

This chapter is based on my previous publications, which are references [77 – 79, 94] in the bibliography.

3.1 Motivation and background

In this chapter, I will focus on the possibilities of MTJs in non-conventional probabilistic computing schemes. More specifically, I will present a method for generating tunable random telegraphic signals in MTJs for asynchronous probabilistic computing methods. In recent years, MTJs with a superparamagnetic free layer (sMTJ) have been proposed as tunable stochastic devices in various asynchronous probabilistic schemes such as p-bits for Boltzmann machines [67] and TRNG units for asynchronous stochastic computing [92]. In these devices, thermal fluctuations alone can drive magnetization switching between the AP and P states. The output of a p-bit is defined by the averaged resistance of the MTJ, which can be controlled with a voltage or current.

A major obstacle for implementation of sMTJs in large-scale networks is that their switching rate and transfer curves are extremely sensitive to variations in device dimensions [93]. In this chapter, we will discuss a solution to avoid the effects of device variations using biasing scheme which we refer to as ‘dual-biasing’. In this method, stochastic switching signals from MTJs are controlled using two individually tunable biases rather than one. In our previous work, we have demonstrated three key features of the dual-biasing method that make it a promising solution for overcoming the effects of device variations. One is that it can be used to generate stochastic switching signals in

thermally stable MTJs as well as unstable MTJs [79], which reduces the influence of device variations since thermally stable MTJs are more robust against these variations. A potential cost of using thermally stable MTJs is the reduction in switching speed. However, this reduction may be avoided since the second key feature of dual-biasing is that the average switching rate of the signals generated can be tuned by over 4 orders of magnitude and reach switching rates above 1 MHz, as demonstrated in our previous work [77-79].

The third, and most unique, feature of dual-biasing what we refer to as two degrees of tunability. This was demonstrated in our previous work [77-79] and refers to our observation that the two biases can control the AP- and P-state dwell times separately. Two degrees of tunability will be discussed in more detail in the following sections, however, the basic implication is that the average output and average switching rate can be tuned independently in each p-bit. These three features show that device variations can be avoided in dual-biased p-bits by either using thermally stable MTJs and using dual-biasing to increase switching rates or by using unstable MTJs and using dual-biasing to correct for device variations. Furthermore, two degrees of tunability in dual-biasing introduces other novel possibilities in neuromorphic computing such as generating neural spiking signals where each signal encodes 2X more information [77], implementing temporal coding in probabilistic neurons [63], and creating neural networks where both neuron and synaptic functions are performed in a single device.

The rest of this chapter is organized as follows. Section 3.2 describes the dual-biasing strategy used and the testing set-up used in our experiments. Section 3.3 presents the results obtained from dual-biasing on single MTJ devices, which demonstrates the two degrees of tunability feature. The data presented in this section were published in [78]. Results from

dual-biasing on two MTJs connected in series are presented in section 3.4, contents of which are published in [94]. In this section, the signals generated have bipolar representation, making them promising TRNG units in stochastic computing. Section 3.5 provides an in-depth exploration of the results in 3.3 and 3.4 which investigates 1) the physical mechanism behind the two degrees of tunability feature of dual-biasing and 2) the effectiveness of dual-biasing on MTJs with a wide range of thermal stabilities. The contents presented in section 3.5.2 are published in [78] and the contents presented in section 3.5.3 are published in [79]. Finally, this chapter is summarized in section 3.6.

3.2 Experimental set-up

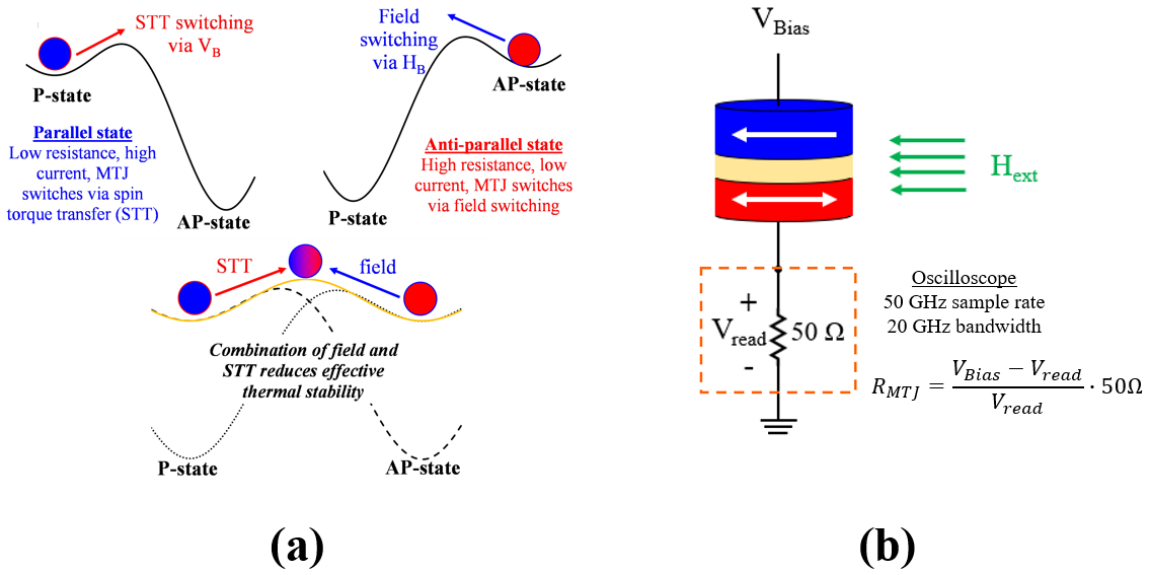


Fig. 3.1. (a) Illustration of the evolution of the energy barrier between the AP and P states during the dual-biasing process. **(b)** Schematic of the dual-biasing experiment.

One way to implement the dual-biasing method is via an external magnetic field (H_{bias}) and a DC voltage (V_{bias}) with specific restrictions on their magnitudes and orientations. The biases applied to the MTJ must be set so that H_{bias} favors the P-state and V_{bias} favors the AP-state. The mechanism for generating telegraphic signals is the evolution of the

energy barrier between the AP- and P-states, as illustrated in Fig. 3.1(a). This evolution occurs because the current through MTJ is larger in the P-state than in the AP-state, meaning that the influence of the STT effect is more significant in the P-state. There are certain combinations of magnitudes of V_{bias} and H_{bias} where the MTJ is never in an energetically favorable state, which causes the magnetization to continuously toggle between the two states, thus generating stochastic signals.

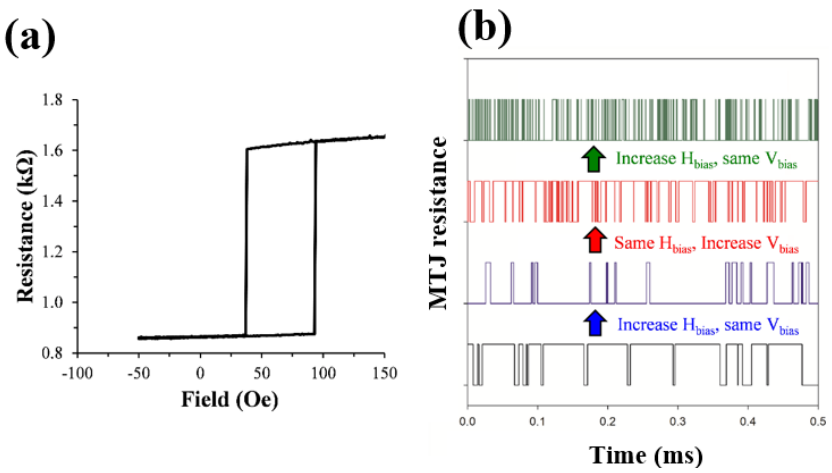


Fig. 3.2. (a) Field switching minor loop of the MTJ tested and illustration of the orientation of H_{bias} and V_{bias} of the dual-biasing method (inset). **(b)** Time-domain data showing the effects of H_{bias} and V_{bias} on both the average resistance state and average switching rate.

The random telegraphic switching signals are measured using a Tektronix DPO 72004C mixed signal oscilloscope, which has a bandwidth of 50 GHz and an input resistance of 50 Ω. In this experiment, the MTJ is connected in series with the oscilloscope's input so that the MTJ creates a voltage divider circuit with the 50 Ω input resistance, as illustrated in Fig. 3.1(b). The voltage measured on the oscilloscope will be lower when the MTJ is in the AP-state than in the P-state. Note nano-sized MTJs typically have resistances on the order of 1 kΩ in both states. Furthermore, the DC bias voltages we used in these experiments ranged from 200 mV and 600 mV. Since the MTJs resistance >>

50 Ω, the voltage measured at the oscilloscope were on the order of 1 mV (in both states), meaning that the voltage across the MTJ is approximately the same in both the AP and P-states.

Our experiment was performed on an MTJ with a nanopillar stack and an elliptical cross section with nominal dimensions of 130 x 50 nm^2. The MTJ has a stack structure of PtMn (15)/SAF/ MgO (0.85)/$Co_{20}Fe_{60}B_{20}$ (1.8), where the thicknesses given are in units of nm and the SAF (synthetic anti-ferromagnetic) layers have a structure of $Co_{70}Fe_{30}$ (2.3)/Ru (0.85)/$Co_{40}Fe_{40}B_{20}$ (2.4). The device has a resistance-area product of 3.5 Ωμm^2, a coercivity (H_C) of 27.5 Oe, and a tunneling magnetoresistance ratio (TMR) of 90.7%, where the AP- and P-state resistances are 1.65 kΩ and 865 Ω respectively. Using the field switching characteristics of the MTJ, we calculated the zero-V_{bias} thermal stability factor (Δ_0) to be 27.2, which means that the anisotropy field (H_K) is 166 Oe. The hysteresis plot for our MTJ is shown in Fig. 3.2(a).

Measurements were collected at H_{bias} = 60, 65, 70, 80, and 90 Oe. Note that the values for H_{bias} include the stray field (H_{stray}) of the MTJ's reference on the free layer (determined by the offset along the horizontal axis of the field switching plot), which was -62 Oe, therefore, the actual bias fields were 122, 127, 132, 142, and 152 Oe. Each trial contained enough data points to capture at least 100 switching events in both the AP-to-P and P-to-AP directions. Each trial was measured for at least 2 ms in 80 ps intervals. From each waveform, I acquired the average AP- and P-state dwell times ($\langle\tau_{AP}\rangle$ and $\langle\tau_P\rangle$, respectively) as well as the time-averaged output, which is refered to as the AP-rate. The AP-rate was defined as the percentage of time the MTJ spent in the AP-state and is represented numerically in a [0,1] range.

Figure 3.2(b) shows the telegraphic signals for three sets tested. The bottom and middle signals shown in Fig. 3.2(b) were obtained at the same H_{bias}, however V_{bias} is larger for the middle signal. Since V_{bias} was set to favor the AP-state, it is unsurprising that the middle signal appears to spend more time in the AP-state than the bottom signal, which indicates that the middle signal has a higher AP-rate. The top and middle two signals shown in Fig. 3.2(b) were obtained at the same V_{bias}, however, H_{bias} is larger for the top signal. This plot shows that the top signal appears to spend more time in the P-state, therefore, has a lower AP-rate). Furthermore, one interesting observation is that the top signal has significantly more switching events in the same time window, implying that increasing H_{bias} also causes the overall switching rate to increase. This observation is explained in further detail in the following section.

3.3 Results

The AP-rate vs V_{bias} plots at all H_{bias} values tested are shown in Fig. 3.3(a). This plot indicates a trade-off between low bias voltages and wider AP-rate versus V_{bias} curves. At lower H_{bias} values, telegraphic switching signals can be generated at lower bias voltages. However, this also means that the AP-rate is tunable through a narrower range of bias voltages. The data from Fig. 3.3(a) demonstrates the AP rate can be tuned from 0 to 1 within a 35 mV range of bias voltages for H_{bias} = 60 Oe. This range increases to 60 mV for H_{bias} = 80 Oe and 100 mV for H_{bias} = 90 Oe.

The tunability of the AP rate is not unexpected and has already been demonstrated in several previous experiments [71, 90]. However, the interesting result from our data is seen in the relation between H_{bias} and V_{bias} versus $\langle \tau_{AP} \rangle$ and $\langle \tau_P \rangle$. Recall the observation from the plots in Fig. 3.2(b), where increasing H_{bias} and V_{bias} appear to influence the AP-rate,

however, overall switching rate also increases. Based on the orientation of H_{bias} and V_{bias}, it would be expected that increasing H_{bias}(V_{bias}) would decrease $\langle\tau_{AP}\rangle$ ($\langle\tau_P\rangle$) but increase $\langle\tau_P\rangle$ ($\langle\tau_{AP}\rangle$), meaning that the overall switching rate would not change. However, the increase in the switching rates seen in Fig. 3.2(b) suggests that H_{bias} and V_{bias} have asymmetric influences on $\langle\tau_{AP}\rangle$ and $\langle\tau_P\rangle$.

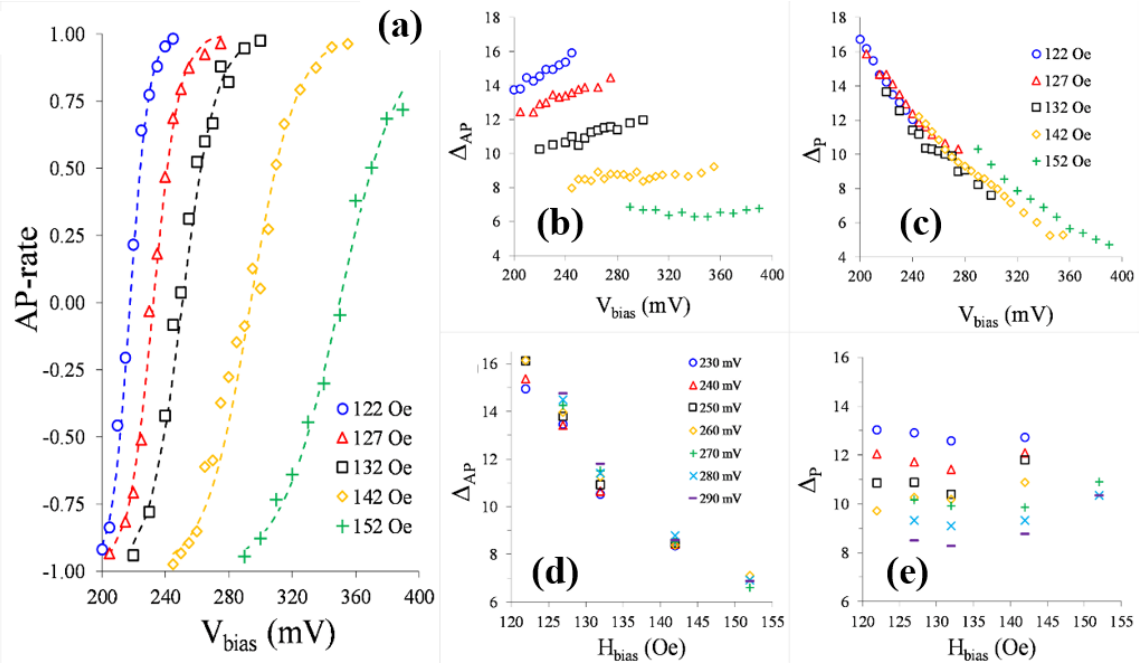


Fig. 3.3. (a) AP-rate curves vs V_{bias} at all H_{bias} values tested. **(b,c)** V_{bias} dependence of the AP- and P-state components of the effective thermal stability factors at fixed H_{bias} values. **(d,e)** H_{bias} dependence of the AP- and P-state components of the effective thermal stability factors at fixed V_{bias} values. Images were extracted from [78].

These observations are substantiated in Fig. 3.3(b) – 3.3(e), which show plots of $\langle\tau_{AP}\rangle$ and $\langle\tau_P\rangle$ vs V_{bias} and H_{bias}. Note that in these plots, $\langle\tau_{AP}\rangle$ and $\langle\tau_P\rangle$ are represented in terms of the AP- and P-state components of the effective thermal stability factors, Δ_{AP} and Δ_P, where $\Delta_{AP(P)} = \Delta_0 \ln(\langle\tau_{AP(P)}\rangle/\tau_0)$. Figures 3.3(b) and 3.3(c) shows that changes in Δ_{AP} with V_{bias} are much smaller than changes in Δ_P with V_{bias}. The opposite relation is shown in Fig. 3.3(d) and 3.3(e), where changes in Δ_{AP} with H_{bias} are much smaller than changes

in Δ_P with H_{bias}. It should be noted that V_{bias} has a small effect on Δ_{AP} and H_{bias} has a small effect on Δ_P, however, these effects become insignificant when comparing the changes in Δ_{AP} and Δ_P with H_{bias} and V_{bias} in terms of pulse widths. Figures 3.3(b) and 3.3(c) show that changes in Δ_P with V_{bias} are approximately 3 to 6 times larger than changes in Δ_{AP} with V_{bias}, depending on H_{bias}, which means that increasing V_{bias} decreases $\langle \tau_P \rangle$ by 3 – 4 orders of magnitude yet $\langle \tau_{AP} \rangle$ only increases by 1 order of magnetude. Alternatively, increasing H_{bias} decreases $\langle \tau_{AP} \rangle$ by ~4 orders of magnitude when $\langle \tau_P \rangle$ increases by 1 order of magnitude. This asymmetry on the influence of V_{bias} and H_{bias} on $\langle \tau_{AP} \rangle$ and $\langle \tau_{AP} \rangle$ gives rise to the two degrees of tunability capability in the dual-biasing method.

This discrepancy in how the two biases effect $\langle \tau_{AP} \rangle$ and $\langle \tau_P \rangle$ separately allows for independent control over the AP rate and the average switching rate. This capability means that dual-biasing has several promising prospects for probabilistic computing applications using asynchronous processing methods. For example, if implemented in a spiking neural network where the MTJs are used to emulate spiking neural signals, the two degrees of tunability feature increases the information capacity in the rate coding scheme. This is because the average spike rate and the spike count now are independent variables that can be processed separately within the same decoding scheme. Additionally, this introduces a method of overcoming the effects of devices variations in large-scale p-bit networks through on-board corrections.

3.4 Bipolar Random Spiking Results

In the previous section, I showed that dual-biasing on a single MTJ generates random telegraphic signals with two degrees of tunability, indicating that dual-biased MTJs are promising devices as p-bits in Boltzmann machines and stochastic spiking neurons in

spiking neural networks. In these applications, the MTJ behaves as a single, asynchronous stochastic unit with a tunable output probability. However, another method of generating random numbers is through a random source with an output probability fixed at 0.5. A stochastic bit-stream, X, is then generated by connecting a digital input, x, and the output of the random source to a digital comparator. The probability of any bit in X being '1' is directly correlated with x. In this section, I demonstrate that a RNG with a fixed output probability of 50% is built when dual-biasing is applied to two MTJs connected in series, as shown in Fig. 3.4(a). This configuration generates bipolar signals with three logic states, meaning that the numeric output can be represented within a [-1, +1] range rather than a [0, +1] range. Note that an output probability of 50% corresponds to a bipolar value of 0.

Bipolar encoding has several advantages over unipolar encoding in probabilistic and neuromorphic applications. For example, in stochastic computing (SC), bipolar encoding expands the numeric representation range of the stochastic bit-streams from [0, +1] to [-1, +1], which enables efficient processing of a broader range of arithmetic functions such as scaled subtraction and hyperbolic tangent function [43]. Furthermore, X. Yang et al showed that bipolar encoding also has advantages in spiking neural networks (SNN) as it enables transmission of both excitatory and inhibitory neural signals, thus reduces the total circuit size by a half and improves learning efficiency [95]. Another example where bipolar encoding is promising is for neural activation functions in Hopfield networks, where M. Mansor et al showed that using a bipolar sigmoid activation function for combinatorial optimization problems improves output and reduces learning time when compared to networks using the conventional McCulloch-Pitts function [96].

Typically, bipolar algorithms are still encoded using binary bits, where a binary bit '0' represents a bipolar bit '-1'. This creates several short-comings in computing schemes using bipolar encoding. In SC, bipolar encoding reduces the numeric resolution of the stochastic bit-streams by one half [43]. The resolution can be improved by doubling the size of the bit-streams, however, this would unfortunately double the computation delay. Furthermore, in SNNs and Hopfield networks, bipolar encoding doubles the number of transmission lines between neurons since separate channels are needed for excitatory and inhibitory neural signals.

The method demonstrated in this section (illustrated in Fig. 3.4(a)) would alleviate these short-comings since the bipolar signals with three voltage levels are used to represent bipolar data, therefore, the physical signals can be mapped directly to the numeric representation. The orientation and polarity of V_{bias} and H_{bias} are the same in this experiment as those for dual-biasing on single MTJ devices (recall sections 3.2 – 3.3). Note that H_{bias} is a global field meaning that H_{bias} is equal in both MTJs and therefore, the two MTJs should be similar in terms of R_{AP}, R_P, TMR, H_C, and H_{STRAY}. The field switching (R-H) plots of both MTJs tested are shown in Fig. 3.4(b), whic shows that R_{AP} = 2.35 kΩ and 2.43 kΩ, R_P = 960 Ω and 1050 Ω, TMR = 145% and 132%, H_C = 35 Oe and 28 Oe, and H_{STRAY} = -73 Oe and -74 Oe for the top and bottom MTJ, respectively. The MTJs tested are elliptical nanopillars with in-plane anisotropy, RA products of 3.5 Ω·μm^2, and nominal dimensions of 130 x 50 nm^2. The stack structure is PtMn (15)/SAF/MgO (0.8)/ $Co_{20}Fe_{60}B_{20}$ (2.0), where layer thicknesses are in nanometers and the SAF structure is comprised of $Co_{70}Fe_{30}$ (2.5)/Ru (0.85)/$Co_{40}Fe_{40}B_{20}$ (2.4).

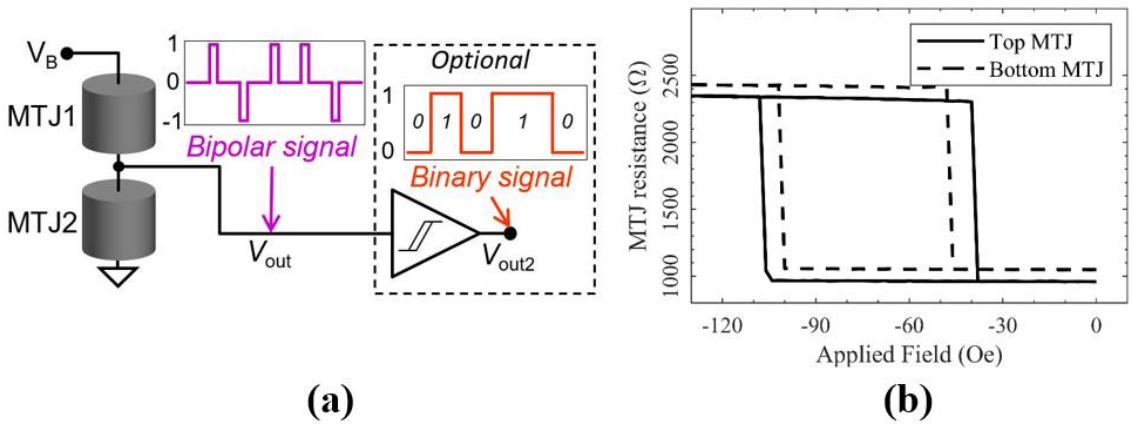


Fig. 3.4. (a) Experimental set-up for generating bipolar random numbers. **(b)** Field switching plots for both top and bottom MTJs tested. Images extracted from [94].

Measurements were collected in four sets at H_{bias} = 5, 10, 15, and 20 Oe and at V_{bias} ranging between 780mV to 1.4V in each set. The time-domain data of the bipolar random signals were measured and recorded by a Micsig tBook mini TO1104 table oscilloscope with sampling rates ranging from 20 kHz to 1 GHz. Figure 3.5(a) shows an example of the time-domain signals obtained at H_{bias} = 15 Oe and V_{bias} = 1060 mV, which clearly shows that the output signals have three distinct voltage levels. Since the two MTJs in series serve as a voltage divider, the lowest signal level occurs when the combination of states of the two MTJs (represented as [top state, bottom state]) is [AP, P], the highest signal level occurs for [P, AP], and the middle level occurs at either [P, P] or [AP, AP]. For simplicity, the voltage signal data was converted into a series of ternary values of +1, -1, and 0, which corresponds to the highest, lowest, and middle voltage level, respectively. Figure 3.5(b) shows the signals from Fig. 3.5(a) in terms of bipolar bits.

The data obtained was analyzed from two perspectives. One was in terms of the average dwell times of the high and low voltage spikes (T_H and T_L, respectively), which is the average time the MTJs spend in the +1 and -1 bipolar state before switching. The second

was in terms of the average output ($\langle y^* \rangle$), which is the time-averaged value of the bipolar data, and can be calculated from T_H and T_L in the expression $\langle y^* \rangle = ((-1) * N_L * T_L + N_H * T_H)/T_{meas}$, where N_L and N_H are the number of low and high spikes measured, respectively, and T_{meas} is the total measurement time.

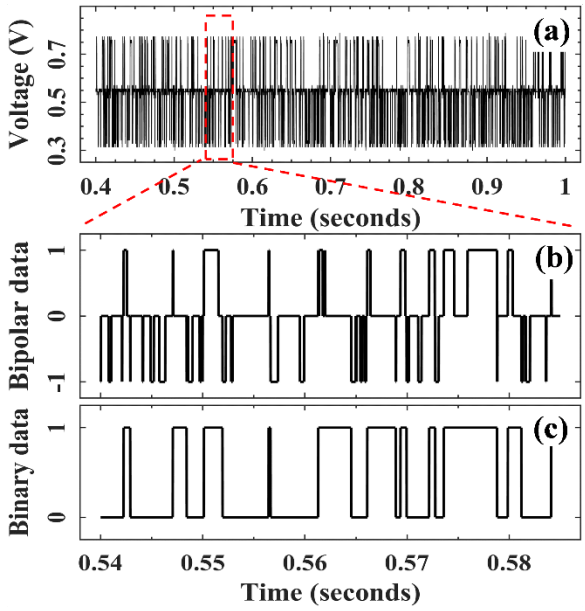


Fig. 3.5. Time domain data at a bias field of 15 Oe and bias voltage of 1060 mV of **(a)** output voltages measured from oscilloscope, **(b)** bipolar representation of the output signal, and **(c)** binary representation of the output signal. Images extracted from [94].

Additionally, the output signals were converted into binary data in the same way a Schmitt trigger would do to the MTJs' bipolar signal, as shown in Fig. 3.5(c). The numeric representation of the binary data is condensed from a range of [-1, 1] for bipolar data to a [0, 1] range. For this conversion, only +1 and -1 spikes activate binary switching. When the bipolar signal is at -1, the binary state is 0 and will remain at 0 until the bipolar signal switches to +1, at which point the binary state becomes 1. Conversely, the binary state will remain at 1 until the bipolar signal switches to -1. Note that the binary state is not affected if the bipolar state switches between -1 and 0 or between +1 and 0. As with the bipolar

data, the binary data was also analyzed in terms of the average high and low dwell times (T_1 and T_0, respectively) and the average output ($\langle y \rangle$).

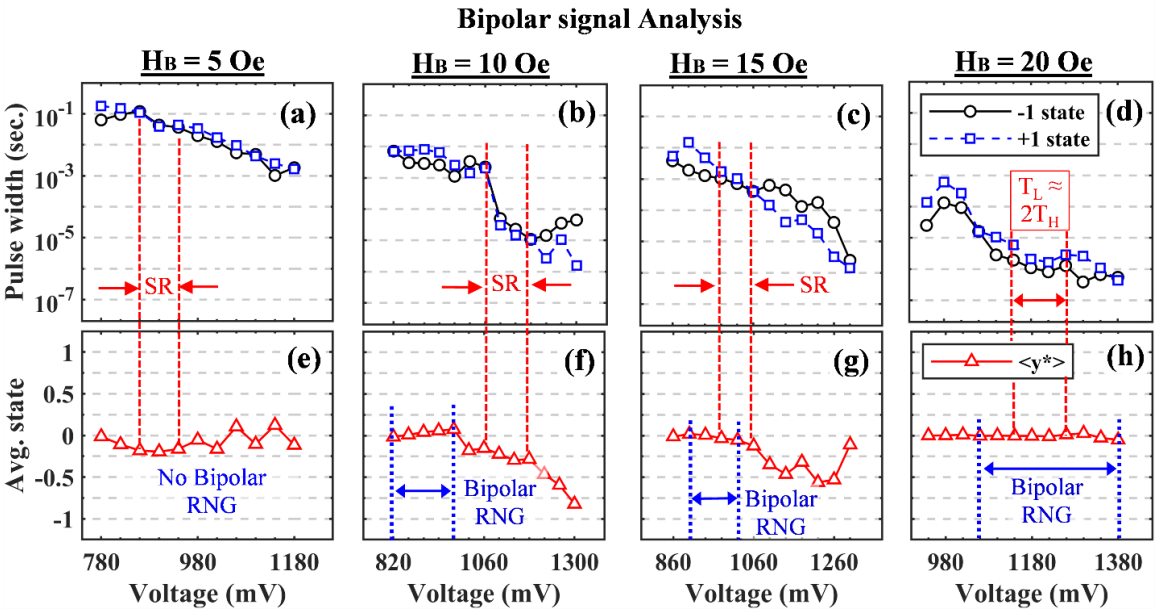


Fig. 3.6. Average dwell times for both low-state and high-state pulses from bipolar signals at bias fields of **(a)** 5 Oe, **(b)** 10 Oe, **(c)** 15 Oe, and **(d)** 20 Oe. At each bias field, the synchronization region (SR), where $T_H \approx T_L$, is indicated. Average bipolar state at bias fields of **(e)** 5 Oe, **(f)** 10 Oe, **(g)** 15 Oe, and **(h)** 20 Oe. Each bias field shows the range of data points which passed the three NIST STS tests used in our analysis, which is labelled "Bipolar RNG". Images extracted from [94].

A proper bipolar RNG will produce signals where $\langle y^* \rangle \approx 0$. Additionally, for RNGs to be considered in many modern applications (such as cryptography), they must possess several additional characteristics such as unpredictability, non-repeating patterns, and no correlation between consecutive bits. Therefore, I applied three tests from the standardized National Institute of Standards and Technology Statistical Test Suite (NIST STS [97]) to the bipolar data to determine if the signals generated satisfy cryptographic quality requirements. The three tests applied were the Runs test, Non-overlapping template matching test, and the Discrete Fourier Transform (DFT) test, which detects too many (or

too few) occurrences of runs of +1 or -1 consecutive bits, non-periodic patterns, and periodic features, respectively.

Measurements for T_H and T_L from the bipolar data is shown in Fig. 3.6(a) – 3.6(d). These figures show that there is a range of V_B values where $T_H \approx T_L$, which we refer to as the synchronization region (SR) at each H_B. However, at H_B = 20 Oe, the SR is difficult to identify since $T_H \approx T_L$ very briefly at V_B = 1060 mV and 1380 mV. However, within the range of V_B = 1140 – 1260 mV, there appears to be a strong correlation between T_H and T_L, where $T_H \approx 2T_L$. For the data at H_B = 20 Oe, this region was selected as the SR. The V_B range for each SR is shown in Table 3.1.

Figures 3.6(e) – 3.6(h) show the measurements for $\langle y^* \rangle$ at each H_B and highlights the V_B range where the data sets passed the three NIST STS tests used in our analysis, which are labelled "Bipolar RNG". These figures illustrate that $T_H = T_L$ does not guarantee that $\langle y^* \rangle = 0$. For example, at H_B = 5, 10, and 15 Oe, $\langle y^* \rangle$ = -0.25, -0.2, and -0.15, respectively, within the SRs. This indicates that -1 pulses occur more frequently, thus causing $\langle y^* \rangle$ to skew towards -1 when $T_H = T_L$. Conversely, each data set, except H_B = 5 Oe, has points where $\langle y^* \rangle \approx 0$ at V_B values outside the designated SR. This is because T_H and T_L are measured independently to the rate at which +1 and -1 spikes are generated. Based on our definition of $\langle y^* \rangle$, the criteria for T_H and T_L to produce $\langle y^* \rangle \approx 0$ is $N_H T_H = N_L T_L$ or $N_L/N_H = T_H/T_L$.

The V_B range that produced $\langle y^* \rangle \approx 0$ and passed the NIST STS tests applied to our data is listed in Table 3.1 at each H_B. This table shows that the range of V_B values that produced $\langle y^* \rangle \approx 0$ increased as H_B increased. However, the most notable results are from the H_B =

20 Oe data set, where $\langle y^* \rangle \approx 0$ and passed all three NIST STS tests for most of the V_B range, despite the fact that $T_H > T_L$ at nearly all V_B. It should be noted that Fig. 3.6(g) and (h) show a few data sets, namely sets where H_B is high and V_B is relatively small, generated signals where $\langle y^* \rangle \approx 0$ but failed at least one of the NIST STS tests. For these sets, the dwell times between consecutive spikes is much larger than T_H and T_L, thus causing the signal to fail the DFT test. The results shown in Fig. 3.6(h) demonstrate that dual-biasing on two MTJs connected in series can produce signals with true bipolar representation for bipolar RNGs that are robust against V_B variations at certain H_B values. Additionally, the fact that $\langle y^* \rangle \approx 0$ even when $T_H \neq T_L$ indicates that these signals are also robust against variations in intrinsic device properties between the two MTJs.

Table 3.1. Overview of bipolar and binary synchronization ranges.

Bias Field	Bipolar Synchronization range	Bipolar RNG range	Binary Synchronization range	$\langle y \rangle$ value in synchronization range
5 Oe	860 – 940 mV	NA	940 – 1180 mV	0.45 – 0.6
10 Oe	1060 – 1180 mV	820 – 940 mV	900 – 1060 mV	0.45
15 Oe	980 – 1060 mV	900 – 1020 mV	940 – 1180 mV	0.3
20 Oe	1140 – 1260 mV	1060 – 1380 mV	1100 – 1220 mV	0.2 – 0.25

In the following discussion, I will explain how the robustness of $\langle y^* \rangle$ at H_B = 20 Oe can be attributed to the two degrees of tunability capability of the dual-biasing method. Recall the results presented in section 4.3, which showed that $|d\Delta_{AP}/dV_B| << |d\Delta_P/dV_B|$ at all H_B. Furthermore, as H_B increased, $d\Delta_{AP}/dV_B$ decreased and eventually reached zero, meaning V_B had no influence Δ_{AP}. The influence of V_B on two MTJs in series will be similar to a single MTJ, except for the fact that V_B is divided between the top and bottom MTJs (V_{TOP} and V_{BOT}, respectively) which are dependent on the state configuration of the

two MTJs. When the MTJs are in the [P, P] or [AP, AP] state configuration ([top state, bottom state]), then $V_{TOP} = V_{BOT}$ and a bipolar value of 0 is produced. When one of the MTJs switches, the MTJ in the AP-state will have a larger voltage and bipolar values of +1 and -1 will be produced for MTJ state configurations of [P, AP] and [AP, P], respectively.

Two degrees of tunability means that changes in V_{TOP} and V_{BOT} will significantly influence Δ_P^{TOP} and Δ_P^{BOT} (Δ_P for top and bottom MTJs, respectively) whereas changes in Δ_{AP}^{TOP} and Δ_{AP}^{BOT} will be negligible. Since $T_{AP(P)} \propto \exp(\Delta_{AP(P)})$, T_P^{BOT} (T_P^{TOP}) will increase by several orders of magnitude more than T_{AP}^{TOP} (T_{AP}^{BOT}) will increase when a +1 (-1) spike occurs. This means that MTJs in the [AP, P] or [P, AP] state configurations have a much higher probability of switching to the [P, P] state configuration rather than to the [AP, AP] state configuration. Therefore, T_P^{BOT} (T_P^{TOP}) determines the frequency of +1 (-1) spike initiation and T_{AP}^{BOT} (T_{AP}^{TOP}) determines their duration. Now, N_L/N_H and T_H/T_L can be substituted with T_P^{BOT}/T_P^{TOP} and $T_{AP}^{BOT}/T_{AP}^{TOP}$, respectively, therefore, the criteria for $\langle y^* \rangle = 0$ is expressed as $T_P^{BOT}/T_P^{TOP} = T_{AP}^{BOT}/T_{AP}^{TOP}$ or $\Delta_P^{BOT} - \Delta_P^{TOP} = \Delta_{AP}^{BOT} - \Delta_{AP}^{TOP}$.

Considering that dual-biasing provides a high degree of tunability of Δ_{AP} and Δ_P, this condition will likely be met under some V_B and H_B conditions for any pair MTJs fabricated from the same stack as long as some overlap exists in the V_B and H_B switching ranges between the two MTJs. However, to meet this condition over a large V_B range, then $d\Delta_P/dV_B$ and $d\Delta_{AP}/dV_B$ should be equal between the two MTJs. Recent work from our research group found that $|d\Delta_P/dV_B|$ decreases slightly with H_B but is not influenced by H_C [79] (described in further detail in section 4.5). Therefore, even though $H_C^{TOP} > H_C^{BOT}$, it is likely that $d\Delta_P/dV_B$ is equal between the two MTJs at all H_B since $H_{stray}^{TOP} \approx H_{stray}^{BOT}$ (recall

Fig. 3.4(b)). Additionally, this recent work from our research group showed that MTJs with larger H_C have larger $|d\Delta_{AP}/dV_B|$ at the same H_B, however, $d\Delta_{AP}/dV_B$ eventually reached 0 for all MTJs, regardless of H_C, when H_B was increased to sufficiently large values [79]. This means that this method can produce true random bipolar numbers centered around zero even in the presence of device variations.

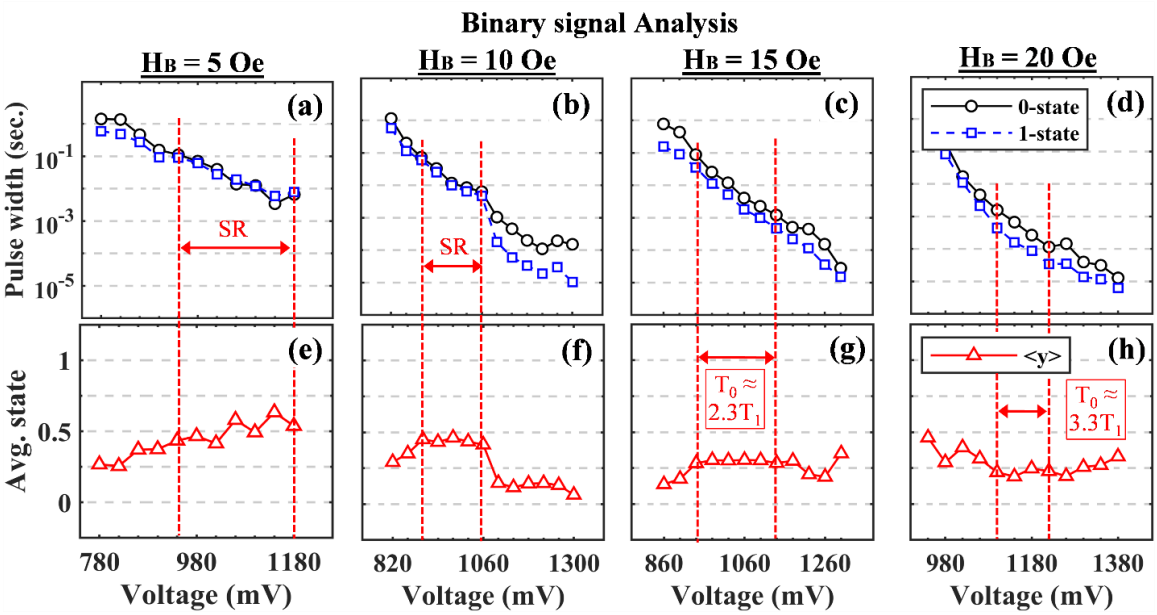


Fig. 3.7. Average dwell times for both low-state and high-state pulses from binary signals at bias fields of **(a)** 5 Oe, **(b)** 10 Oe, **(c)** 15 Oe, and **(d)** 20 Oe. At each bias field, the synchronization region (SR), where $T_1 \approx T_0$, is indicated. Average bipolar state at bias fields of **(e)** 5 Oe, **(f)** 10 Oe, **(g)** 15 Oe, and **(h)** 20 Oe. Images extracted from [94].

Measurements for T_1 and T_0 are shown in Fig. 3.7(a) – 3.7(d), which also indicates the SR, where strong correlations are observed between T_1 and T_0. At H_B = 5 Oe and 10 Oe, SR is defined when $T_1 \approx T_0$, at H_B = 15 Oe, SR is defined when $T_0 \approx 2.3T_1$, and at H_B = 20 Oe, SR is defined when $T_0 \approx 3.3T_1$. The V_B range for each SR for binary data is shown in Table 3.1.

Figures 3.7(e) – 3.7(h) shows the measurements for $\langle y \rangle$ at each H_B. At H_B = 5 Oe, $\langle y \rangle$

mirrors the $\langle y^* \rangle$ data within the binary SR, however, $\langle y \rangle$ never consistently settles to 0.5. At H_B = 10 Oe, 15 Oe, and 20 Oe, $\langle y \rangle$ saturates to values of 0.45, 0.3, and 0.2, respectively (listed in Table 3.1) which indicates that $T_0 > T_1$ for these sets and that T_1 decreases with H_B at a faster rate than T_0. Small variations in $\langle y \rangle$ at H_B = 10 Oe and 15 Oe suggests that reliable binary random numbers can be generated within the SRs. For the binary signals to be considered in most applications involving RNGs, $\langle y \rangle$ should saturate at 0.5, however, there are some probabilistic computing and stochastic computing applications that could potentially utilize this feature to generate constant $\langle y \rangle$ that does not equal 0.5.

The results of the binary data are indicative of the performance of our method in unipolar encoding schemes with synchronous operations since the +1 and -1 spikes can trigger random bit generation. Alternatively, the bipolar data is indicative of the performance of our method in bipolar encoding schemes with asynchronous operations. The results presented for $\langle y \rangle$ and $\langle y^* \rangle$ (particularly at H_B = 20 Oe) illustrates that the method presented is best suited for computing schemes with bipolar encoding and asynchronous read and write operations. This is because dual-biasing ensures a fixed output that is resilient to V_B changes and MTJ device variations. Furthermore, the signals generated in this method have true bipolar representation, thus alleviating the short-comings of bipolar encoding described earlier in this section.

3.5 Discussions

Two degrees of tunability demonstrated in the previous sections is potentially a promising feature for applications such as large-scale p-bit networks, stochastic spiking neurons, and bipolar RNGs. While the results discussed in the previous section provided

strong proof that dual-biasing is capable of generating signals with two degrees of tunability, they did not reveal the physical mechanism that explains this feature. Without an understanding of this mechanism, we cannot understand the possibilities and limitations of dual-biasing. This section is divided into three sub-sections. In section 3.5.1, I will describe quantify the expected influence of H_{bias} and V_{bias} using the Stoner-Wohlfarth model and compare these calculations with the results presented in section 3.3. Then in section 3.5.2, I will explain how two degrees of tunability is achieved in dual-biasing by fitting the data shown in section 3.3 to the Néel-Brown model. Lastly, I will discuss the trade-offs of dual-biasing on sMTJs vs thermally stable MTJs by presenting results from dual-biasing measurements on multiple MTJs with varying Δ_0.

3.5.1. Comparison with current magnetic dynamics models

The mechanism that allows for rapid stochastic switching between the AP- and P-states is the evolution of the energy barrier separating the two states, which is illustrated in Fig. 3.8(a) – 3.8(d). These figures provide a very basic description of how H_{bias} and V_{bias} influence the energy (E) with respect to the orientation of the free layer magnetization (θ) using the Stoner-Wohlfarth model [98].

At thermal equilibrium, $E(\theta)$ is described using Eq. 3.1, where K_U is the anisotropy energy constant and V is the free layer volume. In this condition, $E(\theta)$ will be symmetric between the AP- and P-states, as illustrated in Fig. 3.8(a). When a bias field is applied in the direction that favors P-state orientation, $E(\theta)$ is now expressed using Eq. 3.2, where ϕ represents the angle of H_{bias} relative to the easy-axis (EA). Figure 3.8(b) indicates that the energy barrier in the AP-state ($\Delta E(\pi)$) increases and whereas the energy barrier in the P-state ($\Delta E(0)$) decreases. Applying a bias voltage which favors the AP-state orientation has

the opposite effect on $\Delta E(\pi)$ and $\Delta E(0)$, as illustrated in Fig. 3.8(c). The effect of V_{bias} on $E(\theta)$ is expressed in Eq. 3.3, where $H_{ST}(\theta)$ represents a field term characterized by STT that can be described in terms of V_{bias}, STT efficiency (η), and MTJ resistance ($R_{MTJ}(\theta)$), where $R_{MTJ}(\theta) = R_P\left(TMR\sin^2\left(\frac{\theta}{2}\right)+1\right)$, where R_P is the P-state resistance.

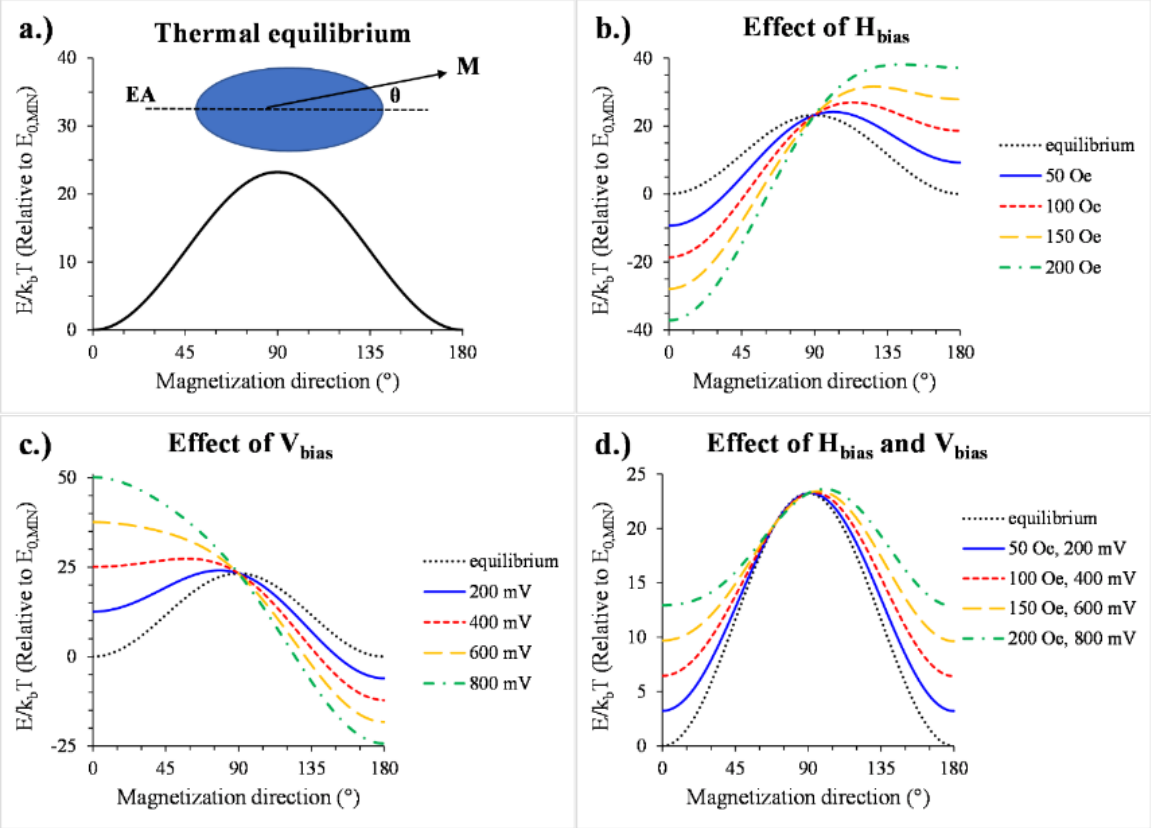


Fig. 3.8. Illustration of the energy barrier between AP- and P- states at **a.)** equilibrium (no external biases), **b.)** $H_{bias} > 0$ (favoring P-state orientation, **c.)** $V_{bias} > 0$ (favoring AP-state orientation, and **d.)** both H_{bias} and $V_{bias} > 0$.

$$E(\theta) = K_U V sin^2(\theta) \tag{3.1}$$

$$E(\theta) = K_U V sin^2(\theta) - H_{bias} M_S V cos(\phi - \theta) \tag{3.2}$$

$$E(\theta) = K_U V sin^2(\theta) - H_{ST}(\theta) M_S V cos(\psi - \theta) \;\rightarrow\; H_{ST}(\theta) = \frac{\eta\hbar}{2eM_S V} \cdot \frac{V_{bias}}{R_{MTJ}(\theta)} \tag{3.3}$$

Using the definitions of $R_{MTJ}(\theta)$ and $H_{ST}(\theta)$ in (3), the influence of $E(\theta)$ with H_{bias} and V_{bias} can be expressed using Eq. 3.4. Note that Eq. 3.4 assumes that $\phi = 0$ and $\psi = \pi$, based on the desired orientations of H_{bias} and V_{bias}. This expression suggests that V_{bias} has a larger influence on $E(0)$ than $E(\pi)$, which is expected considering that the MTJ draws more current in the P-state since the resistance is lower in the P-state than the AP-state. Figure 3.8(d) illustrates that the combination of V_{bias} and H_{bias} causes both $\Delta E(\pi)$ and $\Delta E(0)$ to decrease. This means that dual-biasing can tune the effective thermal stability of an MTJ, meaning that any MTJ can generate random telegraphic switching signals for p-bits, regardless of their intrinsic properties.

$$E(\theta) = M_S V \left\{ \frac{H_K}{2} sin^2(\theta) - \left[H_{bias} - \frac{\eta \hbar V_{bias}}{e M_S V R_P} (TMR - TMR cos(\theta) + 2)^{-1} \right] \cos(\theta) \right\} \quad (3.4)$$

The calculations for Δ_{AP} and Δ_P are shown in Fig. 3.9(a-d), which also shows the experimental data presented in section 3.3 for comparison. These results show two key inaccuracies that are predicted by the Stoner-Wohlfarth model. First is that the Stoner-Wohlfarth model over-estimates values for Δ_P at low H_{bias} but under-estimates values for Δ_P at high H_{bias}. The second, and more important, is that the Stoner-Wohlfarth model predicts that V_{bias} has a much stronger influence on Δ_{AP} then our data shows. While the Stoner-Wohlfarth model can provide a basic understanding in how the dual-biasing method can generate random telegraphic signals in thermally stable MTJs, however, it does not predict the two degrees of tunability capability of dual-biasing. To illustrate how two degrees of tunability is achieved, we will fit the data presented in section 3.3 to the Néel-Brown model.

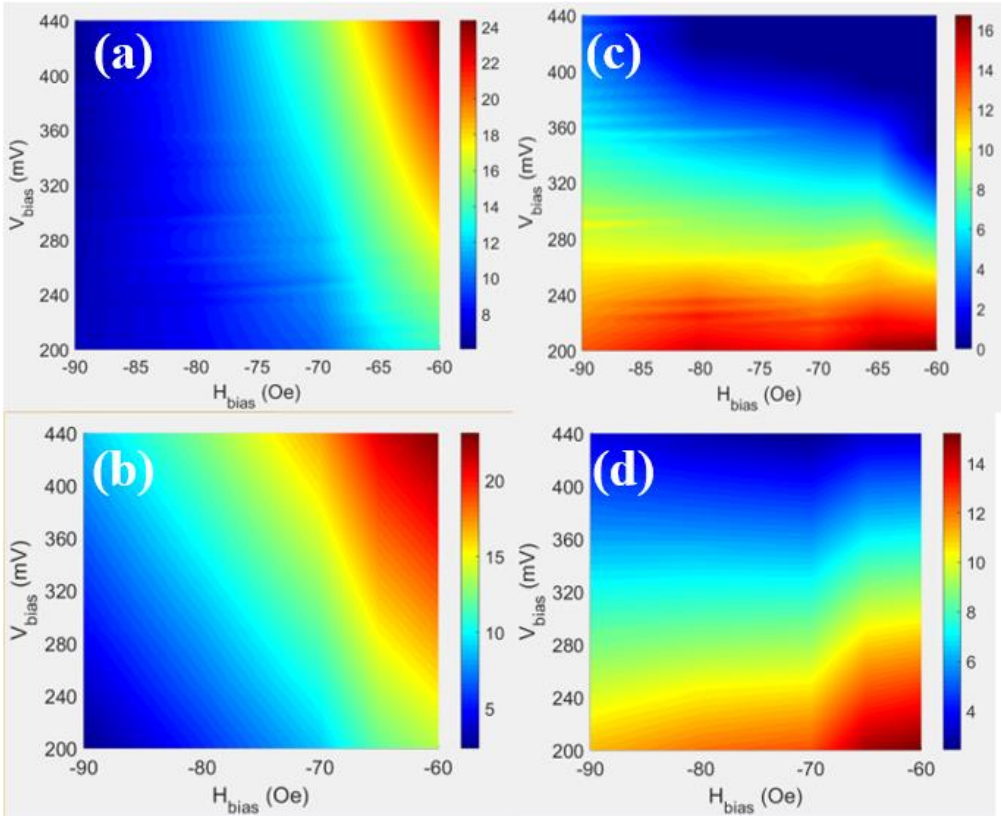


Fig. 3.9. Comparison between experimental data (also presented in Fig. 4.3) and calculations using the Stoner-Wohlfarth model for **(a-b)** Δ_{AP} measured and calculated **(c-d)** Δ_{P} measured and calculated.

3.5.2. Physical Mechanism for two degrees of tunability

In this section, I will explain how dual-biasing achieves two degrees of tunability using the Neel-Brown model [99], which is expressed using Eq. 3.5. In this equation, ↑ and ↓ indicate AP- and P-state components, respectively. A straight-forward application of Eq. 3.5 would predict that $d\Delta_{AP}/dV_{bias}$ is two orders of magnitude larger than what is observed in the experimental data. Therefore, I will investigate the dependence of $V_{C0}^{\uparrow(\downarrow)}$ and $H_K^{\uparrow(\downarrow)}$ on H_{bias} and V_{bias} to determine if these factors are responsible for two degrees of tunability.

$$\Delta_{AP(P)} = \Delta_0 \left(1 \mp \frac{H_{bias}}{H_K^{\uparrow(\downarrow)}}\right)^2 \left(1 \pm \frac{V_{bias}}{V_{C0}^{\uparrow(\downarrow)}}\right) \tag{3.5}$$

The first step of this analysis is to determine the center and width of the AP-rate curves by fitting them to the hyperbolic tangent function expressed shown on the left side of Eq. 3.6. In this equation, $\langle r \rangle$ represents the AP-rate, $\langle r^* \rangle$ represents $\langle r \rangle$ values expanded from a [0, 1] range to a [-1, +1] range using the simple conversion $\langle r^* \rangle = 2\langle r \rangle - 1$, and V_0 and V_R represent the center and width of the AP-rate curves, respectively. V_0 is defined as V_{bias} when $\langle r \rangle$ = 0.5 and V_R is defined as the difference in V_{bias} at $\langle r \rangle$ = 0.25 and $\langle r \rangle$ = 0.75. The fitting parameters V_0 and V_R were calculated through a linear fit of the right side of Eq. 3.6, which was derived using the definition $\tanh(x) = (e^{2x} - 1)/(e^{2x} + 1)$. Figure 3.10(a) shows the calculations for V_0 and V_R for each H_{bias} set. Here, the relation between these fitting parameters and H_{bias} is assumed to be linear, which is a valid approximation for the range of H_{bias} values tested since the linear correlation coefficients for V_R vs H_{bias} and V_0 vs H_{bias} were determined to be 0.993 and 0.986 respectively.

$$\langle r^* \rangle = \tanh\left(\frac{V_{bias} - V_0}{V_R}\right) \quad \leftrightarrow \quad \frac{V_{bias} - V_0}{V_R} = \frac{1}{2}\ln\left(\frac{1 + \langle r^* \rangle}{1 - \langle r^* \rangle}\right) \tag{3.6}$$

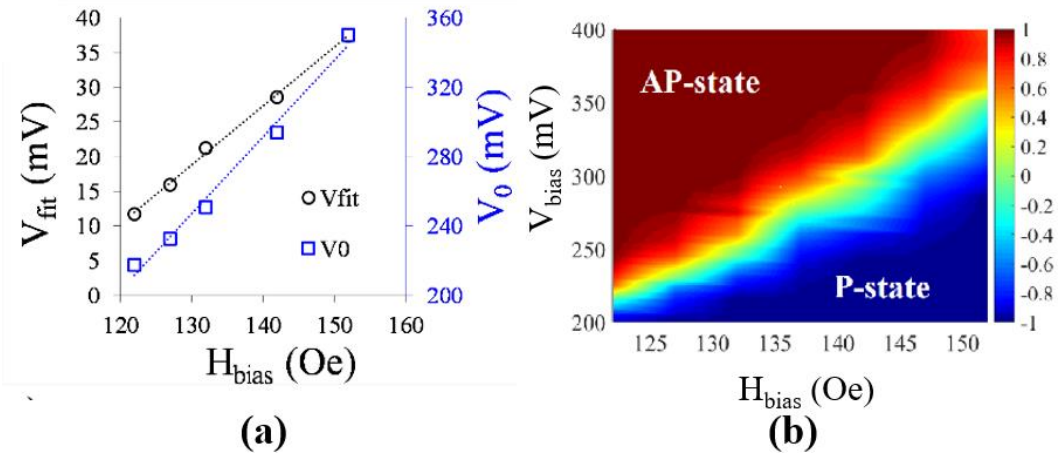


Fig. 3.10. (a) AP-rate versus V_{bias} at several fixed bias fields, **(b)** fitting parameters V_{fit} and V_0 with respect to H_{bias} and **(c)** contour plot of the AP-rate over all the bias fields and bias voltages tested. Images extracted from [78].

The next step in this analysis is to express $\langle r \rangle$ in terms of τ_{AP} and τ_P, using Eq. 3.7. From this expression, $\langle r^* \rangle$ can be expressed in terms of Δ_{AP} and Δ_P, as shown in Eq. 3.8. When coupling the expressions shown in Eq. 3.6 and Eq. 3.8, Δ_{AP} and Δ_P can be expressed in terms of V_{bias} and H_{bias} directly, independent Eq. 3.5.

$$\langle r \rangle = \frac{\langle \tau_{AP} \rangle}{\langle \tau_{AP} \rangle + \langle \tau_P \rangle} \leftrightarrow \langle r^* \rangle = 2\frac{\langle \tau_{AP} \rangle}{\langle \tau_{AP} \rangle + \langle \tau_P \rangle} - 1 = \frac{\langle \tau_{AP} \rangle - \langle \tau_P \rangle}{\langle \tau_{AP} \rangle + \langle \tau_P \rangle} \tag{3.7}$$

$$\langle r^* \rangle = \frac{\exp(\Delta_{AP} - \Delta_P) - 1}{\exp(\Delta_{AP} - \Delta_P) + 1} = \tanh\left(\frac{\Delta_{AP} - \Delta_P}{2}\right) \leftrightarrow \Delta_{AP} - \Delta_P = \ln\left(\frac{1 + \langle r^* \rangle}{1 - \langle r^* \rangle}\right) \tag{3.8}$$

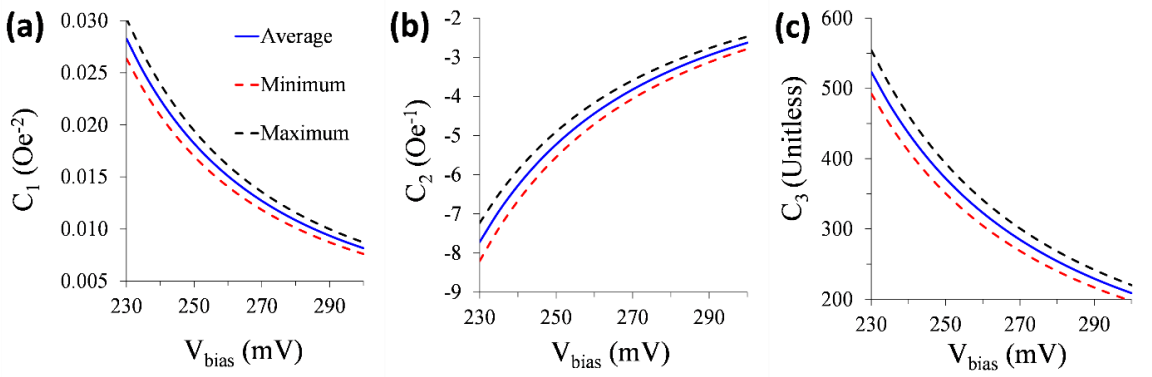


Fig. 3.11. Calculations for the fitting coefficients **(a)** C_1, **(b)** C_2, and **(c)** C_3 in Eq. 4.9 vs V_{bias}. For each plot, the black and red dotted lines represent the maximum and minimum error bounds (95% confidence) and the solid blue line is the average value within the error bounds. Images were extracted from [78].

The next step is to express the AP-rate data in terms of $H_{K\uparrow}$, $H_{K\downarrow}$, $V_{C0\uparrow}$, and $V_{C0\downarrow}$ by combining Eq. 3.5 and Eq. 3.8. Using Eq. 3.5, $\Delta_{AP} - \Delta_P$ was expressed in terms of H_{bias} with Eq. 3.9, which is a second degree polynomial with coefficients C_1, C_2, and C_3. The full expressions for the dependence of these coefficients on V_{bias} is shown in Eq. 3.10. Figures 3.11(a) – 3.11(c) show the calculations of these coefficients, which were determined from second-order polynomial fits of $\Delta_{AP} - \Delta_P$ with H_{bias} at fixed V_{bias} values from 230 to 300 mV. Each curve fitted contained a total of 17 data points, with $\langle r^* \rangle$ = -0.8

to +0.8 in steps of 0.1, which were converted to $\Delta_{AP} - \Delta_P$ using Eq. 3.8. The results for V_0 and V_R presented in Fig. 3.10(a) were used to find H_{bias} for each combination of $\langle r^* \rangle$ and V_{bias}. These calculations allow for $H_{K\uparrow(\downarrow)}$ and $V_{C0\uparrow(\downarrow)}$ to be determined at each combination of H_{bias} and V_{bias}.

$$\Delta_{AP} - \Delta_P = C_1 H_{bias}^2 + C_2 H_{bias} + C_3 \tag{3.9}$$

$$\left\{ \begin{array}{c} C_1 = \left(\frac{\Delta_{0\uparrow}}{H_{K\uparrow}^2} - \frac{\Delta_{0\downarrow}}{H_{K\downarrow}^2}\right) + \left(\frac{\Delta_{0\uparrow}}{H_{K\uparrow}^2 V_{C0\uparrow}} + \frac{\Delta_{0\downarrow}}{H_{K\downarrow}^2 V_{C0\downarrow}}\right) V_{bias} \\ C_2 = -2\left[\left(\frac{\Delta_{0\uparrow}}{H_{K\uparrow}} + \frac{\Delta_{0\downarrow}}{H_{K\downarrow}}\right) + \left(\frac{\Delta_{0\uparrow}}{H_{K\uparrow} V_{C0\uparrow}} - \frac{\Delta_{0\downarrow}}{H_{K\downarrow} V_{C0\downarrow}}\right) V_{bias}\right] \\ C_3 = (\Delta_{0\uparrow} - \Delta_{0\downarrow}) + \left(\frac{\Delta_{0\uparrow}}{V_{C0\uparrow}} + \frac{\Delta_{0\downarrow}}{V_{C0\downarrow}}\right) V_{bias} \end{array} \right\} \tag{3.10}$$

The next step in this analysis is to evaluate the dependence of $H_K^{\uparrow(\downarrow)}$ and $V_{C0}^{\uparrow(\downarrow)}$ on V_{bias} and H_{bias}. Since there are four variables being solved for and only three equations, $H_K^{\uparrow(\downarrow)}$ and $V_{C0}^{\uparrow(\downarrow)}$ cannot be solved for directly with Eq. 3.10 unless at least one of these variables is calculated using a separate method. Rearranging Eq. 3.5 to solve for H_{bias} shows that $H_{bias} \propto (\Delta_{AP})^{1/2}$ and that $H_K^{\uparrow}$ can be determined from the intercept through a linear fit of a H_{bias} vs $(\Delta_{AP})^{1/2}$ plot. This analysis was done at several fixed bias voltages ranging from 230mV to 300mV. Note that this can also be done for $H_K^{\downarrow}$, however Δ_P has a very weak dependence on H_{bias}, which significantly impacts the accuracy of the calculations.

The results for $H_K^{\uparrow}$ calculations are shown in Fig. 3.12(a), which shows that $H_K^{\uparrow}$ increases linearly with V_{bias} at a rate of ~0.11 Oe/mV. Furthermore, the intercept of the linear fit is 164.9 Oe, which is approximately equal to the anisotropy field measured at V_{bias} = 0, which was 166 Oe. These results are consistent with previous experiments on VCMA, which demonstrate that magnetic anisotropy has a linear dependence on voltage and increases when the polarity of the voltage is set to favor the AP-state (such as in our

experiment) [100]. From the $H_K^{\uparrow}$ calculations, calculations shown in Fig. 3.11(a-c), and Eqn. 4.10, $H_K^{\downarrow}$ can be determined. These calculations revealed that $H_K^{\downarrow} \approx H_K^{\uparrow}$ at all bias voltages, which implies that $\Delta_{0\uparrow} \approx \Delta_{0\downarrow}$, since $\Delta_{0\uparrow(\downarrow)} = H_K^{\uparrow(\downarrow)} M_S Vol/2k_bT$.

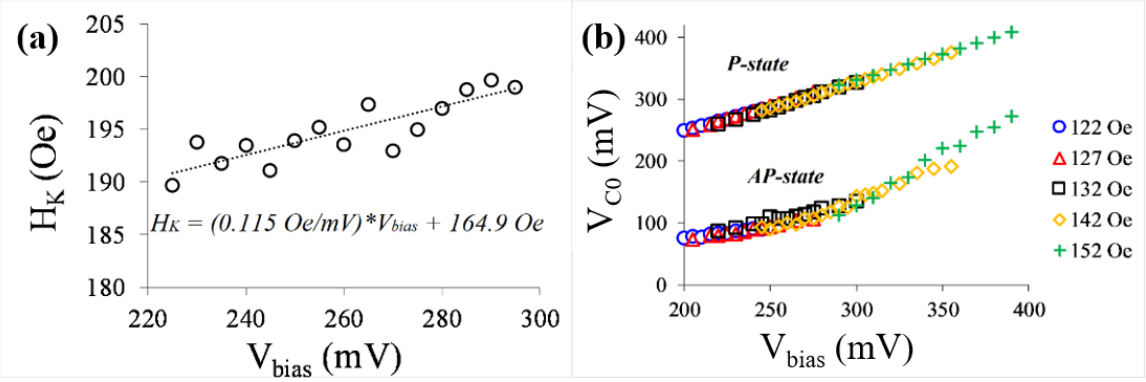


Fig. 3.12. (a) $H_K^{\uparrow}$ vs V_{bias} plots determined from linear fits of H_{bias} vs $(\Delta_{AP})^{1/2}$ plots. **(b)** Plots of the intrinsic critical switching voltage with respect to voltage bias. These values were calculated using the raw Δ_{AP} and Δ_P measurements as well as the H_K calculations. Images were extracted from [78].

Calculations for V_{C0} needs more careful consideration since it is dependent on both V_{bias} and H_{bias}. Furthermore, the influence of H_{bias} and V_{bias} on V_{C0} may be noticeably different for AP-state and P-state components. The definition of V_{C0} is shown in Eq. 3.11, which shows that $V_{C0}^{\uparrow(\downarrow)} \propto H_{eff}^{\uparrow(\downarrow)}$ where $H_{eff}^{\uparrow(\downarrow)} = \pm H_{bias} + H_K^{\uparrow(\downarrow)} + 2\pi M_S$. Note that Eq. 3.11 is the same as Eq. 2.1, except that V_{C0} in Eq. 3.11 is dependent on the resistance state of the MTJ. It should be noted that $2\pi M_S \gg H_{bias}$, so for now, we will consider the changes in H_{bias} to have a negligible effect on V_{C0}.

$$V_{C0\uparrow(\downarrow)} = \frac{2e\alpha M_S Vol R_{AP(P)}}{\eta \hbar} H_{eff\uparrow(\downarrow)} \tag{3.11}$$

There are two variables in Eq. 3.11 (other than H_K) that are dependent on V_{bias} which are R_{AP} and the STT efficiency (η). The change in R_{AP} with V_{bias} is a consequence of the MTJ's rectifying behavior, which means that the TMR ratio is also voltage dependent,

which can be described using the expression $T(v_t) = T_0(1 + v_t^2)^{-1}$. In this expression, $T(v_t)$ is the TMR ratio at a given V_{bias}, T_0 is the TMR ratio at zero bias voltage, and $v_t = V_{bias}/V_H$, where V_H is the voltage at which the TMR becomes half of its original value, which is 440 mV for the device tested here. Since TMR is dependent on V_{bias}, η will also be dependent on V_{bias} since η is a function of the spin polarization (P) of the current through the MTJ, where $P = [T(v_t)/(2 + T(v_t))]^{1/2}$ and $\eta = 2P/(1 + P^2)$. Eq. 3.11 can be used to predict the influence of V_{bias} on $V_{C0\uparrow}$ and $V_{C0}^{\downarrow}$ since $V_{C0}^{\uparrow} \propto R_{AP}/\eta$ and $V_{C0}^{\downarrow} \propto \eta^{-1}$.

The initial calculations for $V_{C0}^{\uparrow}$ indicated that $V_{C0}^{\uparrow} \propto V_{bias}^2$. To check the validity of this, $V_{C0}^{\uparrow}$ values were re-calculated for each H_{bias} set using our Δ_{AP} and Δ_P measurements and Eq. 3.5. Note that the initial calculations ignored the influence of H_{bias} on V_{C0}, however, these calculations do not. Figure 3.12(b) shows the calculations for $V_{C0}^{\uparrow}$ and $V_{C0}^{\downarrow}$, both of which show a linear relation with V_{bias}. Note that H_{bias} does not significantly influence $V_{C0}^{\uparrow}$ at any V_{bias}, however, it does affect the rate at which $V_{C0}^{\uparrow}$ increases with V_{bias}. At H_{bias} = 60 Oe, $V_{C0}^{\uparrow}$ increases with V_{bias} at a rate of 0.35 mV/mV, however at H_{bias} = 90 Oe, this rate increases to 1.63 mV/mV. On the other hand, the rate at which $V_{C0}^{\downarrow}$ increases with V_{bias} is not significantly influenced by H_{bias}.

From a mathematical perspective, $dV_{C0}^{\uparrow}/dV_{bias}$ is dependent on $d[R_{AP}/\eta]/dV_{bias}$ ($d\eta^{-1}/dV_{bias}$ in the P-state) which is multiplied by H_{bias} in Eq. 3.11. In the AP-state, increasing H_{bias} causes $d[R_{AP}/\eta]/dV_{bias}$ to increase, however, in the P-state, increasing H_{bias} will cause $d\eta^{-1}/dV_{bias}$ to decreases, which appears to be negligible based on the results in Fig. 3.12(b). These results suggest that the reason our method can tune Δ_{AP} and Δ_P separately is because of the asymmetrical effect that V_{bias} and H_{bias} has on $V_{C0}^{\uparrow}$ and $V_{C0}^{\downarrow}$.

To illustrate how H_K and V_{C0} affects the dependence of $\Delta_{AP(P)}$ on H_{bias} and V_{bias}, Eq.

3.5 will be separated into two components, h and ν. For the purpose of this analysis, these components are defined as $h_{\uparrow(\downarrow)} = (1 \mp H_{bias}/H_{K\uparrow(\downarrow)})^2$ and $\nu_{\uparrow(\downarrow)} = 1 \pm V_{bias}/V_{C0\uparrow(\downarrow)}$. Since $\Delta_{AP(P)} = \Delta_{0\uparrow(\downarrow)} * h_{\uparrow(\downarrow)} * \nu_{\uparrow(\downarrow)}$, the percent change in $\Delta_{AP(P)}$ can be expressed as a sum of the percent differences in $h_{\uparrow(\downarrow)}$ and $\nu_{\uparrow(\downarrow)}$, which were calculated with respect to V_{bias} ($d_V\nu$, d_Vh, and $d_V\Delta_{AP(P)}$) and with H_{bias} ($d_H\nu$, d_Hh, and $d_H\Delta_{AP(P)}$). Calculations for $d_V\nu$, d_Vh, and $d_V\Delta_{AP(P)}$ were obtained at all H_{bias} values and $d_H\nu$, d_Hh, and $d_H\Delta_{AP(P)}$ were obtained at V_{bias} = 220, 240, 260, 280, and 300 mV. The values for H_{bias} and V_{bias} used for calculating these percent differences are determined using data shown in Fig. 3.10(a) for ⟨r*⟩ between -0.8 to +0.8.

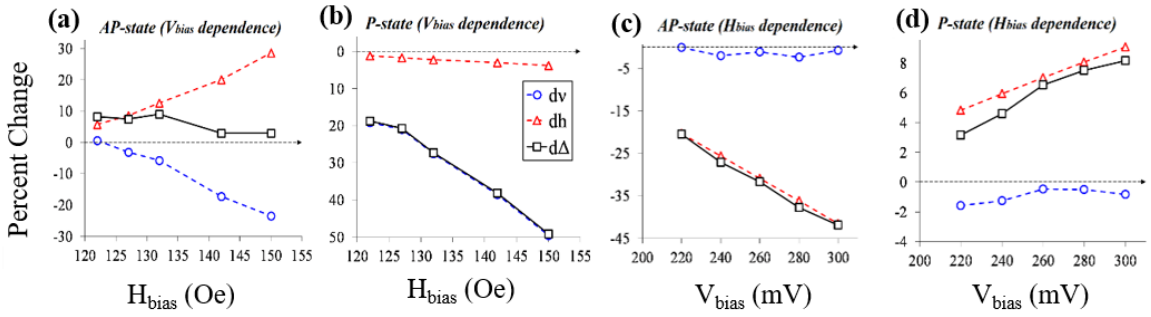


Fig. 3.13. Percent changes in ν, h, and $\Delta_{AP/P}$ with respect to V_{bias}, where the **(a)** AP-state and **(b)** P-state components are calculated at fixed fields of 122, 127, 132, 142, and 152 Oe, and with respect to H_{bias}, where the **(c)** AP-state and **(d)** P-state components are calculated at fixed bias voltages of 220, 240, 260, 280, and 300 mV. In each case, the range of V_{bias} or H_{bias} values in which these percent changes are being calculated are set in order to produce AP-rates between -0.8 to +0.8, which are determined using our fitting parameters of the AP-rate data. Images were extracted from [78].

The AP- and P-state component of the $d_V\nu$, d_Vh, and $d_V\Delta_{AP(P)}$ are shown in Fig. 3.13(a) and (b), respectively. In the AP-state, $h_\uparrow$ increases by 5.8% and 28.5% for H_{bias} = 60 Oe and 90 Oe, respectively. Conversely, $\nu_\uparrow$ decreases by 3.1% and 23.5% for H_{bias} = 65 Oe and 90 Oe, respectively. Surprisingly, $\nu_\uparrow$ increases slightly by 0.7% when H_{bias} = 60 Oe. Note that Eq. 3.5 appears to suggest that $\nu_\uparrow$ should increases with V_{bias} however, when the

Vbias dependence of $V_{C0\uparrow}$ is considered, then Eq. 3.5 shows that increasing V_{bias} could either increase or decrease $\nu_{\uparrow}$ depending on $dV_{C0}^{\uparrow}/dV_{bias}$. Fig. 3.13(a) shows that as H_{bias} increases, changes in $\nu_{\uparrow}$ and $h_{\uparrow}$ cancel each other out, which explains why Δ_{AP} has a negligible dependence on V_{bias}.

Figure 3.13(b) shows that $d_v\Delta_P$ is primarily dominated by $d_v\nu_{\downarrow}$ since $|d_vh_{\downarrow}|$ is much smaller than $|d_v\nu_{\downarrow}|$, which explains why Δ_P is not strongly influenced by H_{bias}. Unlike the AP-state, $\nu_{\downarrow}$ will always decrease with V_{bias}, regardless of $dV_{C0}^{\downarrow}/dV$bias. Note that the increase in $|d\nu_{\downarrow}|$ from 19.1% to 49.5% for H_{bias} = 60 Oe and 90 Oe, respectively, may seem counterintuitive since Fig. 3.12(b) shows that $V_{C0}^{\downarrow}$ is not affected by H_{bias}. However, this is likely due to a broader range of V_{bias} at larger H_{bias} values.

The AP- and P-state components of $d_H\nu$, d_Hh, and $d_H\Delta_{AP(P)}$ are shown in Fig. 3.13(c) and (d), respectively. Fig. 3.13(c) shows that $d_H\nu_{\uparrow}$ = -0.1% and -2.1% for V_{bias} = 220 mV and 300 mV, respectively, which are insignificant when compared to changes in $h_{\uparrow}$, which decreases by 20.4% and 41.5% when V_{bias} = 220 mV and 300 mV, respectively. The increase of $|d_Hh_{\uparrow}|$ with H_{bias} for increasing V_{bias} is attributed to a larger range of H_{bias} values at larger V_{bias} and increases in H_K with V_{bias}. Since H_{bias} does not strongly change $V_{C0}^{\uparrow}$ at any given V_{bias} value, it is expected that H_{bias} would not influence $\nu_{\uparrow}$ significantly.

In the P-state, when V_{bias} increases from 220 mV to 300 mV, $h_{\downarrow}$ increases from 4.8% to 9.1% whereas $\nu_{\downarrow}$ decreases from 4% to 0.5%. These results show that $d_Hh_{\downarrow} > 0$, which is predicted in Eq. 3.5, however, since increasing V_{bias} will cause H_K to increase, increasing V_{bias} will also cause $d_Hh_{\downarrow}$ to decrease. The reason that our results show $d_Hh_{\downarrow}$ increasing with V_{bias} is likely due to an increase in the range of H_{bias} between $\langle r^* \rangle$ = -0.8 and +0.8. $V_{C0\downarrow}$ should decrease slightly with H_{bias}, as predicted by Eq. 3.11, which would result in a

decrease in $\nu_{\downarrow}$. However, since H_K increases with V_{bias}, changes in H_{bias} would have a negligible effect on $\nu_{\downarrow}$, which agrees with the results shown in Fig. 3.13(d). At low voltages, $|d_H h_{\downarrow}| \approx |d_H \nu_{\downarrow}|$, but have opposite signs resulting in a negligible change in Δ_P. At higher voltages, $|d_H h_{\downarrow}| >> |d_H \nu_{\downarrow}|$, meaning that $d_H \Delta_P$ is primarily influenced by $d_H h_{\downarrow}$, resulting in an increase in more significant increases in Δ_P.

In summary, the cause for the two degrees of tunability capability observed is primarily the effect that V_{bias} and H_{bias} has on the critical switching voltage, which differed between the AP- and P-states. The analysis described in this section demonstrates both applicability and predictability to our dual-biasing method.

3.5.3. Dependence of Intrinsic Device Properties

One of the novel capabilities of dual-biasing is tuning average switching rate over several orders of magnitude. However, the maximum switching rate achieved from the device in section 3.3 was ~1 MHz, which is 3 orders of magnitude lower than the ideal switching rates of 1 GHz desired for most p-bit designs. However, thermally stable MTJs are not necessary to avoid the effects of device variations in applications involving asynchronous random telegraphic signals since the dual-biasing method can also serve as a method of avoiding these effects through on-board corrections. Therefore, the average switching rates can increase by several orders of magnitude by using sMTJs rather than stable MTJs. In this section, I will discuss all the trade-offs of dual-biasing on sMTJs versus thermally stable MTJs in terms of 1) transfer characteristics, 2) switching speed, and 3) degree of separation between AP and P-state dwell time tunability to determine if two degrees of tunability is still achieved in dual-biased sMTJs.

The data presented in this section was obtained from dual-biasing measurements on 10 MTJs with Δ_0 varying from 10.5 to 35.6. The stack structure of these MTJs is PtMn(15)/SAF/MgO(0.8)/ $Co_{60}Fe_{20}B_{20}$(2.0), where layer thicknesses are in nanometers and the SAF structure comprises of $Co_{70}Fe_{30}$(2.5)/Ru(0.85)/$Co_{40}Fe_{40}B_{20}$(2.4). Each MTJ had an RA product of 3.5 $\Omega \cdot \mu m^2$ and a TMR ratio that ranged between 135% to 150%.

Table 3.2. Intrinsic properties and characterization of all MTJs tested.

Category	MTJ no.	Δ_0	R_P (Ω)
Superparamagnetic MTJs (sMTJs)	1	10.5	1360
	2	11.4	1330
	3	12.0	1100
Thermally unstable MTJs	4	19.0	1290
	5	18.8	1015
Semi-stable MTJs	6	24.9	1260
	7	26.4	1090
	8	28.2	1250
Thermally stable MTJs	9	34.2	570
	10	35.6	620

Table 3.2 lists the intrinsic thermal stability factors (Δ_0) and P-state resistances (R_P) for all 10 MTJs tested. R_P values were obtained from R-H plots and Δ_0 values were calculated from P_{SW} measurements. For MTJs with H_C > 10 Oe (MTJs 6 – 10), synchronized P_{SW} measurements were carried out and Δ_0 was determined using the same process as the one described in section 3.2, except the pulse widths were ≥ 500 ns so that thermal activation model can be used [81]. For MTJs with R-H curves showing random fluctuations near the switching fields, Δ_0 was calculated using dwell time measurements from time-domain data. It should be noted that these measurements should not be confused with the dual-biasing measurements since H_{bias} and V_{bias} do not follow the proper parameters of magnitude and

orientation of the biases that define the dual-biasing method. Instead, τ_{AP} and τ_P are measured at small bias currents (I_{bias}) and at H_{bias} values near the switching fields, which were used to calculate Δ_{AP} and Δ_P. H_{stray} was determined by finding H_{bias} where $\Delta_{AP} = \Delta_P$ at $I_{bias} = 0$, where Δ_0 is the value where Δ_{AP} and Δ_P intersect.

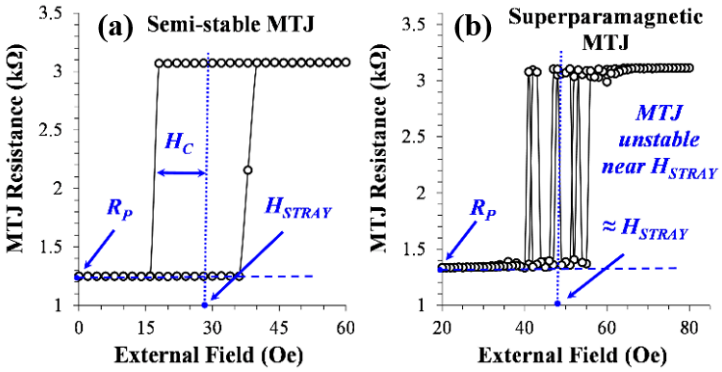


Fig. 3.14. Resistance vs field plots for (**a)** MTJ 6 and (**b)** MTJ 2 (see table 3.2). Images were extracted from [79].

In this analysis, MTJs are divided into four categories based on Δ_0 which were 1) superparamagnetic MTJs (sMTJs), 2) thermally unstable MTJs, 3) semi-stable MTJs, and 4) thermally stable MTJs. These categories are determined by the intrinsic retention times of the MTJs (determined from Δ_0), which are < 1 ms, ~100 ms, ~2-20 minutes, and ~7-33 days for sMTJs, unstable MTJs, semi-stable MTJs and stable MTJs, respectively.

AP-rate curves from the dual-biasing measurements, examples of which are shown in Fig. 3.15(a), were acquired for all MTJs with respect to both V_{bias} and H_{bias} and then characterized in terms of their center and width. The centers of the AP-rate curves are defined as the V_{bias} value which produced an AP-rate of 0.5 (V_0) and the widths of the AP-rate curves are defined as the V_{bias} range between $\langle r \rangle \approx 0.75$ and at $\langle r \rangle \approx 0.25$ (V_R). Note that wide transfer curves are desired for devices as tunable random signals generators

because it improves input resolution and noise tolerance of output. V_0, and V_R were obtained by fitting the AP-rate curves to the same hyperbolic tangent function as in the previous section (recall Eqn. 3.8).

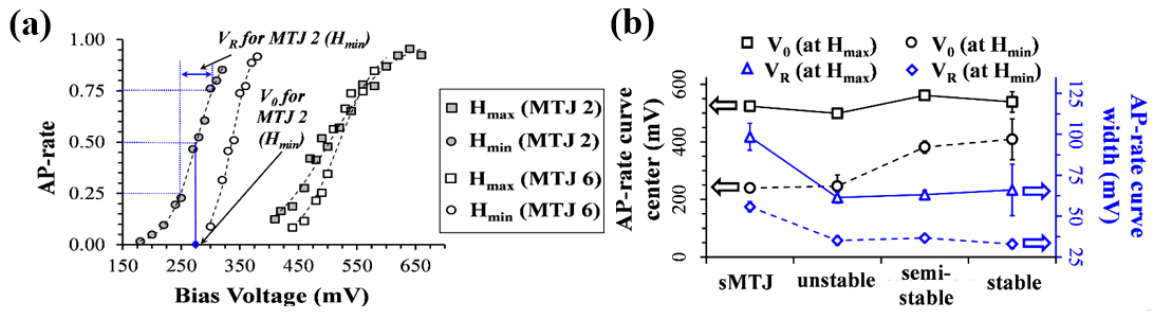


Fig. 3.15. (a) AP-rate versus bias voltage curves at H_{min} and H_{max} for a semi-stable MTJ and an sMTJ (MTJs 6 and 2 from table 3.2). H_{min} = 60 Oe and 30.5 Oe and H_{max} = 80 Oe and 55.5 Oe for MTJ 6 and 2, respectively. **(b)** Line plots showing the center (V_0) and width (V_R) of the voltage dependent transfer curves for each MTJ category at H_{min} and H_{max}. Images were extracted from [79].

The line plots shown in Fig. 3.15(b) shows the average V_0 and V_R measured at minimum and maximum H_{bias} (H_{min} and H_{max}, respectively) for each MTJ category. For this analysis, H_{min} is defined as the smallest H_{bias} value needed to generate stochastic switching in the MTJ and H_{max} is defined as the largest H_{bias} value tested where the AP-rate vs V_{bias} curve did not saturate (AP-rate saturation will be discussed later). This data shows that V_0 increases as Δ_0 increases at H_{min}. However, at H_{max}, V_0 is nearly constant between all sets. This cut-off in V_0 can be attributed to the rectifying properties of MTJs, where V_{bias} causes their TMR ratio to decrease. This creates an upper limit on V_{bias} for the dual-biasing method since it relies on current fluctuations caused by resistance changes (recall previous section). Therefore, the limiting factor for V_{bias} is the TMR ratio and not Δ_0, which is why V_0 is not dependent on Δ_0 at H_{max} in Fig. 3.15(a). The difference in V_0 at H_{min} and H_{max} is larger for

sMTJs than any other MTJs, which demonstrates that dual-biasing of sMTJs provides more flexibility and control over the AP-rate. The V_R data in Fig. 3.15(b) show that dual-biased sMTJs produced wider AP-rate curves. There were no significant differences in V_R between unstable, semi-stable, and stable MTJs, but was at least 2 times larger for sMTJs at both H_{min} and H_{max}. Our data shows that V_R increases with H_{bias} for all MTJs.

In terms of transfer curve characteristics, sMTJs showed superior performance to all others. Not only do they demonstrate better energy efficiency with lower minimum V_0 compared to the other three categories, but they also have the same maximum V_0 values, thus providing a larger tunability range. Furthermore, sMTJs had significantly wider transfer curves than the others, demonstrating superior output resolution and noise tolerance. Note that V_R can be improved by increasing H_{bias} for any MTJ.

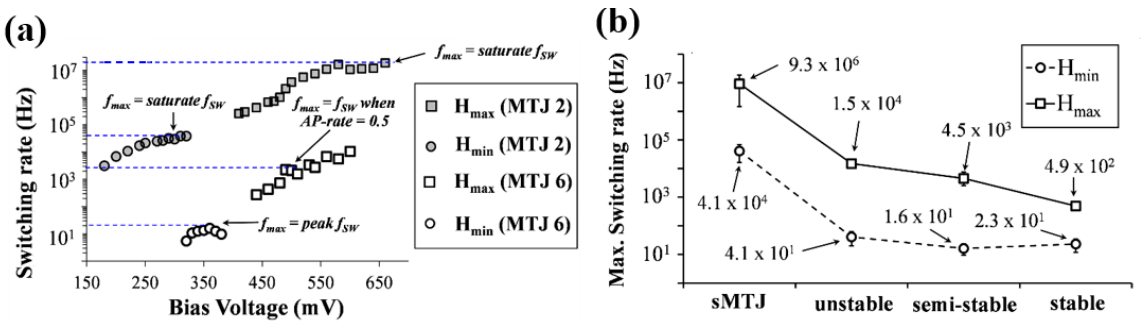


Fig. 3.16. (a) Average switching rate vs bias voltage at H_{max} and H_{min} for MTJs 6 and 2 (see table 3.2) and **(b)** Maximum switching rate (f_{max}) achieved for each MTJ category. Images were extracted from [79].

Next, switching rate data was evaluated and the maximum f_{SW} (f_{max}) achieved was determined at both H_{min} and H_{max} for each MTJ. Figure 3.16(a) shows examples of f_{SW} vs V_{bias} plots at H_{min} and H_{max}, which illustrates three different behaviors of f_{SW} with V_{bias}. One is that f_{SW} reaches a maximum and then decreases, the second is that f_{SW} increases with V_{bias} until reaching a saturation point, and the third is that f_{SW} increases exponentially

with V_{bias}. Therefore, the method of calculating f_{max} depended on the behavior of f_{SW} vs V_{bias}. If f_{SW} peaked or saturated, then f_{max} was simply f_{SW} at the peak or saturation value. However, if f_{SW} increases exponentially with V_{bias}, then f_{max} was defined as the f_{SW} value when the AP-rate = 0.5. This is because the largest f_{SW} occurred at V_{bias} values that produce an AP-rate > 0.9, which is not particularly useful for most applications.

Figure 3.16(b) shows the line plots of f_{max} at H_{min} and H_{max} for all four categories. Unsurprisingly, f_{max} was the largest for sMTJs, which achieved $f_{max} \approx 10^4$ Hz – 10^7 Hz. The remaining three MTJ categories showed $f_{max} \approx$ 10 Hz – 100 Hz at H_{min}. This is because time-domain data was not collected if the dwell times > 50ms due to bandwidth limitations. Note that f_{max} at H_{min} for sMTJs and unstable MTJs is on the same order of magnitude as their intrinsic dwell times whereas semi-stable and stable MTJs required larger biases to reach 10 Hz. These plots show that unstable MTJs and sMTJs have a broader tuning range of f_{max} than semi-stable and stable MTJs.

Figures 3.17(a-b) shows that Δ_{AP} and Δ_{P} are influenced by V_{bias} and H_{bias} to some degree. However, Fig. 3.17(a) shows that changes in Δ_{P} with V_{bias} are more than 6X larger than changes in Δ_{AP} over the same V_{bias} range for both the semi-stable MTJs and sMTJs, which means that P-state pulse widths decrease by 2 – 3 orders of magnitude while AP-state pulse widths do not even increase by 1 order. The opposite behavior is shown in Fig. 3.17(b), where changes in Δ_{AP} with H_{bias} are approximately 2 – 3X larger than changes in Δ_{P} over the same H_{bias} range, meaning that that AP-state pulse widths decrease by approximately 2 orders of magnitude while P-state pulse widths increase by less than 1 order of magnitude. These discrepancies are large enough to claim two degrees of tunability, however, it also means that the two degrees of tunability feature can be

quantified to some degree.

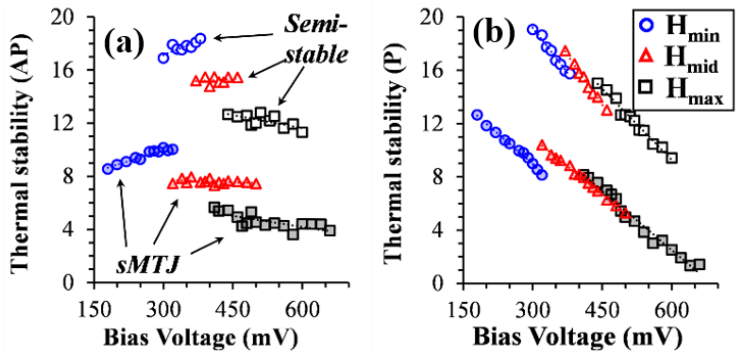


Fig. 3.17. Effective thermal stability factor versus bias voltage at H_{max}, H_{min}, and H_{mid} for **(a)** AP-state and **(b)** P-state components for MTJs 6 and 2 (see table 3.2). Note that H_{mid} = 70 Oe for MTJ 6 and 45.5 Oe for MTJ 2.

In our analysis, two degrees of tunability was quantified using a metric we refer to as the degree of separation in Δ_P and Δ_{AP} tunability, which we divided into V_{bias} dependence and H_{bias} dependence components (D_V and D_H, respectively). D_V and D_H are expressed using slopes of Δ_P and Δ_{AP} with V_{bias} and H_{bias}, as seen in Eq. 3.12. D_V is the order of magnitude decrease measured in τ_P for every 1 order of magnitude increase in τ_{AP} with V_{bias} and D_H is the order of magnitude decrease measured in τ_{AP} for every 1 order of magnitude increase in τ_P with respect to H_{bias}. Note that $\Delta_{AP(P)}$ is proportional to H_{bias}^2 according to the Néel-Brown model (recall Eq. 3.5), therefore, D_H is defined in terms of the changes in $\sqrt{\Delta_{AP}}$ and $\sqrt{\Delta_P}$ with H_{bias}.

$$D_V = \left|\frac{d\Delta_P/dV_{bias}}{d\Delta_{AP}/dV_{bias}}\right| \; and \; D_H = \left|\frac{d\sqrt{\Delta_{AP}}/dH_{bias}}{d\sqrt{\Delta_P}/dH_{bias}}\right| \tag{3.12}$$

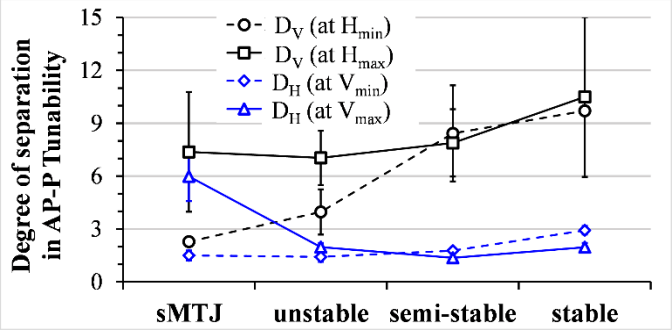


Fig. 3.18. Degree of separation between Δ_{AP} and Δ_P tunability with voltage and field (D_V and D_H, respectively) for each MTJ category. D_V represents the order of magnitude decrease in τ_P for every 1 order of magnitude increase in τ_{AP}, and D_H represents the order of magnitude decrease in τ_{AP} for every 1 order of magnitude increase in τ_P. Image extracted from [79].

The line plots in Fig. 3.18 show D_V and D_H for all four MTJ categories, where D_V for both H_{min} and H_{max} are shown and D_H for both V_{min} and V_{max} are shown. This plot shows that D_V improves as Δ_0 increases, particularly at H_{min}. Note that this trend not as significant at H_{max} since D_V improves with from H_{min} to H_{max} for sMTJs and unstable MTJs but does not change significantly with H_{bias} for semi-stable and stable MTJs. This figure indicates that the advantages of dual-biasing sMTJs shown in Fig. 3.16(b) and 3.15(b) comes at the cost of smaller D_V. However, $D_V > 2$ for sMTJs at all H_{bias} values meaning that the two degrees of tunability capability is still present. The D_H curves in Fig. 3.18 shows that stable MTJs have the largest D_H at V_{min} whereas sMTJs have the largest D_H at V_{max}. Therefore, even though D_V is reduced in sMTJs, two degrees of tunability is not sacrificed since D_H is improved.

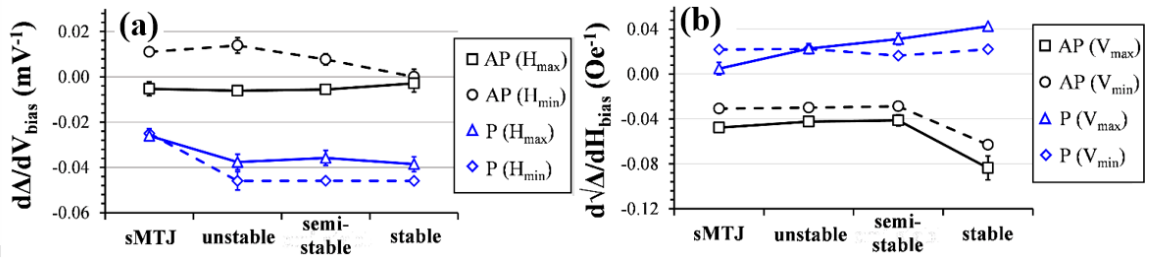


Fig. 3.19. Slopes Δ_{AP} and Δ_P with respect to **(a)** V_{bias} and **(b)** H_{bias} for each MTJ category. Images extracted from [79].

The behavior of D_V and D_H can be explained from the slopes of Δ_{AP} and Δ_P with V_{bias} and H_{bias}, which are shown in the line plots in Fig. 3.19(a) and 3.19(b). Figure 3.19(a) shows that $d\Delta_{AP}/dV_{bias}$ is approximately 10X larger for sMTJs and unstable MTJs then for stable MTJs at H_{min}, which explains why D_V is larger for stable MTJs. Furthermore, the smaller D_V for sMTJs can also be explained by the fact that $|d\Delta_P/dV_{bias}|$ is ~2X lower for sMTJs than for all other categories at both H_{min} and H_{max}. One interesting observation from this plot is that $d\Delta_{AP}/dV_{bias}$ changes sign from H_{min} to H_{max} for all MTJs. This was also observed in section 3.5.2, where H_{bias} caused dV_{C0}/dV_{bias} to increase in the AP-state but not in the P-state [78]. This suggests that MTJs of Δ_0 value has an H_{bias} where $d\Delta_{AP}/dV_{bias} \approx 0$, resulting in a near infinite D_V, which explains the large error bars at H_{max} seen in Fig. 3.19(a). Note that optimum control over Δ_P and Δ_{AP} occurs in MTJs where the influence of H_{bias} on $d\Delta_{AP}/dV_{bias}$ is minimized and the stable MTJs had the smallest variation in $d\Delta_{AP}/dV_{bias}$ from H_{min} to H_{max} along with the largest D_V values. It should be noted that the VCMA efficiency increases as MTJ size decreases [89]. In this study, sizes of the stable MTJs are typically larger than the sMTJs (see the R_P values in table 3.2), therefore, the weak influence of H_{bias} on $d\Delta_{AP}/dV_{bias}$ seen in stable MTJs may be a result of device size and not Δ_0.

Figure 3.19(b) shows that $|d\sqrt{\Delta_{AP}}/dH_{bias}|$ is approximately 2 – 3X larger for stable MTJs than all other MTJs at both V_{min} and V_{max}. The large D_H values seen for stable MTJs at V_{min} is explained by the combination of the large $|d\sqrt{\Delta_{AP}}/dH_{bias}|$ values and $d\sqrt{\Delta_P}/dH_{bias}$ being nearly constant among all MTJs at V_{min}. However, V_{bias} had a different influence on $d\sqrt{\Delta_P}/dH_{bias}$ for each MTJ category. From V_{min} to V_{max}, $d\sqrt{\Delta_P}/dH_{bias}$ decreased significantly for sMTJs, remained constant for unstable MTJs, and increased for semi-stable and stable MTJs. The change in $d\sqrt{\Delta_P}/dH_{bias}$ from V_{min} to V_{max} is partially due to the influence of V_{bias} on the anisotropy field (H_K) and V_{C0} since the VCMA efficiency increases as the MTJs thermal stability factor increases [89].

In this section, we demonstrated that dual-biasing can achieve two degrees of tunability for both sMTJs and stable MTJs. Even though sMTJs could did not achieve the same degree of separation in low-state tunability with voltage, the two degrees of tunability capability was not sacrificed. Additionally, dual-biased sMTJs improved several other performance metrics such as larger degree of separation in high-state tunability with field, enhanced control over average output and switching rate, and better noise immunity in transfer curves.

3.6 Summary

In this chapter, I introduced a method of generating random telegraphic switching signals in magnetic tunnel junctions for probabilistic bits and stochastic computing applications. This dual-biasing method involves applying an external magnetic field that favors P-state orientation and a DC bias voltage that favors AP-state orientation. Through my analysis, I determined that dual-biased MTJs has several advantages over single-biased MTJs for probabilistic computing applications. The most important feature of the dual-

biasing method is the two-degrees of tunability capability, which creates the possibility for networks of interconnected probabilistic bits to overcome the effects of device variations through on-board corrections. Alternatively, two-degrees of tunability could double the information capacity in stochastic spiking neural signals since spike count and average output are independent variables.

First, I demonstrated the tunability of the AP-rate as well as the two degrees of tunability capability of the dual-biasing method on a single MTJ by measuring the AP- and P-state dwell times at multiple bias fields and bias voltages. Changing the bias field caused the AP-state dwell times to change by several orders of magnitude while the P-state dwell time only changed by less than one order of magnitude. Alternatively, changing the bias voltage caused the P-state dwell times to change by several orders of magnitude while barely changing the AP-state dwell times. Furthermore, this difference in the changes in AP- and P-state dwell times with changes in the bias voltage increased as the bias field increased.

After testing the dual-biasing method on a single MTJ device, I performed dual-biasing on two MTJ devices connected in series. This process generates telegraphic switching signals in bipolar format with three logic states: -1, 0, and +1. The results presented demonstrate that at sufficiently large bias fields, the average output of the signals generated become fixed at zero while still passing three tests from the NIST statistical test suite, regardless of the bias voltage. I determined that the two degrees of tunability capability of the dual-biasing method was the explanation behind these results. These results indicate that true random number generators with bipolar encoding can be built using only two MTJs connected in series through dual biasing.

After presenting the single MTJ results and the bipolar signal results from two MTJs connected in series, I attempted to explain the two degrees of tunability capability using the Stoner-Wohlfarth model. I found that the Stoner-Wohlfarth model does explain how the dual-biasing method can generate tunable random switching signals, however, it does not predict significant separation between the AP- and P-state dwell times. I then attempted to fit the AP-rate data and the $\Delta_{AP(P)}$ data to the Néel-Brown model to explain two degrees of tunability. Through my analysis, I determined that the bias voltage influences the AP- and P-state components of V_{C0} differently, which explains why two degrees of tunability is achieved through dual-biasing.

In the final section, I analyzed how the dual-biasing method performs on MTJs with varying intrinsic thermal stability factors. My results demonstrated that dual-biasing on superparamagnetic MTJs had several advantages over dual-biasing on thermally stable MTJs. These advantages include better control over the AP-rate, higher maximum switching rate, and higher degree of separation between AP- and P-state dwell times with respect to the bias field. One of the short-comings of dual-biasing on superparamagnetic MTJs is lower degree of separation between the AP- and P-state dwell times with bias voltage at low bias fields. However, the degree of separation was still large enough to claim two degrees of tunability.

Chapter 4: Interplay between SOT and VCEC effects

This chapter is based on one of my previous publications, which is reference [123] in the bibliography.

4.1 Motivation and background

In chapter 3, I presented the dual-biasing method, which uses the STT effect and a local bias field to generate telegraphic switching signals for p-bits. However, the key disadvantage to this approach in large networks of interconnected p-bits is that it requires a tunable external field for each device, thus increasing the total circuit area. Therefore, alternative switching mechanisms should be investigated to eliminate this short-coming in dual-biased p-bits. Most of this chapter is dedicated to investigating the interplay between SOT and VCEC switching mechanisms for its potential as MRAM cells capable of switching at ultra-low switching current densities (~10^3 A/cm^2), however, I will also discuss the possibility of VCEC and SOT as a method of dual-biasing for p-bit applications. The contents of this chapter were published in [123].

The most common write/read mechanism in modern MRAM cells is STT [101 – 103]. STT-MRAM benefits from high cell density [104 – 106] but suffers from reliability and endurance limitations due to the current flows through the MTJ devices [9]. SOT has superior endurance to STT-MRAM since it does not involve a current through the MTJ, but rather, the applied current passes laterally through a heavy metal layer adjacent to the MTJ. Furthermore, there is no incubation delay time in the SOT switching process, thus enabling ultra-fast switching [107-110]. Several recent studies have demonstrated field-free SOT switching of perpendicular magnets through wedge–shaped stack structures

[111], built-in exchange biases through antiferromagnetic interface [112], and heavy-metal bilayers with opposing spin hall angles [113]. One short-coming in SOT – MRAM with three – terminal designs is their limitations on cell density. This can be avoided using two-terminal designs [114], however, this would result in the same endurance and speed limitations as in STT-MRAM. The key challenge for both STT- and SOT-MRAM is the large switching current densities (J_c) required for switching, where (J_c) ~ 5×10^8 A/cm^2 for SOT [17] and ~10^6 A/cm^2 for STT [16].

Recently, electric-field (E-field) has been proposed to develop spintronic devices with ultralow energy consumption, such as voltage-controlled exchange coupling (VCEC) [18]. VCEC occurs in perpendicular MTJs (p-MTJs) where the free layer is designed using synthetic antiferromagnetic (SAF) structure. MTJs with SAF free layers have additional promising features such as high thermal stability, ultrafast switching speeds, and functionality tuned by multiple forces [115-116]. Unlike voltage-controlled magnetic anisotropy (VCMA) presented unidirectional switching without an external field, VCEC can realize a bidirectional switching and low J_c through E-field modulation of the interlayer exchange coupling (IEC) of SAF structures [115].

Each writing mechanism in MRAM has different challenges and advantages, therefore, MRAM technologies with assisting mechanisms should be developed. This way, the advantages of one write mechanism can be exploited while mitigating its short-comings. In this strategy, two mechanisms are employed, one acts as the principal switching mechanism and the other serves to assist in switching. For example, field-free SOT switching has been achieved with an STT-assist mechanism [117]. Furthermore, significant reductions in write energy and write access time have been reported in both

STT-assisted SOT-MRAM [118-119] and SOT-assisted STT-MRAM [120-121]. Most recently, the ultrafast switching speed ~ 0.27 ns has been reported in SOT switching of perpendicular MTJs (p-MTJs) assisted with STT and voltage-controlled magnetic anisotropy (VCMA), however, the key challenge is still large switching current densities, where (J_c) ~ 2.0×10^8 A/cm^2 for SOT and ~ 2.3×10^6 A/cm^2 for STT [119].

In this work, p-MTJ stacks with bi-layered spin Hall channels (Ta/Pd) and SAF free layers were designed and fabricated. The p-MTJ stacks were then patterned into 150-nm pillars, and MTJ devices were switched through combination of SOT and E-field (VCEC). Bidirectional switching of p-MTJs is obtained with J_c as low as 3×10^3 A/cm^2 for VCEC and 6.8×10^7 A/cm^2 for SOT, where J_C for VCEC is two orders of magnitude lower than the reported values for VCEC-only switching [18] By studying the contribution of SOT and VCEC for MTJ switching, it is found that SOT plays a crucial role for bidirectional VCEC magnetization switching with ultra-low switching current density. J_c of SOT can be reduced further if advanced SOT materials, such as topological insulators, are used [122].

The rest of this chapter is organized as follows. Section 4.2 describes the stack structure of our p-MTJs and the testing set-up used in our experiments. Section 4.3 presents the DC switching results obtained via SOT-assisted VCEC switching, which demonstrates that the SOT effects allow for VCEC switching at ultra-low current densities. Section 4.4 provides an in-depth explanation behind the influence of SOT on VCEC switching. I explore the possibility of utilizing the interplay between SOT and VCEC effects for dual-biasing p-bits in section 4.5. Lastly, this chapter is summarized in section 4.6.

4.2 Measurement set-up

The full SAF p-MTJ stacks consisted of Ta (5.0 nm)/Pd (2.0 nm)/[Co (0.3 nm)/Pd (0.7 nm)]$_3$/Co (0.3 nm)/Ru (0.6 nm)/Ta (0.3 nm)/CoFeB (1.0 nm)/MgO (2.0nm)/CoFeB (1.3 nm)/Ta (0.7 nm)/[Pd (0.7 nm)/Co (0.3 nm)]$_8$, which were patterned into nano-pillars with a 150-nm diameter on top of a bi-layer spin Hall channel with a 250-nm width, as illustrated in Fig. 4.1(a). Metallic electrodes of Ti (10 nm)/Au (100 nm) were deposited using electron beam evaporation. Rather than using field annealing, these devices were annealed using rapid thermal annealing (RTA) at 250 °C for 20 minutes.

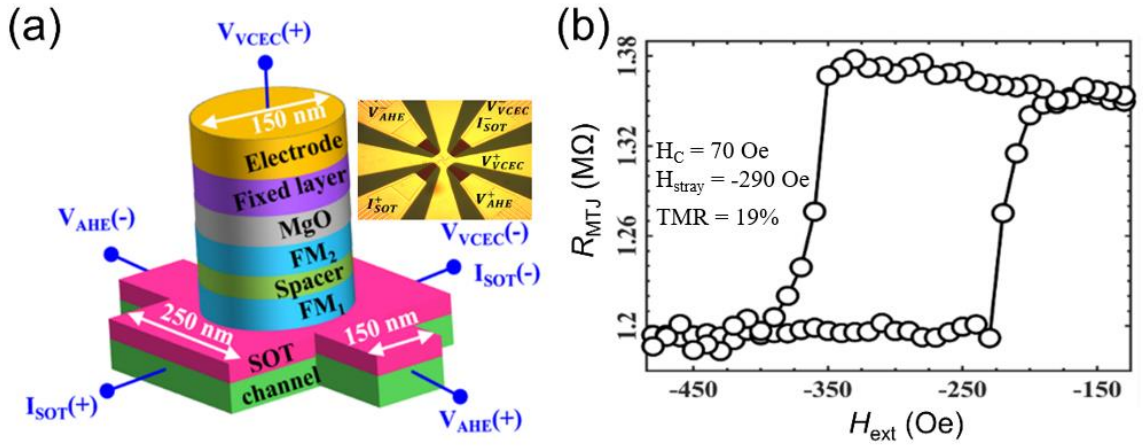


Fig. 4.1. (a) Diagram of probe connection with respect to the SAF p-MTJ device with top view of electrode pattern for all 5 terminals, and **(b)** the out-of-plane minor loop of resistance (R_{MTJ}) vs. external magnetic field (H_{ext}) curve of 150-nm p-MTJ devices with a current of 2 nA (V_{VCEC} ~ 2.5 mV). Images were extracted from the Supplementary Information in [123].

The measurements collected in this chapter can be divided into three categories: (1) transport properties of the p-MTJ devices under an electric-field (VCEC and VCMA effects); (2) influence of both electric-field and SOT effects on p-MTJ devices; and (3) SOT-assisted VCEC switching of p-MTJ devices. Figure 4.1(a) shows the schematic of our

measurement set-up of the SAF p-MTJ devices with 5 probe terminals labelled as V_{VCEC} (+), I_{SOT} (+), V_{VCEC} (-)/I_{SOT} (-), V_{AHE} (+), and V_{AHE} (-).

In the first category of measurements, multiple field-switching curves were collected with an out-of-plane external magnetic field (H_{ext}) at multiple applied voltages (V_{VCEC}). These curves are used to estimate H_C and H_{stray} of the SAF free layer at V_{VCEC}. Figure 4.1(b) shows R_{MTJ} vs. H_{ext} curve of the 150-nm SAF p-MTJ used for the measurements presented in section 4.3 at V_{VCEC} = 2.5 mV. This curve provides the intrinsic values for H_C and H_{stray} which are 70 Oe and -290 Oe, respectively.

In the second category of measurements, field-switching measurements were repeated at multiple applied SOT currents (I_{SOT}), which were used to measure H_C and H_{stray} for various combinations of V_{VCEC} and I_{SOT}. For these measurements, a third electrical probe at I_{SOT} (+) was used to apply I_{SOT} from a Keithley 6221 current source. V_{VCEC} and I_{SOT} have a shared ground contact (labelled V_{VCEC}(-)/I_{SOT}(-) in Fig. 4.1(a)) which may create a large voltage drop across the p-MTJ pillar when $V_{VCEC} < 0$ since the output voltage from the source meter (V_{OUT}) is not equal to V_{VCEC} when I_{SOT} is applied, but rather $V_{OUT} = V_{VCEC} + 0.5I_{SOT}\times R_{CH}$, where R_{CH} is the resistance of the spin Hall channel, which was measured to be 1.8 kΩ. By re-arranging the expression for V_{OUT}, we can express V_{VCEC} as $V_{VCEC} = V_{OUT} - 0.5I_{SOT}\times R_{CH}$. When $V_{OUT} \leq -1.0$ V and $I_{SOT} \geq 1.0$ mA, $|V_{VCEC}|$ becomes ≥ 2.0 V, which is near the breakdown voltage of the p-MTJ devices. To avoid breakdown of the MgO barrier, a current (I_{VCEC}) was applied rather than a voltage in these measurements and V_{VCEC} was taken to be the average of $I_{VCEC}\times R_P$ and $I_{VCEC}\times R_{AP}$.

In the third category of measurements, R_{MTJ} vs. V_{VCEC} curves were collected in the same method as STT switching curves are collected (recall section 2.2.2). V_{VCEC} was from negative-to-positive values, then from positive-to-negative values with step sizes ~15 mV and held for 1 ms at each step. These curves were obtained at multiple I_{SOT} values and with a constant out-of-plane magnetic field (H_{shift}) to overcome H_{stray} from the top fixed layer. Note that H_{shift} does not assist in VCEC switching, but rather, it allows the R_{MTJ} vs V_{VCEC} curve to be centered around $V_{VCEC} \sim 0$. As with the second category of measurements, I_{VCEC} replaces V_{VCEC} to avoid large voltage drops across the p-MTJ pillar at $V_{VCEC} < 0$. Three electrical probes at V_{VCEC} (+), I_{SOT} (+), and V_{VCEC} (-)/I_{SOT} (-) were used for these measurements. Furthermore, the R_{MTJ} values obtained were corrected post-measurements to account for I_{SOT} using Eq. 4.1, where V_{OUT} is the voltage measured from the Keithley 2400 source meter.

$$R_{\mathrm{MTJ}} = \frac{V_{\mathrm{OUT}} - \frac{1}{2}(I_{\mathrm{VCEC}} + I_{\mathrm{SOT}})R_{\mathrm{CH}}}{I_{\mathrm{VCEC}}} \tag{4.1}$$

In addition to the corrections for R_{MTJ} provided by Eq. 4.1, the R_{MTJ} should also be corrected for the thermal effects that I_{SOT} has on R_{CH}, which was done using Eq. 4.2, where R_{MTJ}* is the corrected value of R_{MTJ} and R_T is the correction factor which represents the change in $0.5R_{CH}$ due to the thermal effects of I_{SOT}. For each data set, R_T was determined by setting the constraint $R_{MTJ}(+I_{VCEC}) = R_{MTJ}(-I_{VCEC})$. Note that this constraint is only valid for $|I_{VCEC}|$ values where the p-MTJ is in the same resistance state for both $+I_{VCEC}$ and $-I_{VCEC}$.

$$R^{*}_{\mathrm{MTJ}} = R_{\mathrm{MTJ}} - \frac{I_{\mathrm{VCEC}} + I_{\mathrm{SOT}}}{I_{\mathrm{VCEC}}} R_{\mathrm{T}} \tag{4.2}$$

4.3 Results

4.3.1 Transport properties with E-field

Field switching curves were obtained at multiple V_{VCEC} values between -0.75 V and +0.75 V, examples of which are shown in Fig. 4.2(a), to evaluate effects of V_{VCEC} on H_C and H_{stray} at I_{SOT} = 0. Figures 4.2(b) and 4.2(c) show that H_C and H_{stray} decrease with V_{VCEC} at a rate of ~ 30.0 Oe/V and ~ 9.9 Oe/V without I_{SOT}, respectively.

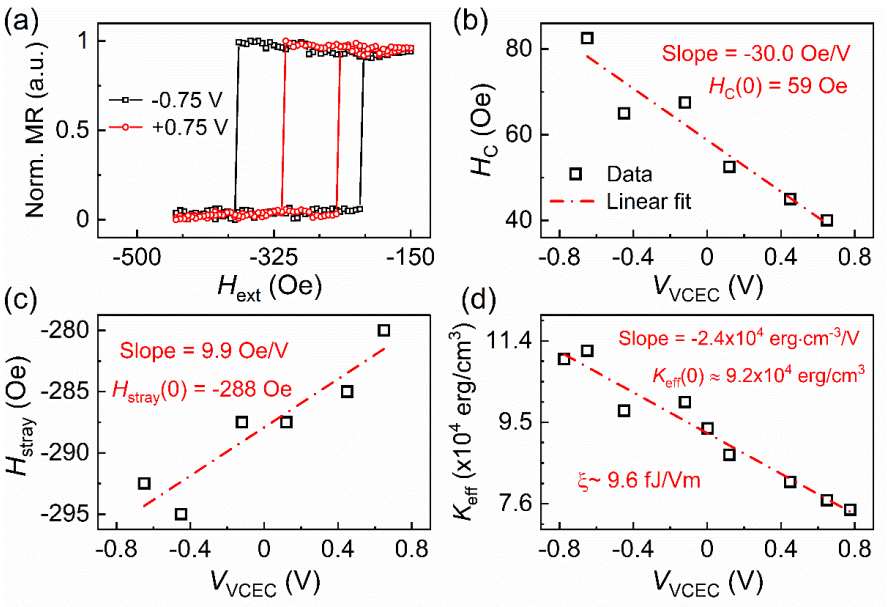


Fig. 4.2. (a) R_{MTJ} vs. H_{ext} curves at V_{VCEC} = ±0.75 V. Measurements and linear fits of **(b)** H_C vs. V_{VCEC} and **(c)** H_{stray} vs. V_{VCEC}. **(d)** Magnetic anisotropy (K_{eff}) vs. V_{VCEC} obtained from H_C measurements fitted through the Kurkijärvi-Fulton-Dunkelberger equation. Image was extracted from the Supplementary information in [123].

The behavior of H_C vs V_{VCEC} shown in Fig. 4.2(b) is consistent with results reported in [100,124-125], however, these results alone are indistinguishable from VCMA effects and do not demonstrate the presence of VCEC. On the other hand, the behavior of H_{stray} vs V_{VCEC}, shown in Fig. 4.2(c), cannot be attributed to the VCMA effects and does indicate

that IEC changes. It should be noted that STT effects can also cause H_{stray} to shift [126-127], however, our SAF p-MTJ devices have 2.0 nm thick MgO tunneling barriers, therefore, changes in H_{stray} cannot be attributed to STT effects.

Using the data shown in Fig. 4.2(b), K_{eff} was calculated using the Kurkijärvi-Fulton-Dunkelberger equation [89], as shown in Eq. 4.3. In this equation, τ_0 is the attempt time, υ is the field ramp rate, which is 100 Oe/second, V is the volume of the CoFeB free layer, M_S is the saturation magnetization of the SAF free layer, which was assumed to be 1100 emu/cc, k_B is Boltzmann constant, and T = 300 K. K_{eff} was determined by calculating σ at each field value and finding the K_{eff} value that produced the maximum σ occurred at the measured H_C. Figure 5.4(d) shows the K_{eff} vs. V_{VCEC} plot, which was used to calculate the VCMA coefficient (ξ).

$$\sigma = \left\{\frac{1}{\tau_0 \upsilon} exp\left[-\frac{K_{eff}V}{k_B T}\left(1-\frac{HM_S}{2K_{eff}}\right)^2\right]\right\} x\ exp\left\{-\int_0^H \frac{1}{\tau_0 \upsilon} exp\left[-\frac{K_{eff}V}{k_B T}\left(1-\frac{H'M_S}{2K_{eff}}\right)^2\right] dH'\right\} \quad (4.3)$$

4.3.2 SOT and VCEC effects

After characterizing the H_C vs. V_{VCEC} and H_{stray} vs. V_{VCEC} behavior at I_{SOT} = 0, I repeated these measurements at various SOT currents. Figures 4.3(a) and 4.3(b) show field switching loops for V_{VCEC} = +0.5 V and -0.5 V, respectively, at I_{SOT} = 0.25 mA, 0.75 mA, and 1.25 mA (J_{SOT} ~ 1.4×10^7 A/cm^2, 4.3×10^7 A/cm^2, and 7.1×10^7 A/cm^2). These plots show that I_{SOT} clearly influences the switching behavior of SAF p-MTJs, particularly at negative V_{VCEC}. The H_C vs. V_{VCEC} and H_{stray} vs. V_{VCEC} measurements at I_{SOT} = 0 mA, 0.5 mA, and 1.0 mA are shown in Figs. 4.3(c) and 4.3(d), which show that H_C and H_{stray} both

decreased with V_{VCEC} at all I_{SOT} values shown. However, the slope of H_C vs. V_{VCEC} and H_{stray} vs. V_{VCEC} both decreased as I_{SOT} increased. When I_{SOT} = 1.0 mA, H_C is nearly constant with V_{VCEC}. These results illustrate that I_{SOT} affects how the E-field influences the switching behavior of the p-MTJs.

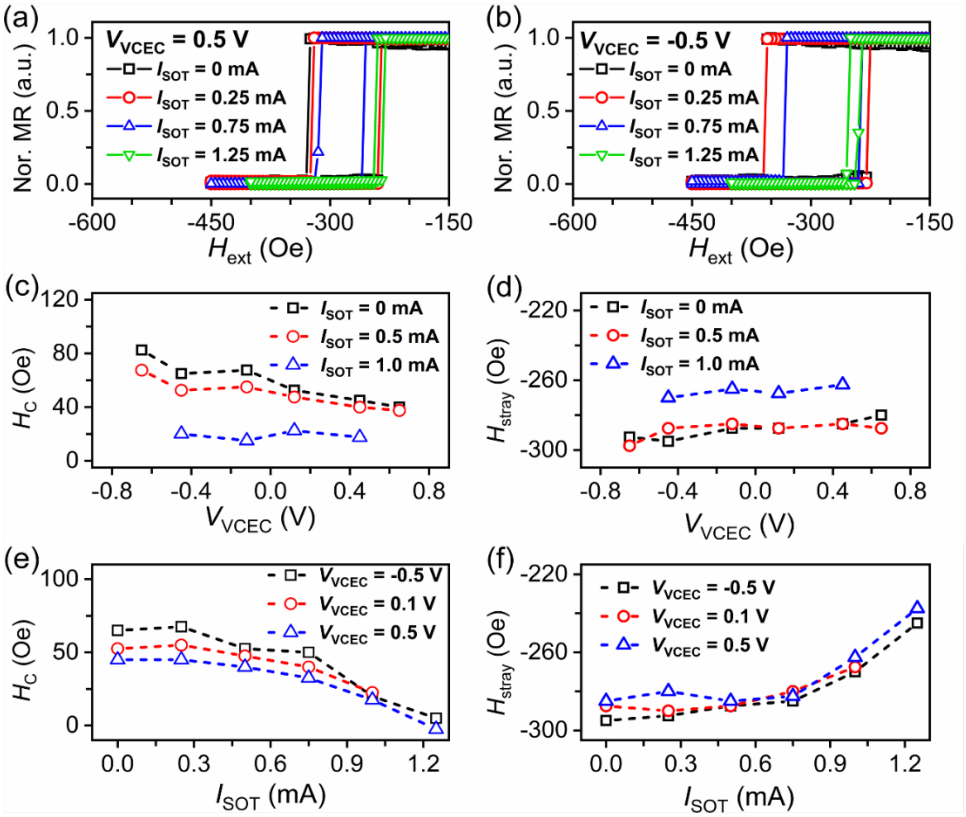


Fig. 4.3. (a, b) External magnetic field (H_{ext}) vs. normalized TMR loops of 150-nm SAF p-MTJ pillars with a [Co (0.3 nm)/Pd (0.7 nm)]$_3$/Co (0.3 nm)/Ru (0.6 nm)/Ta (0.3 nm)/CoFeB (1.0 nm) SAF free layer and the 250-nm Ta (5.0 nm)/Pd (2.0 nm) bi-layered spin Hall channel at I_{SOT} = 0, 0.25 mA, 0.75 mA, and 1.25 mA (J_{SOT} = 0, 1.4×10^7 A/cm^2, 4.3×10^7 A/cm^2, and 7.1×10^7 A/cm^2) with V_{VCEC} of +0.5 V and -0.5 V. **(c, d)** H_C and H_{stray} as a function of V_{VCEC} at I_{SOT} = 0, 0.5 mA, and 1.0 mA (J_{SOT} = 0, 2.9×10^7 A/cm^2, and 5.7×10^7 A/cm^2), where each set shows the slope of H_C and H_{stray} with V_{VCEC}, which were obtained through linear fits. **(e, f)** H_C and H_{stray} as a function of I_{SOT} at V_{VCEC} = -0.5 V, 0.1 V, and +0.5 V, where both H_C and H_{stray} show non-linear behavior with I_{SOT}. Image was extracted from [123].

Figures 4.3(e) and 4.3(f) show the dependence of H_C and H_{stray} on I_{SOT}, respectively, at V_{VCEC} = -0.5 V, +0.1 V, and +0.5 V. These figures show that both H_C and $|H_{stray}|$ decrease non-linearly with I_{SOT} for three cases of V_{VCEC}. Figure 4.3(e) shows that H_C is approximately zero at $I_{SOT} \sim 1.25$ mA ($J_{SOT} \sim 7.1\times10^7$ A/cm^2). Furthermore, Fig. 4.3(f) shows that as I_{SOT} increases from 0 mA to 0.75 mA, H_{stray} only decreases by approximately 10 Oe, however, when I_{SOT} increases from 0.75 mA to 1.25 mA, $|H_{stray}|$ decreases by nearly 40.0 Oe for all V_{VCEC} values. The factors that explain the non-linear behavior of I_{SOT} vs $|H_{stray}|$ and H_C are explained in section 4.4.

4.3.3 SOT assisted VCEC switching

The layer stack is shown on the left-side of Fig. 4.4(a) and an SEM image of the SAF p-MTJ and the schematic of the measurement set-up is shown on the right-side of Fig. 4.4(a). Recall section 4.2 for a full description of the full probe connection layout and the function of each terminal. For our tests, we limited the magnitude of V_{VCEC} to 0.75 V, which corresponds to $|J_{c,VCEC}| = 5.7\times10^3$ A/cm^2 and $|I_{VCEC}| = 1.0$ μA, to avoid voltage-induced breakdown of the MgO barrier.

H_{shift} and I_{SOT} were used to optimize the VCEC switching performance. Decreasing $|H_{shift}|$ increases V_{VCEC} for AP-to-P switching but reduces V_{VCEC} for P-to-AP switching. Increasing I_{SOT} decreases V_{VCEC} for AP-to-P switching and reduces the difference between VCEC for AP-to-P and P-to-AP switching (see section 4.4 for a full discussion on the influence of H_{shift} and I_{SOT} on VCEC switching properties). Tuning these parameters allows for VCEC switching voltages to be minimized and centered around 0 V. At H_{shift} = -245 Oe and $J_{SOT} = 6.8\times10^7$ A/cm^2 (I_{SOT} = 1.2 mA), bidirectional VCEC switching occurs at

V_{VCEC} = +0.51 V and -0.31V ($J_{c,VCEC}$ = +2.8×10^3 A/cm^2 and -1.1×10^3 A/cm^2), as shown in Fig. 4.4(b). Note that large drop in both R_{AP} and R_P with increasing $|V_{VCEC}|$ observed in Fig. 4.4(b) indicate that the devices tested are susceptible to Joule heating effects because of their large resistances.

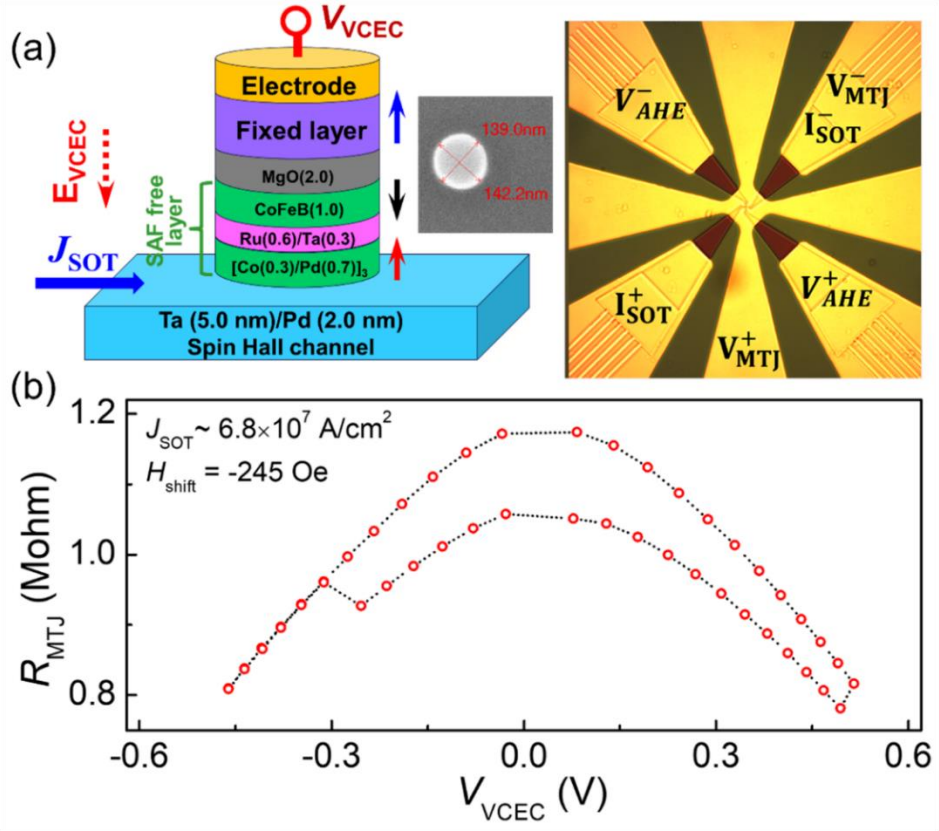


Fig. 4.4. (a) Schematic of p-MTJ devices with SAF free layers switched through SOT + VCEC and the 150-nm MTJ pillar. The SAF free layer has a stack of [Co (0.3 nm)/Pd (0.7 nm)]$_3$/Co (0.3 nm)/Ru (0.6 nm)/Ta (0.3 nm)/CoFeB (1.0 nm) and the bi-layered spin Hall channel has a stack of Ta (5.0 nm)/Pd (2.0 nm) and a width of 250-nm. **(b)** R_{MTJ} vs. V_{VCEC} loop of the 150-nm p-MTJ pillar, where H_{shift} = -245 Oe and J_{SOT} = 6.8×10^7 A/cm^2 (I_{SOT} = 1.2 mA) are utilized to assist VCEC switching. Note that H_{shift} is used to center the R_{MTJ} vs V_{VCEC} curve around V_{VCEC} ~ 0. Image was extracted from [123].

The key result from this plot is that the switching current densities shown are nearly two orders of magnitude lower than the reported for VCEC-only switching [18]. This

demonstrates that SOT has a strong influence on VCEC switching and can be used as an assisting mechanism in VCEC-MRAM. For a more detailed analysis of the influence of I_{SOT} and H_{shift} on VCEC switching, see section 4.4.

These results demonstrate that bidirectional switching of SAF p-MTJ devices can be achieved with ultralow switching current densities of $J_{c,VCEC} = 2.8\times10^3$ A/cm^2 and $J_{SOT} = 6.8\times10^7$ A/cm^2. Through the combination of VCEC and SOT, it was determined that I_{SOT} can significantly influence H_C due to changes magnetic anisotropy and H_{stray} due to changes in the IEC strength in the SAF bottom free layers. Our results reveal that SOT-assisted VCEC switching of SAF p-MTJ devices may pave the way for novel spintronic devices with ultra-low switching currents.

Although ultralow $J_{c,VCEC}$ originated from VCEC is obtained, J_{SOT} is still high, meaning that the overall switching energy does not necessarily improve significantly from VCEC-only switching, despite the two order of magnitude reduction in $J_{c,VCEC}$. A potential solution to reduce J_{SOT} is to replace the heavy-metal (Ta, W, Pt)-based spin Hall channel with an advanced SOT material or other bi-layer stacks such as Topological materials. The enhanced θ_{SH} of these materials would allow for J_{SOT} to decrease by several orders of magnitude while still having a similar influence on IEC in the bottom SAF free layer of p-MTJs, thus maintaining switching current densities as low as 10^3 A/cm^2. Furthermore, the use of H_{shift} as a means of centering the R_{MTJ} vs V_{VCEC} curve around $V_{VCEC} \sim 0$ could be an obstacle for application as a realistic memory device. One way this could be avoided is by exploring alternative material systems in the SAF free layer. For example, [Fe/Pd] could replace [Co/Pd] since VCEC switching has been achieved on p-MTJs with [Fe/Pd] based SAF free layers without H_{shift} [18]. Another way to avoid using H_{shift} is to explore p-MTJ

stacks with an additional bias layer that generates a built-in bias field [128].

4.4 Discussions

4.4.1 Factors influencing H_C by I_{SOT}

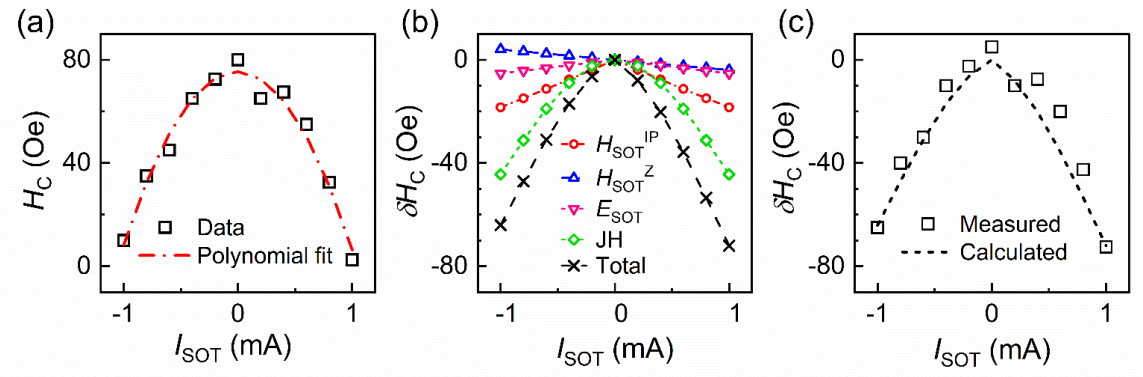


Fig. 4.5. (a) The measured and fitted H_C vs. I_{SOT}, illustrating that H_C is slightly smaller at positive I_{SOT}. **(b)** H_C changes as a function of I_{SOT}, following the origin from H_{SOT}^{IP}, E_{SOT}, JH, H_{SOT}^{Z}, and the combined influence of all four factors. **(c)** The measured and calculated H_C change (δH_C) as a function of I_{SOT}. Image extracted from the Supplementary information in [123].

In this section, I investigate the influence of I_{SOT} on H_C by analyzing data for H_C measured over a wide range of I_{SOT} values, results of which are shown in Fig. 4.5(a). There are four potential reasons that are explored to explain the relationship between H_C and I_{SOT} shown in Fig. 4.5(a) as well as Fig. 4.3(e) in the previous section. These are (1) the in-plane component of the effective field generated by I_{SOT} via the Rashba and spin Hall effects (H_{SOT}^{IP}); (2) the perpendicular component of the effective field generated by I_{SOT} (H_{SOT}^{Z}); (3) the E-field generated by electron accumulation at the CoFeB/MgO interface (E_{SOT}); and (4) I_{SOT} induced Joule heating (JH). Figures 4.5(a) and 4.3(e) in the previous section shows a quadratic relationship between with negative concavity H_C and I_{SOT}. This relation should be considered when determining which factor is the dominating influencing factor on H_C.

Sharrock's equation [129] was used to determine the expected behavior of H_C with I_{SOT}, which is shown in Eq. 4.4(a), where t_R is the time step used to obtain field switching curves (10 ms). The thermal stability factor, defined as $\Delta = K_{eff}V/k_bT = H_K M_S V/2k_bT$, was expressed as a function of I_{SOT}, which means the Eq. 4.4 can be simplified to Eq. 4.5, where $C_0 = 2k_bT/M_SV$ and $C_1 = \sqrt{\ln(t_R/0.693\tau_0)}$. Additionally, the Neel-Brown model, shown in Eq. 4.6, was used to determine $\Delta(I_{SOT})$, where H_A^* is a generalized field term, H_{K0}^* is a generalized anisotropy field term, and Δ_0 is the thermal stability factor at I_{SOT} = 0.

$$H_C = H_K\left[1 - \sqrt{\frac{k_bT}{K_{eff}V}\ln\left(\frac{t_R}{0.693\tau_0}\right)}\right] = \frac{2\Delta(I_{SOT})k_bT}{M_SV}\left[1 - \sqrt{\frac{1}{\Delta(I_{SOT})}}\sqrt{\ln\left(\frac{t_R}{0.693\tau_0}\right)}\right] \quad (4.4)$$

$$H_C = C_0\left[\Delta(I_{SOT}) - C_1\sqrt{\Delta(I_{SOT})}\right] \quad (4.5)$$

$$\Delta(I_{SOT}) = \Delta_0\left(1 - \frac{H_A^*}{H_{K0}^*}\right)^2 \quad (4.6)$$

S. Fukami *et al* demonstrated that H_C decreased with I_{SOT} in AFM/FM bilayer systems, which was attributed to H_{SOT}^{IP}, where H_C changes (δH_C) $\propto$ -$|I_{SOT}^2|$ (positive concavity) [112]. In devices where the AFM layer thickness was 6.0 nm or less, the influence of I_{SOT} on H_C could not be attributed to an exchange bias field (H_{ex}) since H_{ex} was not observed in these samples. Additionally, JH was not considered to be a factor in their results since changes in temperature with I_{SOT} were calculated to be less than 10 K. Using Eq. 4.5 and 4.6, the relationship between δH_C and I_{SOT} predicted by the influence of H_{SOT}^{IP} can be expressed using Eq. 4.7, where D_{Hsot}^{IP} is the slope of H_{SOT}^{IP} vs. I_{SOT} and δH_C represents the H_C changes. The value for D_{Hsot}^{IP} was estimated to be ~13.0 Oe/mA based

on measurements of H_{SOT}^{IP} in Ta-based spin Hall channels reported in previous experimental work [117].

$$\delta H_C\left(H_{SOT}^{IP}\text{ dependence}\right) = C_0\left[\Delta_0\left(1-\frac{D_{Hsot}^{IP}|I_{SOT}|}{H_K}\right)^2 - C_1\sqrt{\Delta_0}\left(1-\frac{D_{Hsot}^{IP}|I_{SOT}|}{H_K}\right)\right] \quad (4.7)$$

The presence of H_{SOT}^{Z} in bi-layered spin Hall channels consisting of two heavy metals with opposite θ_{SH} signs was first demonstrated by Q. Ma *et al* from the changes in H_{stray} with I_{SOT} [113]. The influence of H_{SOT}^{Z} on H_C was not discussed in this study, however, the data presented showed that H_C decreased with I_{SOT}. The results presented in [113] found that $H_{SOT}^{Z} \propto I_{SOT}$, therefore, the relationship between δH_C and I_{SOT} predicted by the influence of H_{SOT}^{Z} can be expressed using Eq. 4.8, where D_{Hsot}^{Z} is the slope of H_{SOT}^{Z} vs. I_{SOT}. Note that Eq. 4.6 and 4.7 are similar, except δH_C in Eq. 4.8 is dependent on the sign of I_{SOT}, therefore, D_{Hsot}^{Z} was determined by measuring the difference in H_C at I_{SOT} = +1.0 mA and -1.0 mA. Figure 4.5(a) shows that H_C is approximately 7.0 Oe larger at I_{SOT} = -1.0 mA than at I_{SOT} = +1.0 mA, meaning that D_{Hsot}^{Z} is approximately 3.5 Oe/mA.

$$\delta H_C\left(H_{SOT}^{Z}\text{ dependence}\right) = C_0\left[\Delta_0\left(1-\frac{D_{Hsot}^{Z}I_{SOT}}{H_K}\right)^2 - C_1\sqrt{\Delta_0}\left(1-\frac{D_{Hsot}^{Z}I_{SOT}}{H_K}\right)\right] \quad (4.8)$$

I_{SOT} could generate an E-field, E_{SOT}, due electron accumulation at the CoFeB/MgO interface. Recall that δH_C is linear with V_{VCEC} (see Fig. 4.3(c)), therefore, the relationship between δH_C and I_{SOT} predicted by the influence of E_{SOT} should also be linear. Furthermore, δH_C should not be dependent on the sign of I_{SOT}. Equation 4.9 shows the relation between δH_C and I_{SOT}, where D_{Hc}^{E} is the slope of δH_C with $|I_{SOT}|$. D_{Hc}^{E} was estimated to be approximately 5.3 Oe/mA by measuring the change in $\delta H_C/\delta V_{VCEC}$ with I_{SOT}.

$$\delta H_{\mathrm{C}}(E_{\mathrm{SOT}}\ \text{dependence}) = D_{\mathrm{Hc}}^{E} \cdot |I_{\mathrm{SOT}}| \tag{4.9}$$

Results presented by L. Liu *et al* showed large reductions in H_{C} with I_{SOT}, where δH_{C} had a quadratic relation with I_{SOT} with negative concavity, which was attributed to JH [130]. Note that the relation between δH_{C} and I_{SOT} shown in [130] matches the results shown in Fig. 4.5(a). The effect of JH predicts that I_{SOT} will cause the temperature of the bottom SAF free layer to increase (δT), where $\delta \mathrm{T} \propto I_{\mathrm{SOT}}^2$. Since $\Delta \propto 1/\mathrm{T}$, Δ will decrease with $1/I_{\mathrm{SOT}}^2$, as shown in Eq. 4.10, where Δ(δT) represents Δ for a given temperature increase of δT, T_0 is the baseline temperature when $I_{\mathrm{SOT}} = 0$, and γ is a constant that represents the change in temperature with I_{SOT}. By inserting Eq. 4.10 in Eq. 4.5), the relationship between δH_{C} and I_{SOT} predicted by JH can be expressed using Eq. 4.11.

$$\Delta(\delta \mathrm{T}) = \frac{H_{\mathrm{K}} M_{\mathrm{S}} \mathrm{V}}{2k_{\mathrm{b}}(T_0 + \delta \mathrm{T})} \rightarrow \Delta(\delta \mathrm{T}) = \frac{\Delta_0}{(1 + \delta \mathrm{T}/T_0)} \rightarrow \Delta(\delta \mathrm{T}) = \frac{\Delta_0}{\left(1 + \frac{\gamma}{T_0} I_{\mathrm{SOT}}^2\right)} \tag{4.10}$$

$$\delta H_{\mathrm{C}}(\text{Joule heating dependence}) = C_0 \left[\frac{\Delta_0}{\left(1 + \frac{\gamma}{T_0} I_{\mathrm{SOT}}^2\right)} - C_1 \sqrt{\frac{\Delta_0}{\left(1 + \frac{\gamma}{T_0} I_{\mathrm{SOT}}^2\right)}} \right] \tag{4.11}$$

Figure 4.5(b) compares the calculated δH_{C} values predicted by $H_{\mathrm{SOT}}^{\mathrm{IP}}$, $H_{\mathrm{SOT}}^{\mathrm{Z}}$, E_{SOT}, and JH as well as the total calculated δH_{C}, which is the sum of all these contributions. These curves show that JH and $H_{\mathrm{SOT}}^{\mathrm{IP}}$ are the dominant factors influencing δH_{C}, whereas $H_{\mathrm{SOT}}^{\mathrm{Z}}$ and E_{SOT} are secondary factors. Figure 4.5(c) shows the total calculated δH_{C} along with the measured δH_{C} for this device.

4.4.2 Factors influencing H_{stray} by I_{SOT}

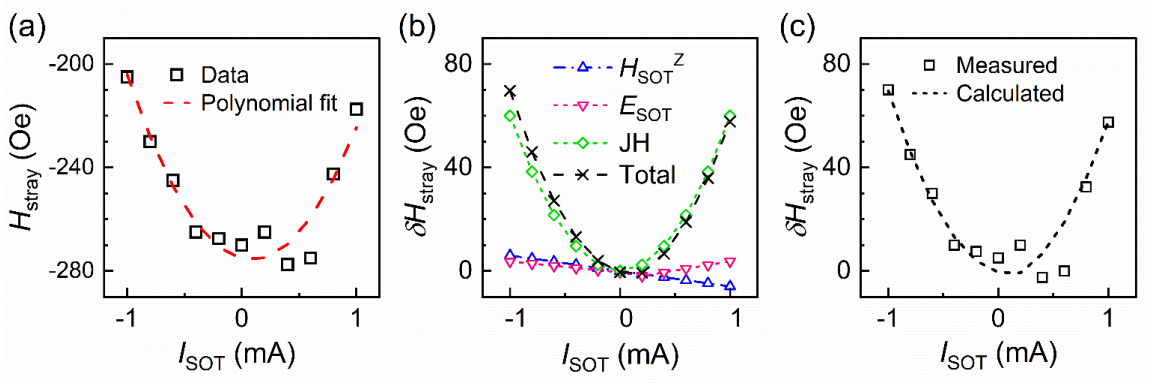


Fig. 4.6. (a) H_{stray} vs. I_{SOT}, illustrating that H_{stray} is larger for positive I_{SOT}. **(b)** H_{stray} as a function of I_{SOT}, following the origin from H_{SOT}^{Z}, E_{SOT}, JH, and the combined influence of all factors. **(c)** The measured and calculated H_{stray} change (δH_{stray}) as a function of I_{SOT}. Image extracted from the Supplementary information in [123].

In the previous section, I examined the influence of I_{SOT} on H_C and in this section, I investigate the influence of I_{SOT} on H_{stray}. The H_{stray} vs I_{SOT} results are shown in Fig. 4.6(a). The same four factors that were explored to explain the relationship between δH_C with I_{SOT} were also explored to explain the relationship between δH_{stray} and I_{SOT} shown in Fig. 4.6(a) as well as Fig. 4.3(f) shown in section 4.3.3. However, the equations used to predict δH_{stray} are not the same as for δH_C. As with the analysis for explaining the relation between δH_C and I_{SOT}, I verified that the expected relationship between H_{stray} and I_{SOT} matches our results, which shows that δH_{stray} has a quadratic relationship with I_{SOT} with positive concavity.

Recall that S. Fukami *et al* showed that increasing I_{SOT} caused H_C to decrease, which they attributed to H_{SOT}^{IP}. However, their results did not show any shift in the center of the

field switching loop with I_{SOT} [112]. For this reason, I assumed that $\delta H_{stray}(H_{SOT}^{IP}$ dependence) = 0.

Since the spin Hall channel is comprised of a bi-layered structure with two heavy metals of opposite θ_{SH} signs, H_{SOT}^{Z} needs to be considered. Q. Ma *et al* showed that H_{SOT}^{Z} generated from these bi-layered spin Hall channels results in shifts in H_{stray}, where $\delta H_{stray} \propto I_{SOT}$ [113], therefore, the relation between δH_{stray} and I_{SOT} as predicted from the effects of H_{SOT}^{Z} can be expressed using Eq. 4.12, where D_{Hs}^{Z} is the slope of H_{stray} with I_{SOT}. Since H_{SOT}^{Z} is dependent on the sign of I_{SOT}, D_{Hs}^{Z} was estimated from the difference between δH_{stray} measured at I_{SOT} = +1.0 mA and -1.0 mA. Figure 5.6(a) shows that δH_{stray} is approximately 15.0 Oe larger at I_{SOT} = +1.0 mA than at I_{SOT} = -1.0 mA, meaning that D_{Hs}^{Z} is approximately -7.0 Oe/mA.

$$\delta H_{stray}\left(H_{SOT}^{Z}\text{ dependence}\right) = D_{Hs}^{Z} I_{SOT} \tag{4.12}$$

Since $H_{stray} \propto V_{VCEC}$, the relation between H_{stray} and I_{SOT} as predicted by the influence of E_{SOT} should also be linear, as shown in Eq. 4.13, where D_{Hs}^{E} is the slope of H_{stray} with I_{SOT}. Note that this equation predicts that δH_{stray} is not dependent on the sign of I_{SOT}. Similar to the method of calculating of D_{Hc}^{E} (recall previous section), D_{Hs}^{E} was estimated by measuring the change in the slope of H_{stray} with V_{VCEC} ($\delta H_{stray}/\delta V_{VCEC}$) with I_{SOT}, where D_{Hs}^{E} was calculated to be approximately 7.1 Oe/mA.

$$\delta H_{stray}(E_{SOT}\text{ dependence}) = D_{Hs}^{E}|I_{SOT}| \tag{4.13}$$

Y. Li *et al* [131] and Y. Henry *et al* [132] studied the temperature dependence of interlayer exchange coupling strength (IEC) in Co-based SAF structures, which demonstrated that IEC decreases linearly with temperature. Since IEC $\propto H_{stray} \cdot M_S \cdot t_f$ and δT

$= \gamma I_{SOT}^2$ (recall previous section), the relationship between δH_{stray} and I_{sot} when considering JH effects predicts that $\delta H_{stray} \propto I_{SOT}^2$, as shown in Eq. 4.14, which accurately represents the behavior shown in Fig. 4.6(a). In this analysis, the change in IEC (δIEC) is D_{IEC}^{T}*δT, where D_{IEC}^{T} is the slope of IEC with temperature. The value for $\gamma \cdot D_{IEC}^{T}$ was estimated to be approximately 0.0045 erg·cm^{-2}/mA^2 by fitting the previous calculations from Eq. 4.12 and 4.13 to the measured δH_{stray} values.

$$\delta H_{\text{stray}}(\text{Joule heating dependence}) = \frac{\gamma D_{\text{IEC}}^{\text{T}}}{t_{\text{F}} M_{\text{S}}} I_{\text{SOT}}^2 \tag{4.14}$$

Figure 4.6(b) compares all the calculated δH_{stray} values as predicted by the effects of H_{SOT}^{Z}, E_{SOT}, and JH as well as the total calculated δH_{stray}, which represents the sum of all these contributions. These curves show that JH is the dominant influence on δH_{stray}, and H_{SOT}^{Z} and E_{SOT} are secondary influences. Figure 4.6(c) shows the total calculated δH_{stray} values and the measured δH_{stray} values for this device.

4.4.3 Influence of H_{shift} on VCEC switching

The H_{stray} values obtained in our measurements arise from a combination of (1) the dipole field from the top fixed layer on the bottom free layers (H_{dip}) and (2) IEC in the bottom SAF free layers. IEC can be calculated from the offset in the minor loop of the $[Co/Pd]_n$ layers in the bottom SAF free layer (H_{ex}) using the expression IEC = $H_{ex} \cdot M_S \cdot t_f$. Our results show that $H_{stray} \propto V_{VCEC}$ which implies that IEC $\propto V_{VCEC}$, assuming that H_{dip} is independent of V_{VCEC}. Recall that the bottom SAF free layers of our devices favor AFM configuration at $V_{VCEC} = 0$. Furthermore, when large positive V_{VCEC} is applied, IEC in the SAF free layer shifts to favor FM configuration and when large negative V_{VCEC} is applied, IEC shifts more towards AFM configuration.

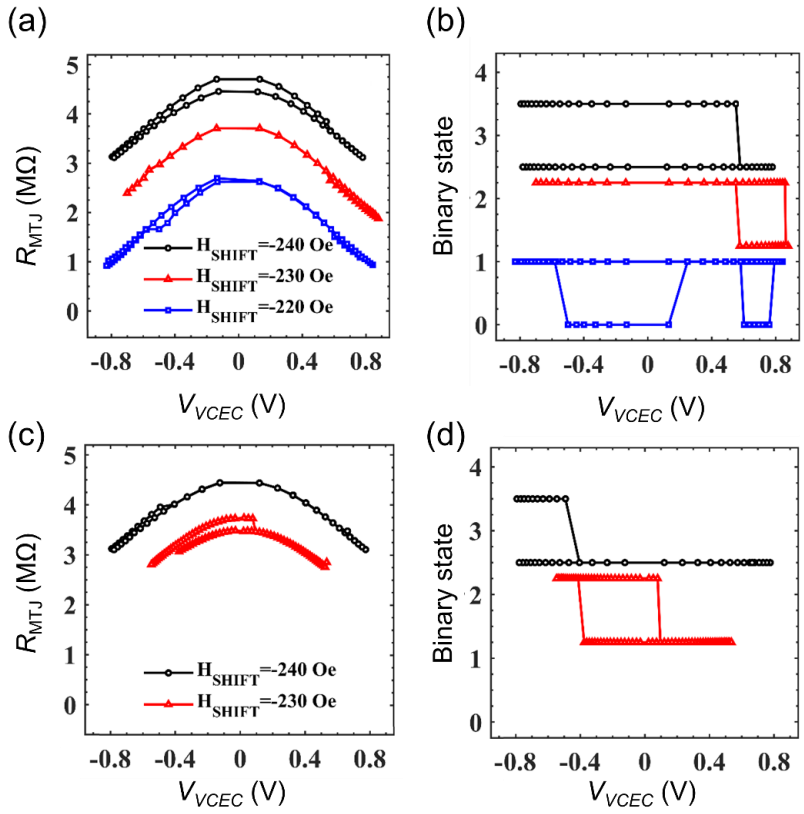


Fig. 4.7. (a) VCEC switching curves at I_{SOT} = 0.8 mA with H_{shift} = -240 Oe, -230 Oe, and -220 Oe and **(b)** the corresponding binary state of the p-MTJs, where AP-state denotes a high binary state. **(c)** VCEC switching curves at I_{SOT} = 0.9 mA with H_{shift} = -240 Oe and -230 Oe and **d)** the corresponding binary state of the p-MTJs. Curves for H_{shift} = -240 Oe and -230 Oe are shifted along the vertical axis in each plot for easy comparison between these curves. Image extracted from the Supplementary information in [123].

IEC can change signs even if H_{stray} does not, however, for IEC to change signs at H_{shift} = 0, the magnitude of H_{stray} should less than the H_C. The data shown in Figs. 4.3(a)-4.3(f) in the previous section indicates that, in the SAF p-MTJ devices tested, there were no conditions for V_{VCEC} and I_{SOT} where this criteria was met. This means that VCEC switching could only be achieved by applying H_{shift} in a direction that would facilitate the transition

from AFM to FM coupling with drastically increasing $|V_{VCEC}|$. By applying H_{shift}, the effective IEC has essentially become IEC = $(H_{ex} - |H_{shift}|)\cdot M_S \cdot t_f$. Note that if $H_{shift} > H_{ex}$, then H_{shift} will cause IEC to favor FM coupling. Therefore, $|H_{shift}|$ was set to be less than or equal to H_{stray}, which ensures that bidirectional VCEC switching can only be attributed to sign changes in IEC and not field switching.

Figures 4.7(a) – 4.7(d) illustrate the influence of H_{shift} on VCEC switching curves. Figure 4.7(a) shows the VCEC switching plots at I_{SOT} = 0.8 mA and H_{shift} = -240 Oe, -230 Oe, and -220 Oe and Fig. 4.7(b) shows the binary state vs. V_{VCEC} for these curves, where the AP-state corresponds to a high binary state. When H_{shift} = -240 Oe, VCEC switching was only observed for AP-to-P switching, which occurs at V_{VCEC} ~ 0.58 V. Bidirectional VCEC switching was achieved at H_{shift} = -230 Oe, where V_{VCEC} required for AP-to-P switching increased to 0.86 V and V_{VCEC} required for P-to-AP was 0.57 V. By decreasing $|H_{shift}|$ from -240 Oe to -230 Oe, IEC strength increases towards AFM coupling, which is why the AP-to-P switching voltage increases and the P-to-AP switching voltage decreased (in terms of absolute value). When $|H_{shift}|$ decreased to -220 Oe, the AP-to-P switching voltage remained approximately the same as for H_{shift} = -230 Oe (~ 0.79 V) and P-to-AP switching voltage increased to 0.60 V. However, the p-MTJ device also switched back to the P-state from V_{VCEC} = 0.13 V to V_{VCEC} = -0.5 V, indicating instability at H_{shift} = -220 Oe.

Figure 4.7(c) shows VCEC switching curves at I_{SOT} = 0.9 mA and H_{shift} = -240 Oe and -230 Oe and Figure 4.7(d) shows the corresponding binary states for the curves in Figure 4.7(c). Bidirectional switching is not achieved at H_{shift} = -240 Oe, only AP-to-P switching occurs at -0.4 V. Note that the AP-to-P switching voltage shifted towards negative values compared to the H_{shift} = -240 Oe curve in Fig. 4.7(a) and 4.7(b), indicating that increasing

I_{SOT} caused IEC to favor FM coupling (see section 5.4.4 for further investigation on the influence of I_{SOT}). When H_{shift} decreases to -230 Oe, bidirectional switching is achieved at switching voltages of 0.1 V and -0.41 V. The data shown in Figures 4.7(a)-4.7(d) illustrate how decreasing H_{shift} can shift IEC strength towards AFM coupling, and therefore, compensates for the influence of I_{SOT} on IEC (see section 4.4.4).

4.4.4 Influence of I_{SOT} on VCEC switching

Figures 4.3(e) and 4.3(f) and the analysis in sections 4.4.1 and 4.4.2 demonstrate that I_{SOT} directly influences the VCEC switching behavior. The reductions in H_C and H_{stray} with I_{SOT} observed in Figs. 4.3(e-f), 4.5(a), and 4.6(a) indicate that I_{SOT} enhances the IEC strength for FM coupling. This means that the p-MTJ becomes more likely to favor P-state orientation as I_{SOT} increases which causes the AP-to-P switching voltage to decrease and the P-to-AP switching voltage to shift towards larger negative values.

Figures 4.8(a)-4.8(d) illustrate the influence of I_{SOT} on the VCEC switching curves. Note that several of these curves are also shown in Figs. 4.7(a)-4.7(d) and re-arranged to highlight their dependence on I_{SOT} rather than H_{shift}. Figures 4.8(a) and 4.8(b) show the VCEC switching curves at H_{shift} = -240 Oe and I_{SOT} = 0.7 mA, 0.8 mA, and 0.9 mA, where Fig. 4.8(a) are show the R_{MTJ} vs V_{VCEC} data and Fig. 4.8(b) shows the corresponding binary state vs. V_{VCEC} data. These data sets show that the AP-to-P switching occurs at V_{VCEC} ~ 0.78 V, 0.58 V, and -0.40 V for I_{SOT} = 0.7 mA, 0.8 mA, and 0.9 mA, respectively. Note that the curve for I_{SOT} = 0.9 mA shows AP-to-P switching occurs at $V_{VCEC} < 0$, which suggests that, at H_{shift} = -240 Oe, the shift in IEC towards FM coupling induced by I_{SOT} is large enough so that the SAF free layer in the p-MTJ favors FM coupling at V_{VCEC} = 0. However, bidirectional VCEC switching was not achieved in any of these data sets because

as I_{SOT} increases, the switching voltages for AP-to-P and P-to-AP switching shifts towards larger negative values, meaning that the center of the overall curve shifts towards larger negative values. This can be avoided be reducing H_{shift}, thus shifting the IEC strength towards AFM coupling (recall section 4.4.3).

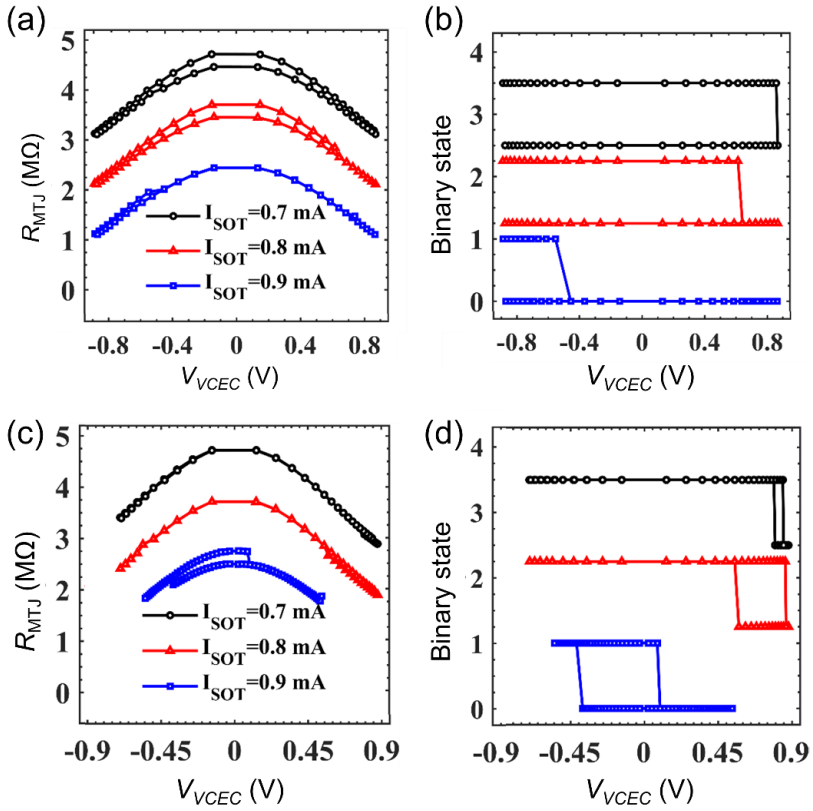


Fig. 4.8. (a) VCEC switching curves at H_{shift} = -240 Oe with I_{SOT} = 0.7 mA, 0.8 mA, and 0.9 mA and **(b)** the corresponding binary state of the p-MTJs, where AP-state denotes a high binary state. **(c)** VCEC switching curves at H_{shift} = -230 Oe when I_{SOT} = 0.7 mA, 0.8 mA, and 0.9 mA and **(d)** the corresponding binary state of the p-MTJs. Curves for I_{SOT} = 0.7 mA and 0.8 mA are shifted along the vertical axis. Image extracted from the Supplementary information in [123].

Figures 4.8(c) and 4.8(d) show the VCEC switching curves at H_{shift} = -230 Oe and I_{SOT} = 0.7 mA, 0.8 mA, and 0.9 mA, where Fig. 4.8(c) shows the R_{MTJ} vs V_{VCEC} data and Fig. 4.8(d) shows the corresponding binary state vs. V_{VCEC} data. Bidirectional switching is achieved at all I_{SOT}, where the AP-to-P switching voltages decreased from 0.86 V to 0.10

V and the P-to-AP switching voltages decreased from 0.79 V to -0.41 V as I_{SOT} increased from 0.7 mA to 0.9 mA. However, the curves for I_{SOT} = 0.7 mA and 0.8 mA show that the switching voltages for both AP-to-P and P-to-AP switching were positive, indicating that, at H_{shift} = -230 Oe, the IEC strength heavily favors AFM coupling, even at small positive V_{VCEC}. When I_{SOT} increases to 0.9 mA, the shift in IEC strength towards FM coupling is large enough for P-to-AP switching voltages to decrease to values less than zero. These curves demonstrate that VCEC switching behavior can be optimized with proper control of I_{SOT} and H_{shift}.

4.5 SOT + VCEC dual-biasing

To reduce the VCEC switching current densities even further, J_{SOT} was increased to 71.4 MA/cm^2 (I_{SOT} = 1.25 mA). However, Fig. 4.9(a) shows that the p-MTJ fluctuated stochastically between the AP- and P-states over a wide range of J_{VCEC}. This behavior can be attributed to two causes. One is the non-linear reduction in H_C with increasing J_{SOT} means that the MA of the SAF free layers eventually reaches zero, where thermal fluctuations will drive random switching in both FM_1 and FM_2.

Second is the reduction in J_{EX} with increasing J_{SOT}. The magnitude of J_{EX} will eventually become so small in both AFM and FM coupling configurations, that thermal fluctuations will drive random toggling between AFM and FM coupling in the SAF free layers. These results illustrate an upper limit for J_{SOT} when SOT is used as an assisting mechanism for VCEC-MRAM applications since this stochastic behavior would result in large write error rates. However, Fig. 4.9(b) indicates that the average MTJ state can be tuned since the fluctuates occur the most frequently around $J_{MTJ} \sim 0$, remains mostly in the P-state with occasional P-to-AP pulses as J_{VCEC} increases to positive values, then remains

mostly in the AP-state with occasional AP-to-P pulses as J_{VCEC} decreases to negative values.

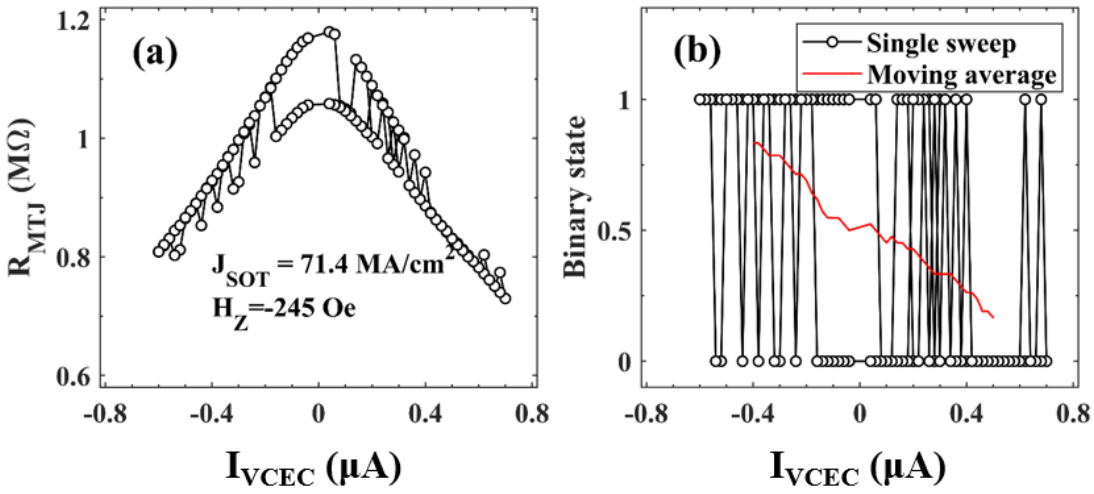


Fig. 4.9. (a) R_{MTJ} vs I_{VCEC} plot showing stochastic VCEC switching at H_Z = -245 Oe and J_{SOT} = 71.4 MA/cm^2 (I_{SOT} = 1.25 mA) and **(b)** binary representation of R_{MTJ} measurements, where the AP-state corresponds to binary state '1' (Black line with circles) and a moving average of the binary value (red line).

The I_{VCEC} versus R_{MTJ} behavior directly emulates the transfer characteristics of probabilistic bits. This means that these results illustrate how SOT + VCEC dual-biasing may have prospects in novel schemes in neuromorphic computing based on probabilistic models. In chapter 3, I presented results for dual-biasing via STT + external field, which demonstrated the two degrees of tunability feature. However, this method would create a significant increase in circuit area in large networks of interconnected p-bits since each p-bit would require a local, tunable field. Therefore, dual-biasing via SOT + VCEC would be a preferred method over STT + field. Unfortunately, the large resistances and low TMR ratios seen in these devices meant that dwell-time measurements could not be obtained in these devices since the signal-to-noise ratio was too low to detect MTJ switching on the oscilloscope. At the present stage of this research, it is unclear if two degrees of tunability

would also be achieved with SOT + VCEC dual-biasing. However, the results seen in Fig. 4.9 shows the transfer characteristics do match those required for probabilistic bits, thus making SOT + VCEC a viable method for dual-biased p-bits.

4.6 Summary

In this chapter, I developed a process for fabricating p-MTJs with SAF free layers on top of bi-layered spin Hall channels. The devices were patterned into 150 nm nanopillars on 250 nm spin Hall channels and had separate signal paths for VCEC and SOT switching. First, I studied the effects of the VCEC voltage and SOT current on the coercivity and the stray field. My results found that both the coercivity and stray field decreased linearly with V_{VCEC} and decreased with I_{SOT}^2. Changes in the coercivity with V_{VCEC} may be a result of a combination of VCEC and VCMA effects, however, changes in the stray field with V_{VCEC} indicates changes in interlayer exchange coupling.

The next set of measurements were VCEC switching measurements which were done using VCEC current sweeps at a constant SOT current and bias field. The bias field was not used to assist in switching, but rather to center the VCEC switching curve around zero. With the assistance of I_{SOT}, bidirectional VCEC switching was achieved at switching current densities of $1 – 3 \times 10^3$ A/cm^2, which is nearly two orders of magnitude lower than the switching current densities reported for VCEC only switching. While these results are promising, a large SOT current and an external bias field were required to achieve these low switching current densities. Therefore, future work should investigate methods to reduce the SOT current density and eliminate the need for an external bias field. Nevertheless, the results reported here provide a pathway for novel VCEC-MRAM cells with ultra-low switching current density.

Next, I discussed the factors of I_{SOT} influencing the coercivity and stray field. I determined that Joule heating and the effective in-plane field generated by I_{SOT} via the Rashba and spin Hall effects were primary factors whereas the E-field generated from charge accumulation by I_{SOT} and the effective out-of-plane field generated from the bi-layered spin Hall channel were secondary factors. This discussion was followed by an examination of the effects of the external bias field and I_{SOT} on the VCEC switching curves. It was determined that increasing I_{SOT} caused IEC in the free layer to favor FM coupling and decreasing the bias field caused IEC to favor AFM coupling. Optimum VCEC switching behavior could be achieved through proper control over both of these biases.

In the final section, I attempted to reduce the VCEC switching current density even further by increasing the SOT current. However, this caused stochastic switching over a large range of VCEC currents, thus making the devices unsuitable for VCEC-MRAM purposes under these conditions. After further investigation of the VCEC switching curves, I determined that the average state of the p-MTJ could be tuned with the VCEC current, which makes them useful in applications involving probabilistic bits. These results show that the dual-biasing method introduced in chapter 3 could be improved further by eliminating the need for a local, tunable external field. Experimental methods of obtaining time-domain data for these devices should be explored to determine if two-degrees of tunability is still achieved in SOT + VCEC dual-biasing as it is with STT + field dual-biasing.

Chapter 5: Stochastic Computational Random Access Memory

The major content of this chapter will be submitted as a manuscript for publication.

5.1 Motivation

Conventional computing schemes, where data is encoded in deterministic binary bits, has several challenges in meeting future demands in neuromorphic computing and other novel applications [20, 133-134]. The tasks required for neuromorphic computing require processing complex functions on large amounts of data. Therefore, one of the key challenges is the costs in circuit area and computation delay incurred in circuits based on conventional CMOS technology [135-136]. Additionally, the data is highly sensitive to thermal noise as transistor size scaling trends towards nanometer-size dimensions, which creates strict limitations on device scaling. To overcome these short-comings, alternative approaches to data encoding and hardware should be investigated.

Recall section 1.3.3 which described how probabilistic computing approaches are attractive methods in several unconventional computing applications for several reasons. One is that they have inherent noise immunity which allows for smaller device size and also makes them more error resilient [38-42]. Another key feature of probabilistic schemes is that complex functions that are common in several neuromorphic computing tasks (such as hyperbolic tangent, exponential functions, square root, etc.) can be implemented with a very small number of logic gates [44, 137].

Stochastic computing (SC) is a probabilistic scheme that has received the most attention in recent years [42]. In SC, data is encoded as probabilities in streams of random bits. There are significant reductions in hardware area for processing circuits since

multiplication, scaled addition, and scaled subtraction of two stochastic bit-streams can be performed using single logic gates [44]. Furthermore, SC is a very promising solution for various image processing tasks since it is robust against random noise. Previous studies have demonstrated that SC-based circuits can achieve above 99% accuracy rates in image recognition tasks at noise injection rates (percentage of bit-flips) above 30% whereas the same tasks processed in conventional computing schemes showed large error rates at noise injection rates below 10% [38].

One of the major short-comings of SC is the hardware area costs required for stochastic bit-stream generation (SBG). The most common methods of random number generation (RNG) with CMOS devices are through linear shift feedback registers (LSFR) [138] or ring oscillators [37]. When these circuits are implemented in SC-based networks, the hardware for SBG can consume up to 80% of the total circuit area and 80% of the total energy consumption [38]. Therefore, SC may not have any reductions in total circuit area and total energy consumption, even with the reductions in circuit area and energy consumption for data processing. Furthermore, LSFRs and ring oscillators and only generate pseudo-random numbers rather than true-random numbers, therefore, the circuit complexity needs to increase for SBG to improve the quality of randomness in the stochastic bit-streams. One attractive solution to overcome this short-coming of SC is to replace CMOS-based RNGs with spintronics-based RNGs [51-54]. This is because a single magnetic tunnel junction (MTJ) can be used as a true RNG (TRNG) with a tunable probability [46-50]. While these solutions significantly reduce the area and energy costs for SBG, implementing spintronic-based RNGs in SC does not reduce the overall delay costs since SBG steps are still performed separately from computation steps.

In this study, I present a SC scheme that uses MTJ-based hardware where SBG and computation steps are completely embedded. Our method exploits the intrinsic stochasticity of MTJs for SBG and utilizes the computational random access memory (CRAM) array to efficiently implement various SC tasks. Previous studies have demonstrated that true in-memory computing can be achieved in the CRAM architecture, thus eliminating circuit area costs and energy consumption for transferring data between memory and computation circuits. However, performing neuromorphic applications that require processing large amounts of data in conventional CRAM (conv-CRAM) will still suffer large area costs as well as noise sensitivity since the data is still encoded in binary bits. The solution presented here overcomes these short-comings by implementing SC within the CRAM array, which we refer to as SC-CRAM. While conv-CRAM embeds memory and logic arrays, SC-CRAM embeds SBG and computation arrays. I demonstrate the effectiveness of SC-CRAM in three neuromorphic applications: Local image thresholding, Bayesian inference for object location, and Bayesian belief network for heart disaster prediction. These tasks were also simulated in conv-CRAM and CMOS-based SC for performance comparison.

The rest of this chapter is organized as follows. Section 5.2 provides an overview of SC-CRAM, including task scheduling, implementation of basic and complex functions, and its advantages over conv-CRAM. In section 5.3, I elaborate on the three example applications where SC-CRAM can thrive. The output accuracy and performance evaluation of SC-CRAM is presented in section 5.4. Lastly, the chapter is concluded in section 5.5.

5.2. Overview of stochastic-CRAM (SC-CRAM)

In this section, I describe the key aspects of our proposed method of implementing SC

in the CRAM architecture. First, I explain basic task scheduling of SC-CRAM, using an AND gate for multiplication as an example. Then I describe the processes for performing more complicated arithmetic functions in SC-CRAM. Lastly, I describe preliminary advantages of SC-CRAM over conv-CRAM.

5.2.1 Task scheduling

The SC-CRAM method proposed combines the processes for MTJ-based SBG described in section 2.4 and logic operations in CRAM described in section 1.3.2. The SBG method in SC-CRAM is similar the method illustrated in Fig. 1.11(a) and 1.11(b), except each cycle includes an additional logic step between the perturb and read steps. Therefore, each cycle in SC-CRAM consists of synchronized reset, perturb, logic, and read pulses.

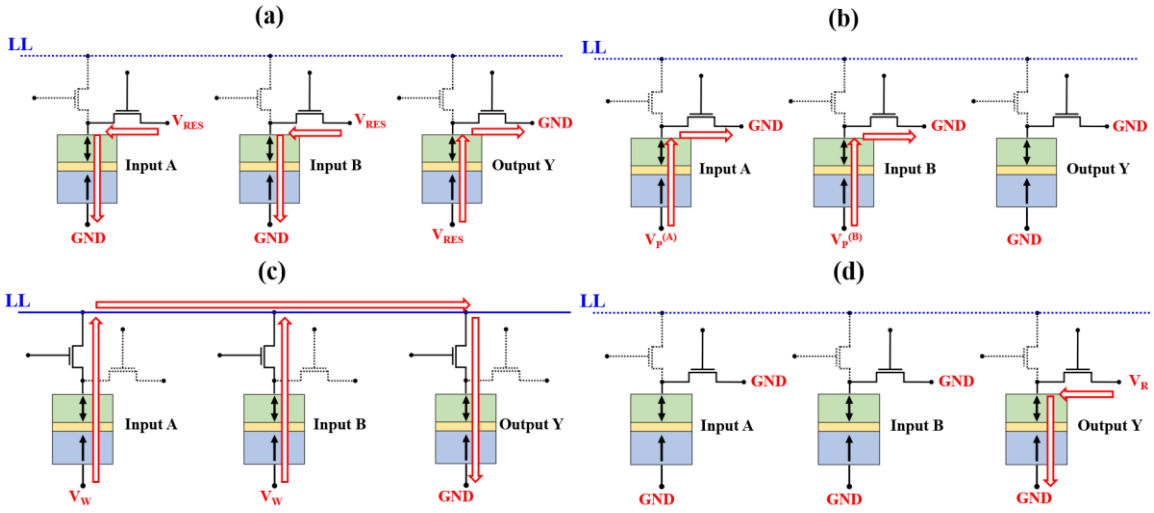


Fig. 5.1. Current paths for **(a)** reset, **(b)** perturb, **(c)** logic, and **(d)** read steps for multiplication of bit-streams A and B using AND logic in SC-CRAM.

Consider the example illustrated in Fig. 5.1(a) – 5.1(d), where AND logic in CRAM is used to multiply inputs A and B. Fig. 5.1(a) illustrates the current paths during the reset step, where V_{RES} initializes the input MTJs (A and B) to the P-state and the output MTJ

(Y) to the AP-state. Note that Y is initialized to the AP-state to meet the output preset criteria for AND logic in CRAM (recall Table 1.1).

The perturb pulses, shown in Fig. 5.1(b), are applied along the memory bit line (MBL) only for input cells A and B, which cause the MTJs to switch probabilistically, with probabilities dependent on $V_P^{(A)}$ and $V_P^{(B)}$. Note that the values for $V_P^{(A)}$ and $V_P^{(B)}$ can be determined three ways depending on the desired task. (1) In machine learning applications, one of the input cells can perform a synaptic function, therefore, V_P is dependent on the synaptic weight determined using learning algorithms. (2) Other tasks may require multiplication of a fixed constant, therefore, V_P is determined to produce a probability equal to the desired constant. (3) The most likely scenario is that V_P is dependent on the input data. For example, in image processing applications, $V_P^{(A)}$ corresponds to intensity at pixel A.

The logic step, shown in Fig. 5.1(c), replicates the logic operation performed in conv-CRAM (recall section 1.3.2). The voltage pulses applied at the bit select line (BSL) at A and B (V_W) follow the same criteria as V_{BSL} described in Table 1.1. During the perturb step, A and B switch probabilistically whereas Y switches deterministically during the logic step.

The final state of Y is read at the end of each cycle with V_R, shown in Fig. 5.1(d). As with the process shown in Fig. 2.8(a), V_R should be set so that it does not affect the state of Y. Alternatively, for tasks involving multiple computation stages, the steps at each computation stage can be overlapped so that the read steps at each stage prior to the final stage can perform the logic steps in the following stages. By adding a logic step in each cycle, SBG and computation are embedded within the same circuit.

5.2.2 Arithmetic functions

The implementation of SC in CRAM allows for various arithmetic functions to be performed efficiently. Previous studies on computation with stochastic bit-streams have demonstrated that a wide variety of functions such as exponential, linear gain function, and hyperbolic tangent can be performed in SC schemes with a minimal number of logic gates [44, 137]. Theoretically, any function performed in CMOS based SC can also be performed in SC-CRAM. However, in this study, we will only focus on the arithmetic functions needed in the three applications described in section 5.3. These functions include multiplication (described in the previous section), scaled addition, absolute valued subtraction, square root, and scaled division.

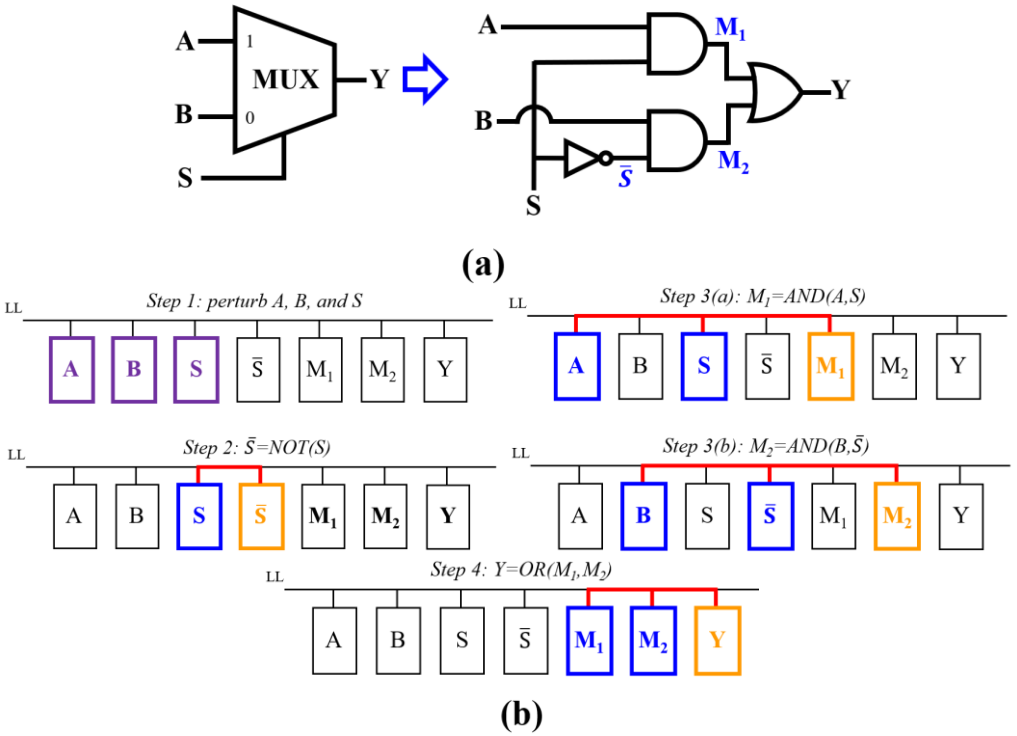


Fig. 5.2. (a) Circuit diagram of a 2-to-1 MUX and (**b)** scheduling operations in SC-CRAM. For simplicity, reset steps and final read steps are not shown.

Scaled addition can be done with either with MAJ3 logic or a MUX. In SC-CRAM, a MAJ3 gate consists of 2 input MTJs with variable probabilities (A and B), 1 input MTJ with a fixed probability of 0.5 (S), and 1 output MTJ (Y). Furthermore, MAJ3 logic in SC-CRAM has the same number of computation steps as AND logic. A MUX can also be implemented in SC-CRAM using the equivalent circuit diagram shown in Fig. 5.2(a). This circuit can be built in CRAM hardware with 2 input MTJs with variable probabilities (A and B), 1 input MTJ with any fixed probability (S), 3 intermediate MTJs ($\bar{S}$, M_1, and M_2), and 1 output MTJ (Y). Steps for implementing MUX logic in SC-CRAM consists of a perturb step on A, B, and S, one NOT logic step on $\bar{S}$, two overlapping AND logic steps on M_1 and M_2, and one OR logic step on Y, as shown in Fig. 5.2(b).

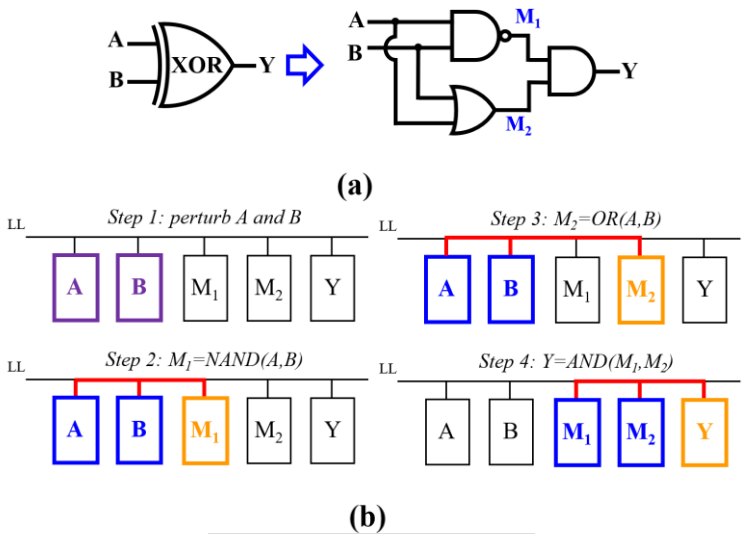


Fig. 5.3. (a) Circuit diagram of an XOR gate and (**b)** scheduling operations in SC-CRAM. For simplicity, reset steps and final read steps are not shown.

The advantage of using MAJ3 logic for scaled addition is that it requires less MTJs and less computation steps than MUX logic. However, S is limited to a value of 0.5 in MAJ3 logic whereas S can be any value for MUX logic. Furthermore, when performing scaled

addition on more than 2 inputs using MAJ3 logic, multiple layers of MAJ3 gates are required, which increases the number of MTJs and computation delay. Alternatively, MUX logic can handle more than two inputs without significantly increasing the number of computation steps. Therefore, for some applications, MUX logic is the preferred method for scaled addition.

Absolute valued subtraction of stochastic bit-streams A and B can be done using XOR logic. In CRAM, an XOR gate can be implemented using the equivalent circuit diagram shown in Fig. 5.3(a). It consists of 2 input MTJs (A and B), 2 intermediate MTJs (M_1 and M_2), and 1 output MTJ (Y). SC-CRAM steps for XOR logic consist of one perturb step on A and B, one NAND logic step on M_1, one OR logic step on M_2, and one AND logic step on Y, as shown in Fig. 5.3(b). It should be noted that XOR logic produces $Y \approx |A - B|$ only when bit-streams A and B have maximum correlation. In most cases, this is undesirable because computation on stochastic bit-streams typically require maximum decorrelation. However, the application described in section 5.3.1 is a special case and maximum correlation can be ensured using AND logic prior to XOR logic (process described in further details in section 5.3.1).

The square root of stochastic bit-stream X can be calculated using AND logic followed by two layers of OR logic, as shown in Fig. 5.4(a). The SC-CRAM circuit consists of 2 input MTJs (X_1 and X_2), 2 input MTJs with fixed probabilities of 0.67 and 0.18 (C_1 and C_2), 2 intermediate MTJs (M_1 and M_2), and 1 output MTJ (Y). SC-CRAM steps consist of one perturb step on X_1, X_2, C_1, and C_2, one AND logic step on M_1, and two consecutive OR logic steps on M_2 and Y, as shown in Fig. 5.4(b). Note that X_1 and X_2 both have probabilities equal to X, however, they are generated independently.

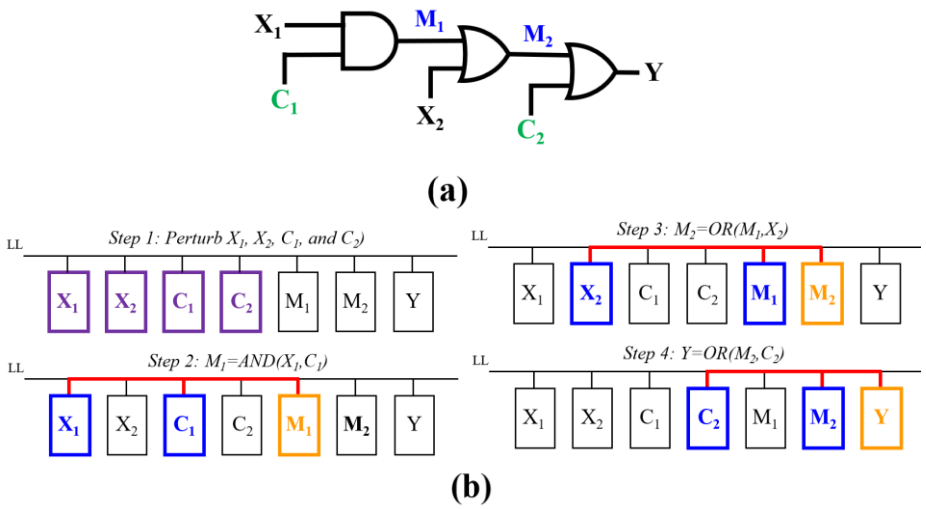
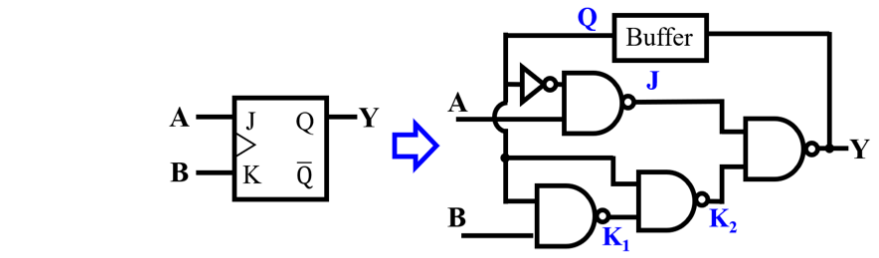
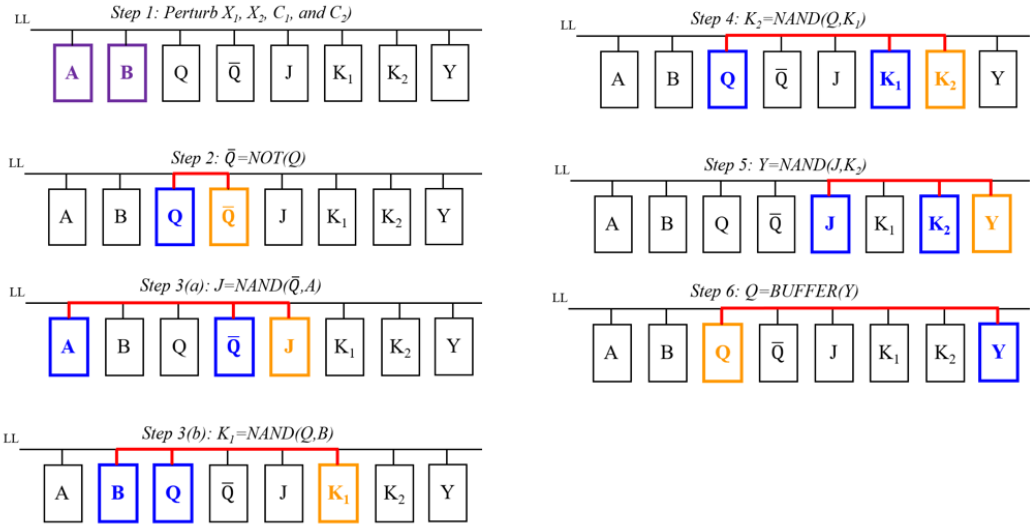


Fig. 5.4. (a) Circuit diagram the square root function in SC-CRAM where $Y \approx \sqrt{X}$, X_1 and X_2 are independent bit-streams with probabilities equal to X, and C_1 and C_2 are bit-streams with fixed probabilities of 0.67 and 0.18, respectively. **(b)** scheduling operations in SC-CRAM. For simplicity, reset steps and final read steps are not shown.

Scaled division of stochastic bit-streams A and B can be calculated using a JK flip-flop, which can be implemented in CRAM using the equivalent circuit diagram shown in Fig. 5.5(a). If A and B are applied to the J and K terminals, respectively, the output Y is expressed as Y ≈ A/(A+B). The SC-CRAM circuit consists of 2 input MTJs (A and B), 10 intermediate MTJs (Q, J_1, J_2, K_1, K_2, M_1, M_2, M_3, S_1, and S_2), and 1 output MTJ (Y). SC-CRAM steps consist of one perturb step on A and B, two consecutive XOR logic steps on M_1 and M_3 (5 logic steps total), one AND logic step on M_2 (overlapped with final XOR logic step on M_3), one XOR logic step on Y (3 logic steps), and one BUFFER step on Q, as shown in Fig. 5.5(b). Note Q is set to '0' for the first cycle.



(a)



(b)

Fig. 5.5. (a) Circuit diagram for a JK flip-flop to compute the function $Y \approx A/(A+B)$ and **(b)** scheduling operations in SC-CRAM. For simplicity, reset steps and final read steps are not shown.

5.2.3 Advantages of SC-CRAM

Table 5.1 shows the number of MTJs (N_{MTJ}), the number of computation steps (N_C), and area-delay product (ADP) for each arithmetic function described in sections 5.2.1 and 5.2.2 in SC-CRAM relative to their implementation in conv-CRAM. The ADPs in Table 5.1 were calculated by multiplying N_{MTJ} and N_C. Furthermore, calculations for each of these performance metrics compared computation on numbers with 8-bit resolution. For conv-CRAM, calculations were done on 8-bit digital numbers and for SC-CRAM, calculations were performed on bit-streams with $N_B = 2^8 = 256$ bits. When calculating

performance metrics in conv-CRAM, we assumed that addition and multiplication was performed using an 8-bit carry ripple adder (CRA) and an 8-bit Wallace tree multiplier (WTM), respectively. We also assumed that subtraction and division in conv-CRAM was performed using full subtractor (FS) circuits and non-restoring array dividers (NAD), respectively. Lastly, we assumed that square root in conv-CRAM was performed using three cycles of the Newton-Raphson (N-R) method.

Table 5.1. Number of MTJs and computation steps for various arithmetic functions in SC-CRAM (relative to conv-CRAM).

Function	Gates used for SC-CRAM	Number of MTJs	Number of computation steps	Area-Delay product
Scaled Addition	MUX	0.080X	14.3X	1.14X
	MAJ3	0.045X	10.7X	0.482X
Multiplication	AND	4.27×10^{-3}X	5.06X	0.022X
Absolute Valued Subtraction	XOR	0.055X	18.0X	1.00X
Scaled Division	JK flip-flop	0.013X	1.99X	0.026X
Square Root	AND-OR-OR	1.16×10^{-3}X	0.490X	5.68×10^{-4}X

**Assume 8-bit resolution for values shown.*

The values in Table 5.1 shows that SC-CRAM uses significantly less MTJs for all functions than conv-CRAM. The reduction in N_{MTJ} in SC-CRAM is the most significant for multiplication and square root functions but the least significant for scaled addition and scaled subtraction functions. This is due to the large number of full adders (FAs) needed for 8-bit WTM circuits and for the N-R method, which is not needed in SC-CRAM. However, N_C is larger in SC-CRAM than in conv-CRAM for all functions, except square root. This result is not unique to SC-CRAM and is observed when comparing any SC schemes to conventional deterministic schemes on binary numbers. The increase in N_C in

SC-CRAM is most noticeable for scaled addition but least significant for multiplication and scaled division. Interestingly, N_C is smaller for SC-CRAM than for conv-CRAM. This is likely because we assumed that conv-CRAM repeated the N-R method three times to approximate the square root function.

For most functions, the reduction in N_{MTJ} in SC-CRAM is larger than the increase in N_C. Therefore, it is not surprising that ADP is smaller for SC-CRAM than for conv-CRAM for most functions other than scaled addition with MUX and absolute valued subtraction. The most significant reduction in ADP for SC-CRAM is shown for multiplication, scaled division, and square root functions. The calculations shown in Table 5.1 indicates several key features of SC-CRAM. One is that the reduction in total circuit area in SC-CRAM out-weighs the increase in computation delay for tasks involving a large number of multiplication and scaled division functions (such as high-order polynomials, Maclaurin expansions of more complicated functions, etc.). Second is that both total circuit area and computation delay may decrease in SC-CRAM for more complex functions, as seen in the calculations for the square root function. While SC-CRAM is not well suited for tasks involving rapid, high-precision calculations, the results in Table 5.1 suggest that SC-CRAM is ideal for applications that involve rough approximations to detect statistical anomalies in large data sets.

5.3. Applications of SC-CRAM

In this section, we describe the three applications that were performed in this study: Local image thresholding for character recognition, Bayesian inference for object location, and Bayesian belief network for heart disaster prediction. These applications are illustrated in Fig. 5.6(a) – 5.6(c).

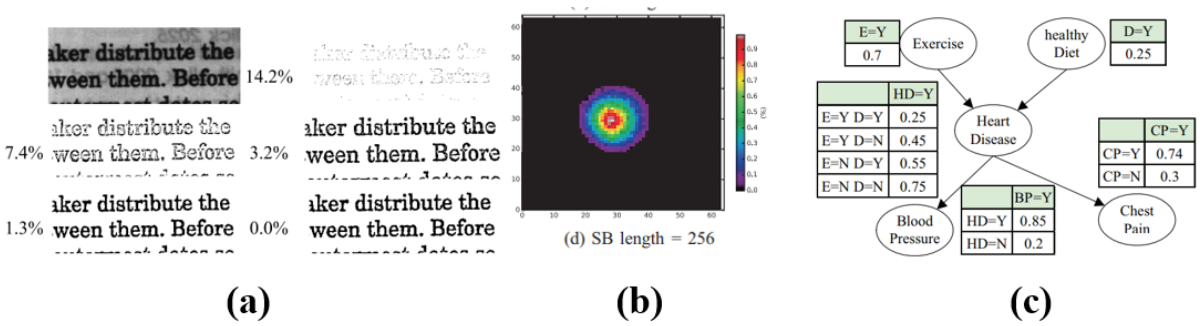


Fig. 5.6. Illustrations of the applications analyzed. **(a)** Effect of local binarization on a degraded image (extracted from [41]), **(b)** plot of data fusion results for object location (extracted from [53]), and **(c)** Bayesian belief network for heart disaster prediction (extracted from [53]).

5.3.1. Local Image thresholding

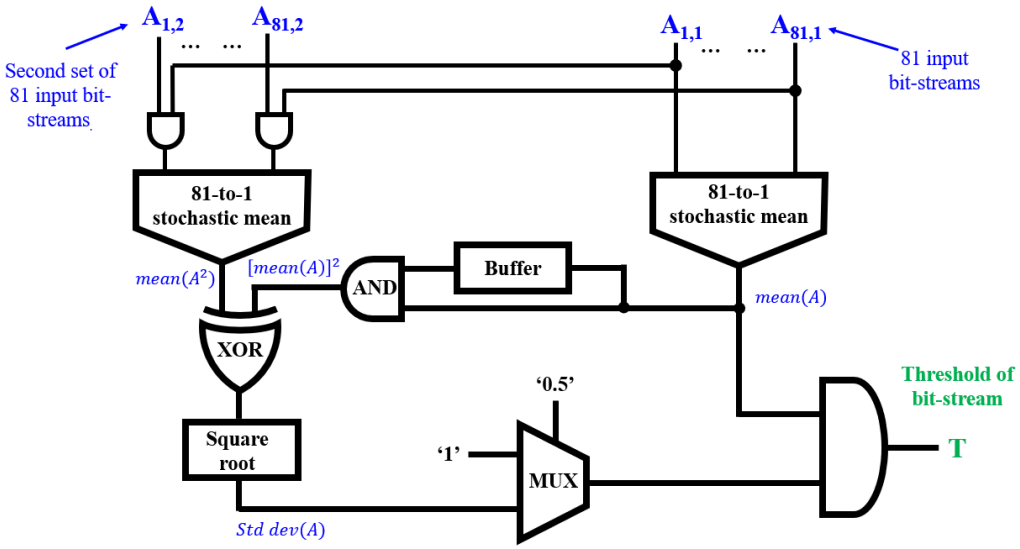


Fig. 5.7. Block diagram for calculating local thresholds in SC-CRAM using input bit-streams, A (based on pixel intensities). Note that bit-streams for $A_{i,1}$ and $A_{i,2}$ are independent but have equal probabilities.

Image thresholding is a very important step for optical character recognition. M. Najafi and M. Salehi applied SC to a local image threshold technique called the Sauvola method [41]. In this method, a window of n x n pixels is defined within a sub-section of the degraded input image. At each n x n sub-section, a threshold (T) is calculated using Eq.

5.1, which is based on the mean and standard deviation (see Eq. 5.2) of the pixel intensity (A) within the window. The pixel is set to '1' if a(x,y) > T(x,y) and '0' otherwise.

$$T(x,y) = \bar{A}(x,y) \cdot \left(\frac{\sigma A(x,y) + 1}{2}\right) \tag{5.1}$$

$$\sigma A(x,y) = \sqrt{|\overline{A^2} - (\bar{A})^2|} \tag{5.2}$$

The circuit for calculating Eq. 5.1 is shown in Fig. 5.7, which assumes a window size of 9 x 9 pixels [41]. The stochastic mean of pixels in A (A_1 to A_{81}) is calculated using MAJ3 logic, which uses 7 layers of MAJ3 gates where the 1st layer has 40 MAJ3 gates plus 1 AND gate, the 2nd, 3rd, 4th, and 5th layers have 20, 10, 5, and 3 MAJ3 gates, respectively, the 6th layer has 1 MAJ3 gate plus 1 AND gate. To calculate Eq. 5.2, the stochastic mean of A^2 ($\overline{A^2}$) needs to be determined. This is done by generating a second set of bit-streams for A, and bit-streams for A^2 are generated using AND logic between the two sets of A bit-streams. The same method of MAJ3 logic used to calculate $\bar{A}$ is used to calculate $\overline{A^2}$. The standard deviation of A also requires a bit-stream for $(\bar{A})^2$, which are generated using AND logic on $\bar{A}$. Note that a BUFFER is needed on one of the inputs of this AND gate, which shifts the bit-stream for $\bar{A}$ by one bit to avoid errors caused by high correlation. Next, absolute valued subtraction of $(\bar{A})^2$ and $\overline{A^2}$ is calculated with XOR logic using the process outlined in Fig. 5.4. The final step in calculating σA is generate bit-streams for the square root of $|\overline{A^2} - (\bar{A})^2|$, which is done using the process outlined in Fig. 5.5. A MUX is used to determine (σA + 1)/2 with the process outlined in Fig. 5 where one of the inputs is set to '1' and the selector input is set to 0.5. Finally, bit-streams for T(x,y) are generated using AND logic between bit-streams at the MUX output and $\bar{A}(x,y)$.

When compared to conventional computing methods, Najafi and Salehi found that the SC implementation reduced the execution time ~5 times, reduced power consumption ~3 times, and reduced hardware usage by ~4 times for computing with bit-streams with 256 bits [41]. The SC implementation achieved higher output accuracies than the conventional implementation when the noise injection rates reach 2% or above. Furthermore, the SC implementation achieved error rates below 5% at noise injection rates of 10% and error rates below 10% at noise injection rates as high as 40%.

The results presented in [41] are very promising and could be improved further with SC-CRAM for three key reasons. (1) The method in [41] uses a CMOS-based LSFR for SBG, which produces large area and delay costs. These costs are eliminated in SC-CRAM since SBG is embedded within computation. (2) Calculations for stochastic means in SC-CRAM can be done using MAJ3 logic rather than MUX gates which reduces the number of computation steps per cycle. (3) The hardware and computation steps for the square root function is significantly reduced in SC-CRAM compared to the method used in [41].

5.3.2. Bayesian Inference system

Data fusion is a process that integrates multiple data sources to find an optimal solution rather than relying on any individual source. The example simulated in this study is same example described in ref. [53], where a Bayesian inference system is used to determine the location of an object using distance (D) and bearing (B) data from three noisy sensors. The Bayesian inference mechanism produces a distribution of object locations by calculating the product series of conditional probabilities. The models for D and B for sensor j at position (x,y) must satisfy the Gaussian distributions described by Eq. 5.3 and Eq. 5.4, respectively. In these equations, μ_{dj} is the Euclidean distance between sensor j and position

(x,y), $\theta_{dj} = 5 + \mu_{dj}/10$, μ_{bj} is the viewing angle of sensor j, and $\theta_{bj} = 14.0626°$. These models are used to calculate the object location probability at position (x,y), p(x,y), using Eq. 5.5.

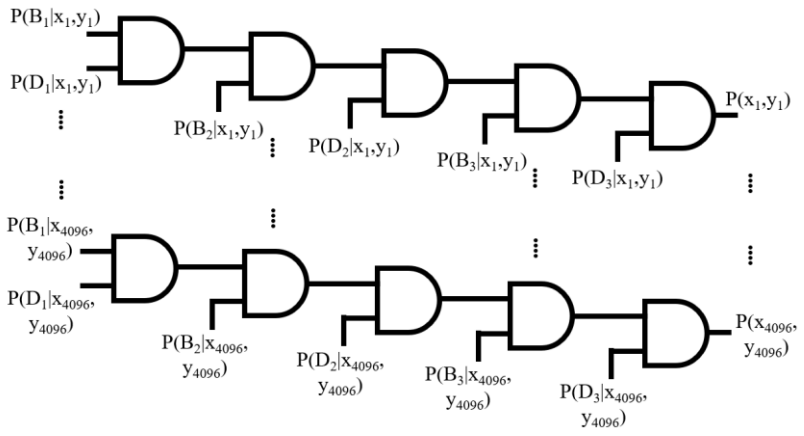


Fig. 5.8. Block diagram of a Bayesian inference system in SC-CRAM for object location on a 64 x 64 2D grid.

$$p(D_j|x,y) = \mathcal{N}(\mu_{dj}, \theta_{dj}) \tag{5.3}$$

$$p(B_j|x,y) = \mathcal{N}(\mu_{bj}, \theta_{bj}) \tag{5.4}$$

$$p(x,y) = \prod_j p(B_j|x,y) * p(D_j|x,y) \tag{5.5}$$

The problem investigated in this study obtains object location data on a 64 x 64 2D grid. The sensors are located at positions (0,0), (0,32), and (32,0) and the actual location of the object is (28,29). The circuit for calculating Eq. 5.5 at all 4096 grid locations is shown in Fig. 5.8. Each grid location uses 5 AND gates to calculate P(x,y), which requires 11 MTJs at each location when implemented in SC-CRAM (3 at the input AND gate and 2 for the remaining 4 layers). The total SC-CRAM circuit consists of 20,480 AND gates which are built from 45,056 MTJs.

5.3.3. Bayesian Belief Network

A Bayesian belief network (BBN) is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph. Figure 5.6(c) shows a BBN for heart disaster prediction (extracted from ref. [53]) which is the example simulated in this study. The parent nodes in this network are factors that cause heart disaster (HD), including exercise (E) and diet (D). The child nodes are clinical manifestations of HD including high blood pressure (BP) and chest pain (CP). Conditional probability tables for each node are shown in Fig. 5.6(c).

The probability of a heart disaster, P(HD), is shown in Eq. 5.6. In this expression, P(HD|E,D) represents the heart disaster probability when only considering E and D, P(BP) and P(CP) represent the heart disaster probabilities when the patient has high blood pressure and exhibits chest pain respectively. Equation 5.7 shows the expression for P(HD|E,D), where P(D), P(E), and P(E,D) represent the HD probabilities when the patient exercises regularly, has a good diet, and both, respectively.

$$P(HD) = \frac{P(BP)P(CP)P(HD|E,D)}{P(BP)P(CP)P(HD|E,D) + P(\overline{BP})P(\overline{CP})P(\overline{HD}|E,D)} \tag{5.6}$$

$$P(HD|E,D) = [P(E,D)P(D) + P(E,\overline{D})P(\overline{D})]P(E) + [P(\bar{E},D)P(D) + P(\bar{E},\overline{D})P(\overline{D})]P(\bar{E}) \tag{5.7}$$

The circuit for calculating (5) is shown in Fig. 5.9. In this circuit, P(HD|E,D) is calculated using three MUX gates, which is implemented in SC-CRAM using the same process outlined in Fig. 5.2. Values for P(BP)*P(CP)*P(HD|E,D) and $P(\overline{BP}) * P(\overline{CP}) * P(\overline{HD}|E, D)$ are calculated using 2 AND gates each, which adds 5 MTJs each. Note that an additional NOT gate is needed to calculate $P(\overline{HD}, E, D)$, which adds 1 more MTJ. The final

value for P(HD) is calculated using a JK flip-flop, which is implemented in SC-CRAM using the same process outlined in Fig. 6.5.

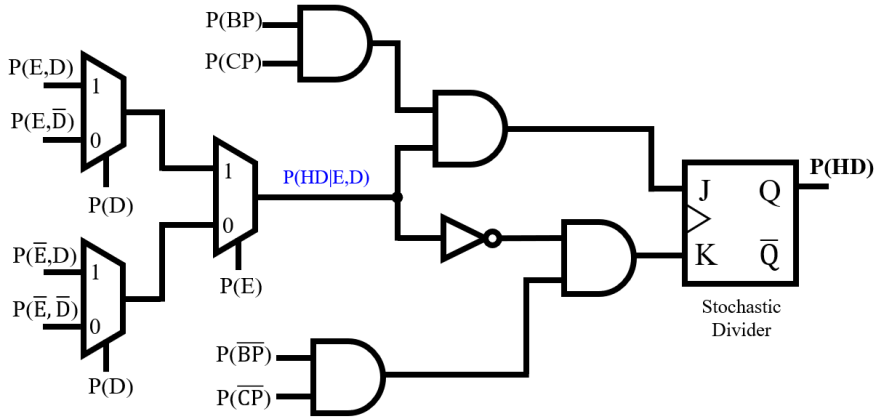


Fig. 5.9. Block diagram of a Bayesian belief network in SC-CRAM for heart disaster prediction.

5.4. Evaluation and results

Table 5.2. MTJ specifications for SC-CRAM performance analysis.

AP-state resistance	R_{AP}	2.5 kΩ
P-state resistance	R_P	1.0 kΩ
TMR ratio	TMR	150%
Thermal stability factor	Δ	60
Intrinsic critical switching voltage	V_{C0}	650 mV
Logic pulse width	t_W	5 ns
Perturb pulse width	t_P	2 ns
Fitting parameter for precessional switching	A_{FIT}	2.1 x 10^9 V^{-1}*s^{-1}

The performance of SC-CRAM for the three applications described in section 5.3 is evaluated. For each application, the output accuracy and noise resilience of SC-CRAM is demonstrated in section 5.4.1 and the execution time, circuit area, and energy usage are

compared with those of an STT-CRAM system using conventional binary computing methods in section 5.4.2. The specifications for the MTJ devices in Table 5.2. Note that A_{FIT} is the fitting parameter for precessional switching, where $\tau^{-1} = A_{FIT}(V - V_{C0})$. For my analysis, I assumed A_{FIT} = 2.1 x 10^9 $V^{-1}*s^{-1}$, which was experimentally determined in [85].

5.4.1 Accuracy Analysis

In this section, MATLAB was used to calculate the accuracy of the six arithmetic functions described in sections 5.3.1 and 5.3.2 as well as the three applications described in section 5.3.3. When analyzing the accuracy of the six arithmetic functions, the output probability was calculated for input probabilities ranging between 0.1 and 0.9. These calculations were repeated to account for variations in device resistances from their nominal ones. The source codes for each function are shown in Appendix A. When analyzing the accuracy of the three applications described in section 5.3.3, the percent error was calculated at multiple noise injection rates between 0% and 40%. For these applications, noise injection was defined as the percentage of bit-flips during the calculations. The source codes for these applications are shown in Appendix B.

Figures 5.10(a) to 5.10(f) shows the accuracy of the SC-CRAM method for multiplication using AND logic, scaled addition using MAJ3 logic, scaled addition using MUX, absolute value subtraction using XOR logic, square root using AND-OR-OR logic, and scaled division using a JK flip flop. Each function was repeated for inputs ranging between 0.1 to 0.9, where the input probabilities for A and B were equal for all functions with two exceptions. One is that input B for the XOR function is equal to A^2, which is done by generating two independent bit-streams for A, then introducing an extra AND gate at

the input to generate an A^2 bit-stream. The other exception is that B is set to be 1-A for the JK flip flop. The error is calculated and averaged over 100 trials and each curve is repeated for variations in the device resistance 5% and 10%. Note that variations in device resistance will also influence V_{C0} and Δ. Recall Eq. 3.11 and recall that $\Delta = H_K*M_S*Vol/2*k_b*T$. From these definitions, it can be approximated that a 1% change in resistance will also cause a 0.1% change in V_{C0} and a 1% change in Δ.

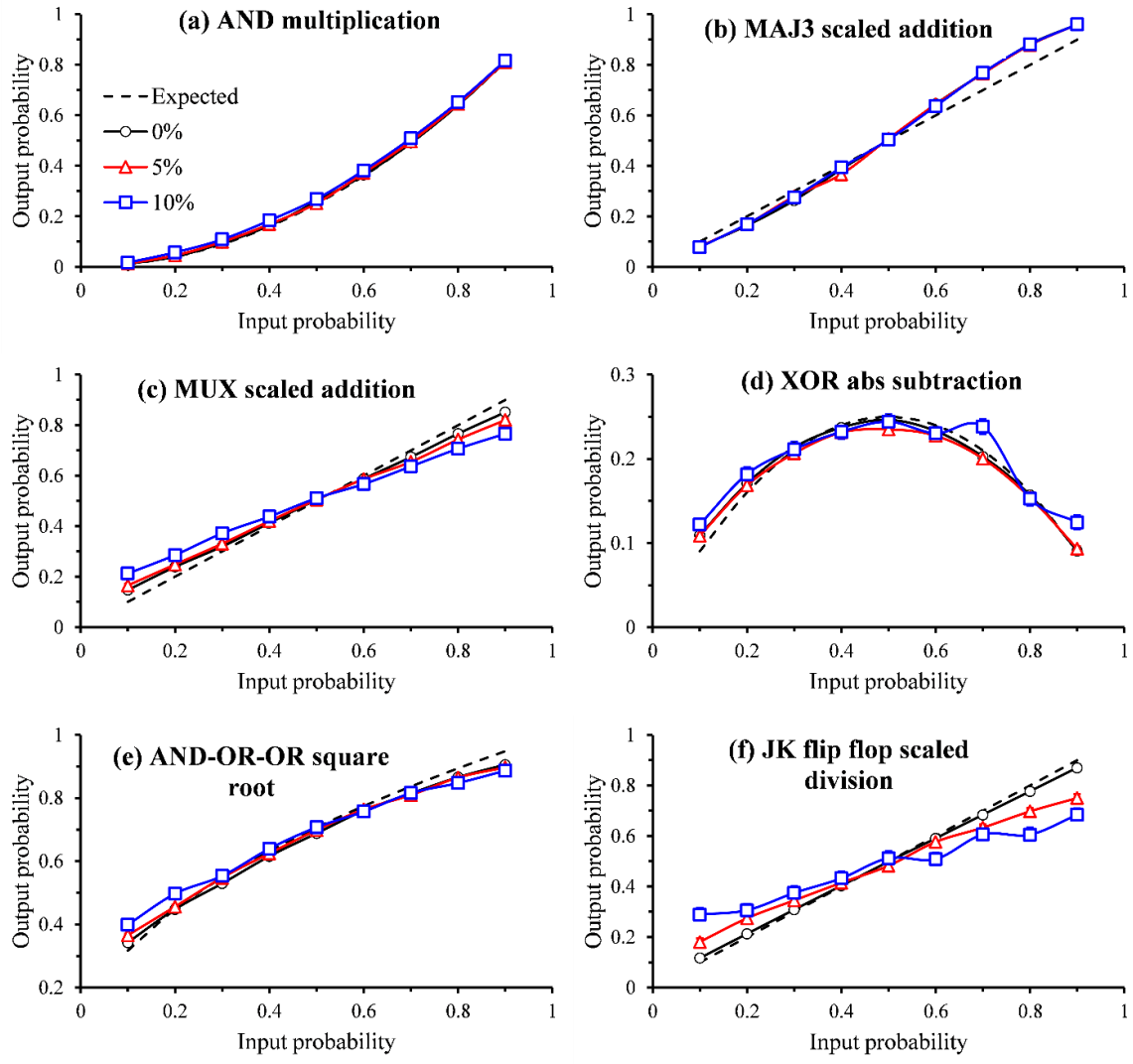


Fig. 5.10. Accuracy of **(a)** AND multiplication **(b)** MAJ3 scaled addition, **(c)** MUX scaled addition, **(d)** XOR absolute value subtraction **(e)** square root with AND-OR-OR logic, and **(f)** scaled division with a JK flip flop in SC-CRAM.

The curves shown in Fig. 5.10(a) shows that the output probability for AND multiplication is nearly equal to the expected value, even for R_{MTJ} variations of 10%. Figures 5.10(b) and 5.10(c) show that scaled addition calculations are accurate and noise resilient for calculations involving input probabilities around 0.5, however, show inaccuracies near input probabilities close to 0 and 1. It should be noted that the calculations are still within 0.05 of the expected value so the calculations can still be considered as accurate approximations. Furthermore, the MUX scaled addition is more susceptible to variations in R_{MTJ}, which is likely due to the additional computation steps required compared to the MAJ3 method of scaled addition. The calculations for absolute valued subtraction and square root also are also accurate within 0.05 of the expected value, however, they also show slight susceptibility to variations in R_{MTJ}. However, the function that shows the most susceptibility to variations in R_{MTJ} is the scaled division function using the JK flip flop, which is also likely due to the additional computation steps required compared to the other functions. While some of these functions show more susceptibility to variations than others, it is not clear if this will ultimately influence the performance of SC-CRAM in applications with larger scale circuits.

Figure 5.11(a) is the degraded gray-scale image of a sign reading ‘SC-CRAM’ that was used to analyze the accuracy of the local image thresholding method in SC-CRAM. Figures 5.11(b) to 5.11(d) show the binarized image at the output of the local image thresholding circuit for noise injection rates of 0%, 10% and 30%. The map of the 2-D plane for object location show the results from the Bayesian inference system at noise injection rates of 0%, 10%, and 30%, are shown in Fig. 5.12(a) to 5.12(c).

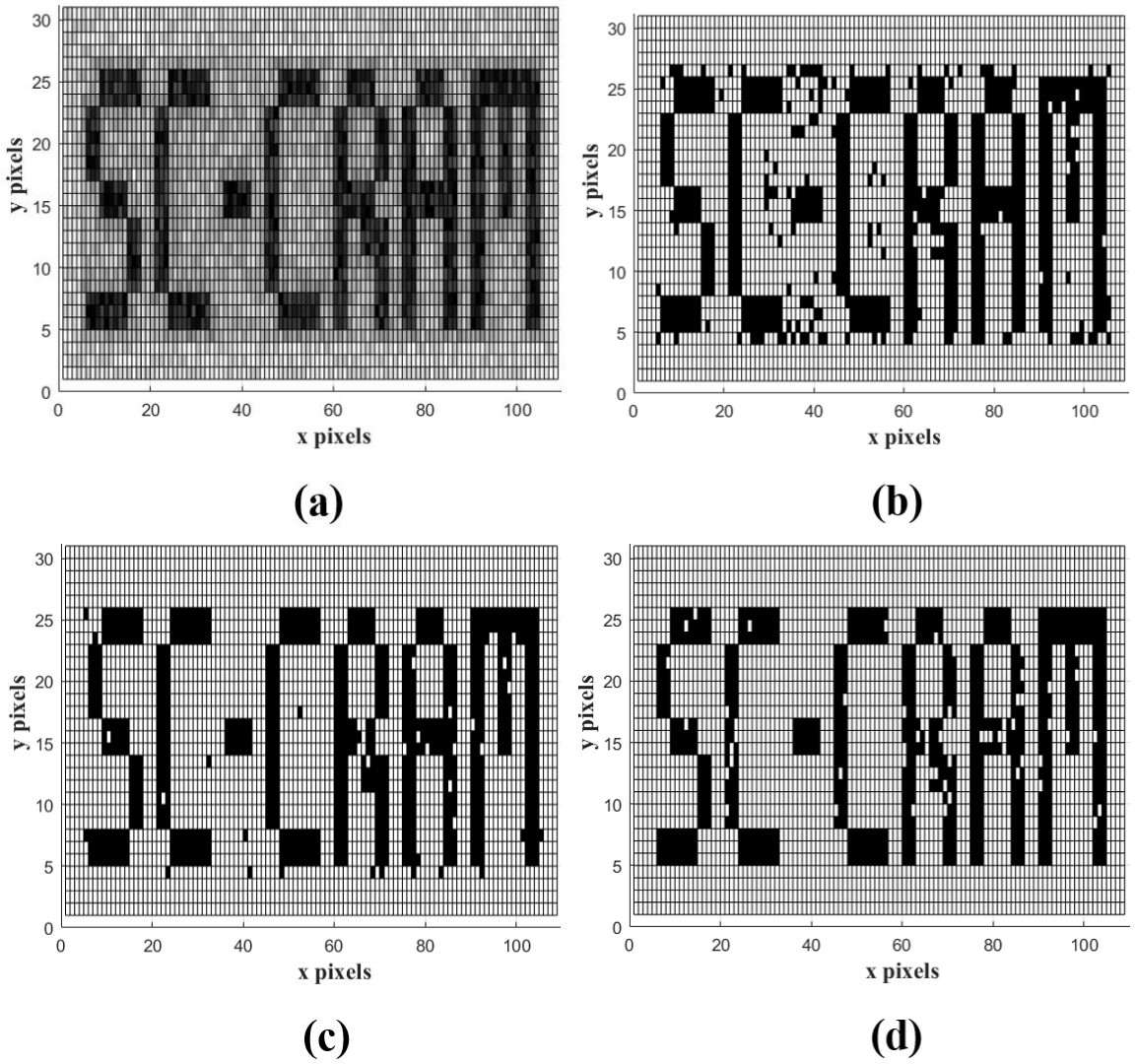


Fig. 5.11. Illustrations of **(a)** degraded gray-scale image for 'SC-CRAM' sign used to test the accuracy of local image thresholding in SC-CRAM. **(b-d)** Binarized image at output of local image thresholding circuit for **(b)** 0% noise injection, **(c)** 10% noise injection, and **(d)** 30% noise injection.

The output errors for local image thresholding, object location, and heart disaster prediction are shown in Fig. 5.13(a), 5.13(b), and 5.13(c), respectively. For local image thresholding, the output error is defined to be the percentage of inaccurate pixels. For the Bayesian inference system, the output error was determined by fitting the 2-D plot to a Gaussian curve and calculating the error in peak position as well as the deviation of the fitted curve, which was defined as the full-width half maximum of the Gaussian fit. For the

Bayesian belief network, the output error was simply defined as the difference between the average of the output bit-stream and the expected values.

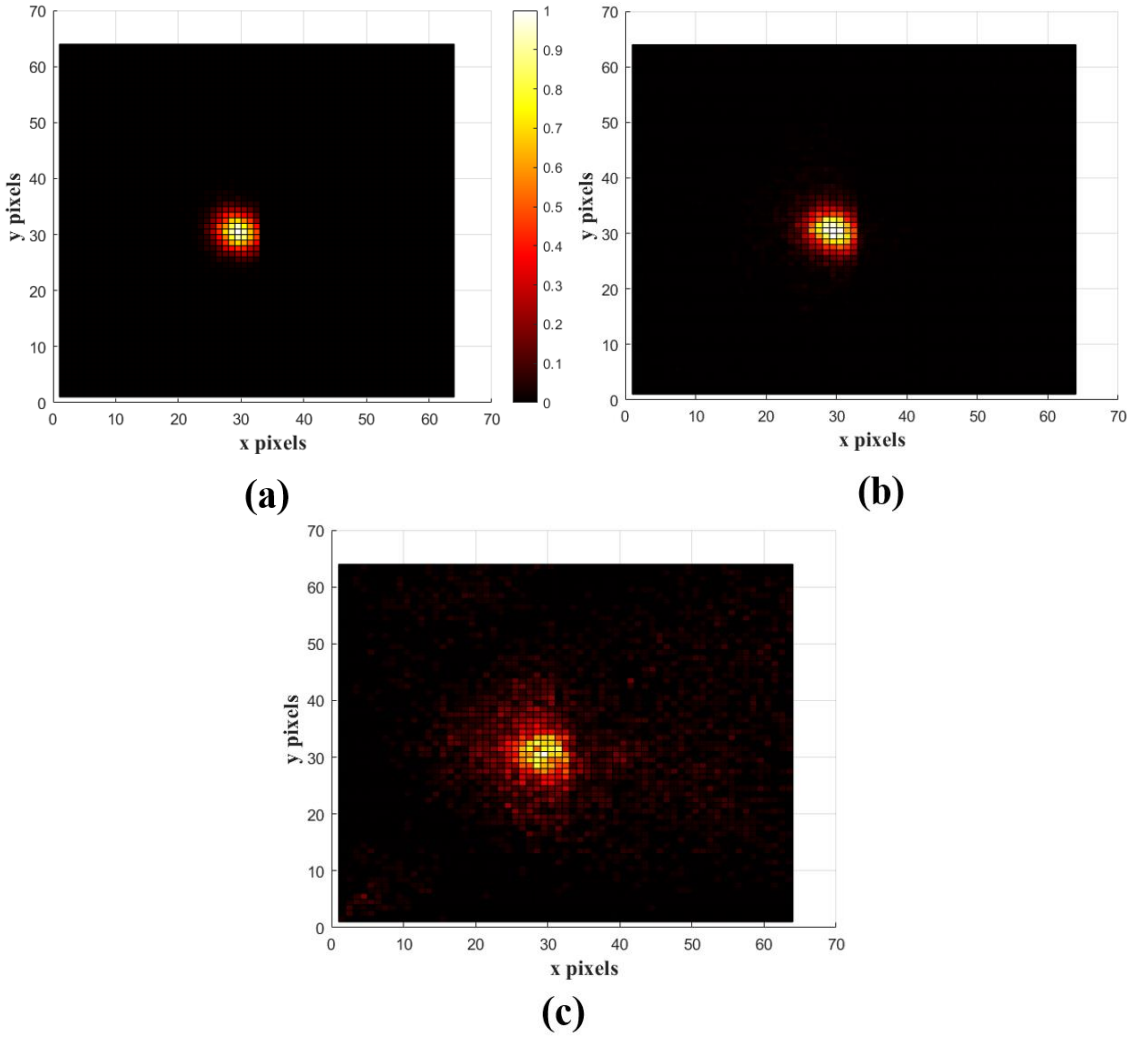


Fig. 5.12. The map of 2-D plane for object location via a Bayesian inference system with **(a)** no noise **(b)** 10% noise injection, and **(c)** 30% noise injection.

Figure 5.13(a) shows that the output error remains below 5% for noise injection rates up to 45%. Additionally, the output error actually decreases as noise injection rate increase from 0% to 10%. Then the output error becomes fixed at around 1% from 10% to 25% and finally increases beyond 25%. Figures 5.11(b) to 5.11(d) indicate that at low noise injection rates, the errors mostly occur in regions between the letters whereas at high noise injection

rates, the errors mostly occur on the letters. This suggests that SC-CRAM may under-estimate the threshold slightly, however, a slight amount of noise injection may correct for these errors.

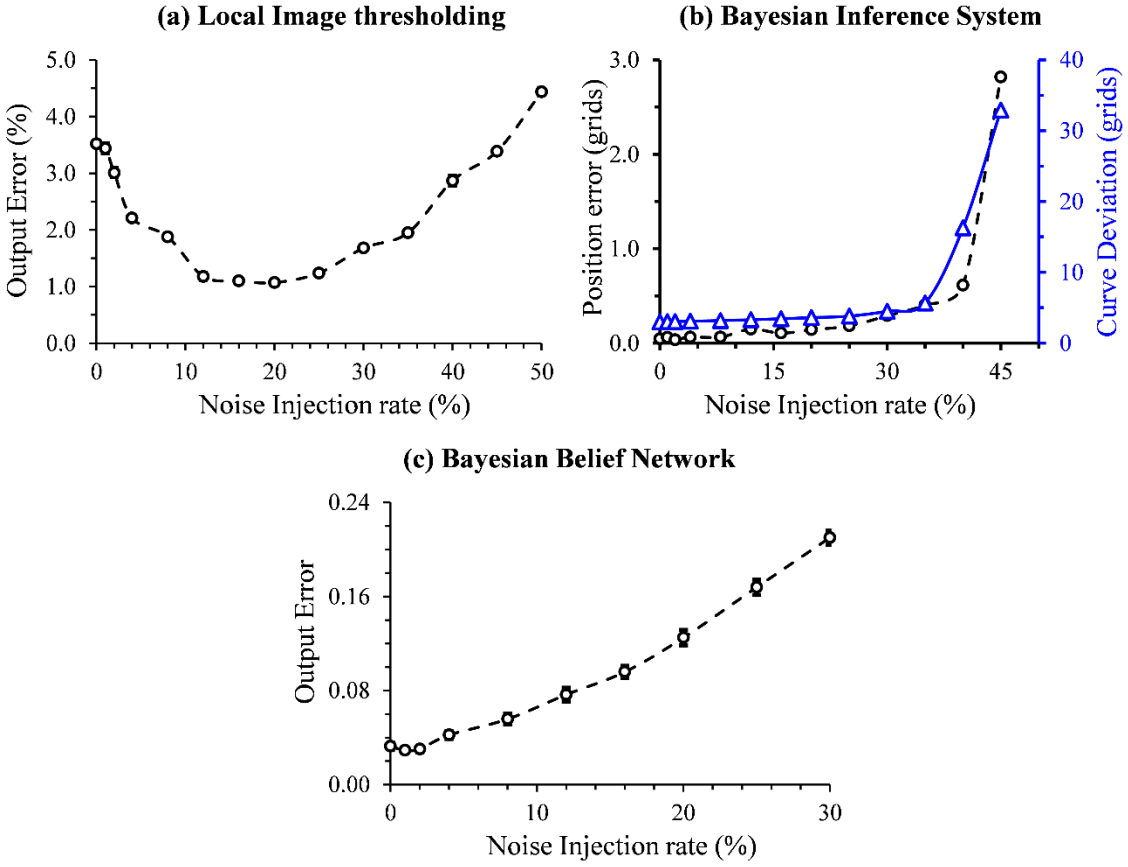


Fig. 5.13. Output accuracy vs noise injection rate for **(a)** local image thresholding **(b)** object location via Bayesian Inference, and **(c)** heart disaster prediction via a Bayesian belief network.

Figure 5.13(b) shows that the SC-CRAM Bayesian inference system can accurately locate the object within 1 grid position for noise injection rates up to 35%. Furthermore, the curve deviation remains below 5 grid points, meaning a low amount of uncertainty in the results. As with the position error, the curve deviation increases sharply for noise injections exceeding 35%. Lastly, Fig. 5.13(c) shows that the Bayesian belief network is the most sensitive to noise injection, however, the output error is still less than the noise

injection rate, which shows some degree of resilience to noise.

5.4.2 Performance evaluation

The performance of SC-CRAM and conv-CRAM were evaluated in terms of circuit area, computation delay, and total energy consumption (E_{TOT}). For each application described in section 5.3.3, the circuit area was defined in terms of the number of CRAM cells used (N_{CELLS}) and the computation delay was defined in terms of N_C. The total energy consumption was calculated by separating the energy consumption for perturb, logic, and reset steps. Note that one difference in E_{TOT} in SC-CRAM and E_{TOT} in conv-CRAM is that E_{TOT} in conv-CRAM does not include the perturb steps. Furthermore, unlike conv-CRAM, the logic and rest steps are repeated N_B times in SC-CRAM, which also need to be accounted for in the energy consumption calculations.

Table 5.3. Performance evaluation of SC-CRAM vs conv-CRAM. The values shown are relative to conv-CRAM.

Application	Number of sub-array rows	Number of sub-array columns	N_{CELLS}	N_C	Energy consumption
Local image thresholding	0.042X	0.014X	0.011X	0.535X	2.86X
Bayesian inference for object location	0.063X	0.037X	0.005X	2.37X	1.28X
Bayesian belief network for heart disaster prediction	0.091X	0.057X	0.0063X	1.20X	1.37X

Table 5.3 shows all of the calculations for N_{CELLS}, N_C, and E_{TOT}. Additionally, table 5.3 shows the number of rows and number of columns in the CRAM subarray. This data

shows that SC-CRAM uses at least 100 times less CRAM cells then conv-CRAM for all three applications. Furthermore, SC-CRAM has about 10 – 100 times less rows in the sub-array than conv-CRAM. This leads to significantly less IR drop due to parasitic effects [139]. However, table 5.3 also shows that SC-CRAM also has more computation steps than conv-CRAM for each application except local image thresholding. It should be noted that significant computation delays are common with any SC method and not specific to SC-CRAM. The reason that N_C is lower for SC-CRAM in local image thresholding is because solving a square root function using the N-R method requires at least 3 iterations to be accurate, which drastically increases N_C. This is also observed from the data in Table 5.1.

One key observation from table 5.3 is that N_{CELLS} for conv-CRAM is at least 100 times larger in conv-CRAM whereas N_C in SC-CRAM is, at most, 2.3 times larger. This indicates that the area-delay product in SC-CRAM is still lower than in conv-CRAM for all three of these applications. This can be seen from the E_{TOT} values in table 5.3. Intuitively, it is expected that the total energy consumption in SC-CRAM would be much larger than in conv-CRAM since the logic and memory write steps in SC-CRAM is repeated for N_B times, with additional perturb steps. However, table 5.3 shows the E_{TOT} for SC-CRAM is still on the same order of magnitude as conv-CRAM for all three applications. This is likely due to the significant reduction in N_{CELLS} in SC-CRAM.

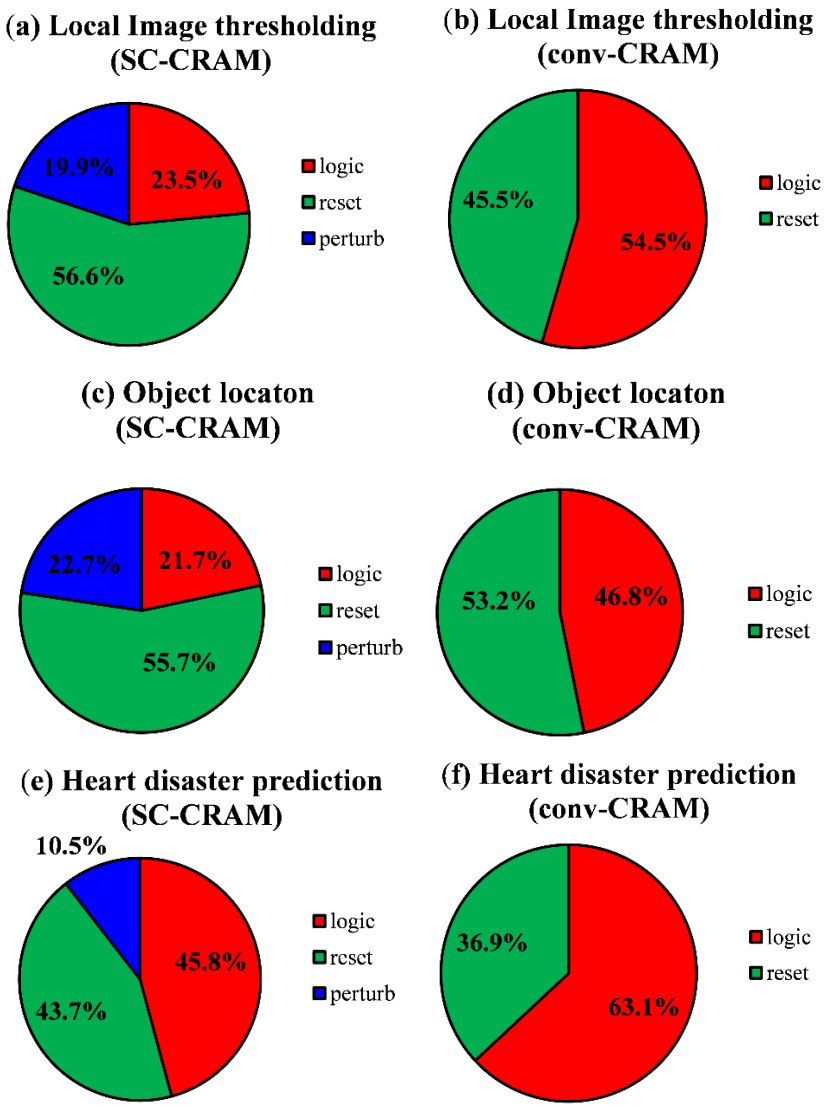


Fig. 5.14. Breakdown of energy consumption in SC-CRAM and conv-CRAM for **(a-b)** local image thresholding **(c-d)** object location via Bayesian Inference, and **(e-f)** heart disaster prediction via a Bayesian belief network.

The pie charts in Fig. 5.14 shows the breakdown of energy consumption for all three applications. These figures show that the reset step (or memory write step) accounts for nearly half of the total energy consumption for both SC-CRAM and conv-CRAM for all three applications. For SC-CRAM, this occurs because the reset step needs to be repeated to generate every bit within the bit-stream and for conv-CRAM, this occurs because of the

large number of MTJs that need to be written to. Unsurprisingly, the logic step accounts for the other half of the total energy consumption in conv-CRAM. In SC-CRAM, the logic step accounts for 25% of the total energy consumption, except for the Bayesian belief network where the logic step accounts for 45.8% of the total energy. This is because the Bayesian belief network requires a perturb step on a relatively small portion of the MTJs within the total circuit.

These charts show that adding the perturb step in SC-CRAM only accounts for a minority of the total energy consumption, whereas generating stochastic bit-streams with conventional CMOS-based methods could increase the total energy consumption by nearly 80% [38].

5.5. Summary

In this chapter, I introduced a method of using the CRAM architecture to generate and compute stochastic bit-streams called SC-CRAM. In SC-CRAM, the stochastic bit-stream generation and processing steps are embedded within the same circuit block. This means that there is no cost in the total circuit area and computation delay to generate stochastic bit-streams. While SC-CRAM does not out-perform conv-CRAM for simple functions, it can perform complex arithmetic functions with both circuit area and area-delay product being reduced by nearly 100 times when compared to conv-CRAM. When considering these reductions combined with the inherent noise tolerance of stochastic computing, SC-CRAM is a promising method for applications involving approximations and detecting statistical anomalies on large data sets. For my analysis, I considered three applications: local image thresholding, Bayesian inference for object location, and Bayesian belief network for heart disaster prediction for my analysis. My results indicate that SC-CRAM

benefits from smaller subarray sizes and similar energy consumption compared to conventional CRAM for these applications.

Chapter 6: Magnetic tunnel junctions for radiation resilient memory

This chapter is based on one of my previous publications, which is references [152] in the bibliography.

6.1 Motivation and background

In the previous chapters, I discussed how the intrinsic stochasticity of MTJs make them promising devices for non-conventional computing schemes. Additionally, in chapter 4, I showed how the interplay of two forces provide a pathway for MRAM cells with ultra-low switching current density. In this chapter, I demonstrate that MTJ devices also have potential as memory and logic cells for future high radiation applications such as space exploration. The contents presented in this chapter were published in [152].

For memory and logic applications in high radiation environments, such as computation on space probes and satellites, ionizing radiation (IR) can cause corruption of data in SRAM and DRAM cells [140]. This is because IR can cause shifts in the threshold voltage (V_{TH}) of the MOSFET devices and increase their leakage current. Total ionizing dose (TID) is one method of quantifying the impact of ionizing radiation and refers to the cumulative effects of exposure to multiple sources [141]-[144]. The shift in V_{TH} is caused by electron-hole pairs generated by large levels of TID, which tends to cause the less-mobile positive charge to build up at the MOSFET's gate oxide. The increase in leakage current is caused by dangling bonds at the MOSFET's gate/oxide interface produced by TID exposure. These dangling bonds can cause the leakage current to increase by one order of magnitude [143] and, in the worst cases, the increase in leakage current can prevent the MOSFET from switching off effectively [143].

There are several mitigation techniques currently used to reduce the effects of TID including fabricating chips on silicon-on-insulator substrates, implementing latch-up protection circuits, physical shielding, and using triple voting redundancy [143]-[144]. Alternatively, the current CMOS-based platform could be replaced with one that is more resilient to the effects of IR, thus reducing the need for these mitigation techniques. Spintronics is a promising candidates to replace CMOS-based technology for radiation-resilient memory and logic applications since recent experimental work has demonstrated that MTJs are exceptionally resilient to TID effects [145]-[148]. In 2012, F. Ren *et al.* showed that MTJs with in-plane magnetic anisotropy could be exposed to doses as high as 10 Mrad without changes in their coercivity, MTJ resistance, and TMR ratio [145]. In that same year, H. Hughes *et al.* tested STT switching characteristics and state retention of in-plane MTJs before and after TID exposure of 1 Mrad and found no differences in STT switching cycles as well as perfect state retention [146].

It should be noted that there are other potential candidates for radiation-resilient memory to replace CMOS-based memory and computing platforms, two main examples are 3D NAND flash memory and resistive random access memory (RRAM). M. Bagatin *et al* studied the effect of TID on 3-D NAND flash cells and found that TID exposure causes significant increases in their stand-by currents along with decreases in V_{TH} after exposed to TID levels of 50 krad (Si) [149], which is nearly 1000 times lower than the TID levels tested for MTJs in [145] and [146]. R. Fang *et al* demonstrated perfect state retention in HfOx-based RRAM cells after TID levels of 5.2 Mrad (Si) [150], which is on the same order of magnitude of the TID levels tested for MTJs. However, memory states in RRAM cells become more vulnerable to random bit flipping under electrical stress.

Recent studies have shown that p-MTJs are slightly more sensitive to the effects of TID than MTJs with in-plane anisotropy. N. Anuniwat tested the effect of a 1 Mrad dose of x-ray radiation on p-MTJs and showed that the exposure produced slight changes in the TMR ratio, which were more significant with larger devices [147]. However, these changes did not impact deterministic STT switching cycles and state retention. In 2019, B. Wang *et al.* studied the effects of TID on H_C and M_S of p-MTJs under doses as high as 20 Mrad [148]. Unlike the in-plane MTJs tested by F. Ren *et al.*, [145] the p-MTJs tested by B. Wang *et al.* showed slight degradation of H_C after TID exposure, however, M_S was unaffected. This suggest that in-plane MTJs and p-MTJs may have somewhat different responses to TID exposure and the small, but apparent, differences warrant further exploration.

In this chapter, I will present experimental results showing the influence of TID on three MTJ properties key for the STT-MRAM write operation, which are thermal stability, intrinsic critical switching voltage, and the write energy. In addition, I investigate secondary effects in these parameters due to MTJ size, switching direction, and switching speed for a more complete understanding of the underlying physical mechanisms. Previous experiments on the effects of TID on MTJs were primarily focused on the MTJ's deterministic switching properties. However, the results presented in this chapter utilize switching probability distribution curves to detect subtle changes in the aforementioned properties for STT-MRAM performance. Note that these experiments only investigate the influence of TID and do not consider single event effects [151].

The rest of this chapter is organized as follows. Section 6.2 describes the p-MTJ devices, the testing set-up, and radiation source used in these experiments. Results from DC current switching and probabilistic pulsed switching measurements are presented in

section 6.3. In section 6.4, I will discuss the implications of the results in section 6.3, then I will investigate potential physical mechanisms to explain these effects. Finally, this chapter is summarized in section 6.5.

6.2 Experimental set-up

TABLE 6.1 List of devices tested for analysis of the effects of total ionizing dose on p-MTJs.

Device No.	D_{MTJ} (approx.) (nm)	R_P (arb. units)	TMR (arb. units)	P_{SW} dir.
1	210	1	1.05	Both
2	210	1.01	1.12	P to AP
3	210	1.03	1.00	P to AP
4	135	1.49	1.04	P to AP
5	135	1.59	1.05	P to AP
6	125	2.41	101	P to AP
7	125	2.62	1.07	P to AP
8	125	2.65	1.13	P to AP
9	85	3.97	1.12	P to AP
10	85	4.35	1.07	P to AP
11	80	6.68	1.12	P to AP
12	80	6.78	1.09	P to AP
13	210	1.01	1.08	AP to P
14	210	1.03	1.12	AP to P
15	160	1.22	1.06	AP to P
16	160	1.23	1.06	AP to P
17	125	2.42	1.14	AP to P
18	125	2.57	1.11	AP to P
19	125	2.47	1.19	AP to P
20	85	4.22	1.11	AP to P
21	85	4.36	1.15	AP to P
22*	80	6.38	1.11	AP to P
23**	80	6.39	1.20	AP to P
24	80	8.28	1.06	AP to P

*DO NOT HAVE P_{SW} MEASUREMENTS BEFORE TID EXPOSURE.

**DO NOT HAVE DATA AFTER 1 MRAD EXPOSURE.

The stack of the p-MTJs tested consist of CoFeB free and fixed layers separated by an

MgO tunneling barrier and were patterned into circular nano-pillars. A total of 24 MTJs were tested, which had nominal diameters (D_{MTJ}) ranging between 80nm to 210nm. Details of all devices tested are shown in Table 6.1.

Two main experiments were conducted on each MTJ: 1) DC current sweeps (explained in section 2.2.2) and 2) pulsed switching measurements (explained in section 2.2.3). A sample plot of the data obtained through DC current sweeps is shown in Fig. 6.1(a). The current required for switching was found to scale roughly, as expected, with area from the smaller MTJs ($D_{MTJ} \leq 100$ nm) to the larger devices ($D_{MTJ} \sim 200$ nm). In these measurements, the current was swept in 1μA steps and held for 100ms for each step. The purpose of the DC current sweeps are to find the DC current required to switch the resistance states of the p-MTJs and to determine their resistance values and TMR ratio.

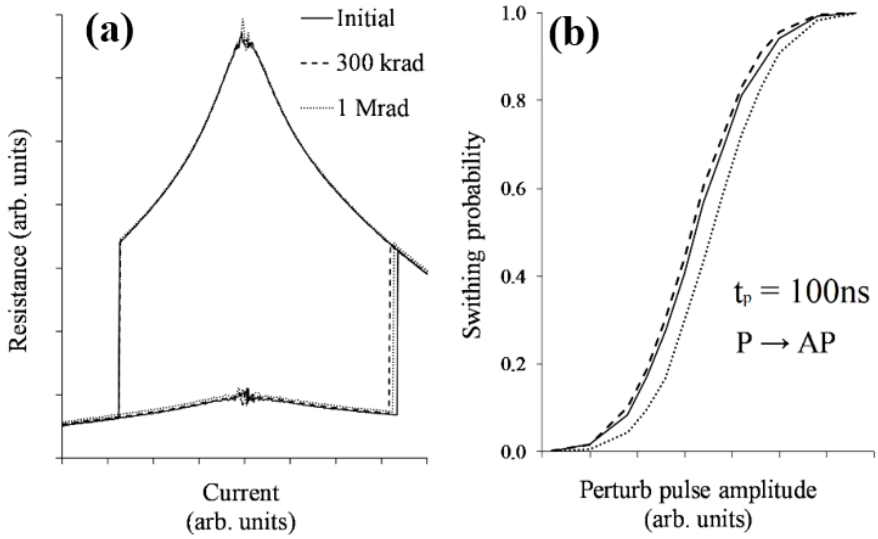


Fig. 6.1. Sample data acquired for device 12 (d_{MTJ} = 80nm). **(a)** STT hysteresis plot showing DC switching properties. **(b)** Switching probability distribution plot measured in the P to AP state switching direction for pulse widths of 100ns. Both plots display the curves obtained before TID exposure, after 300 krad, and after 1 Mrad. Image was extracted from [152].

The pulsed switching measurements in these experiments consisted of a series of at least 5,000 synchronized perturb and reset voltage pulses applied to the p-MTJs, as well as a constant read bias, which was set to 1mV so that it did not influence the state of the p-MTJs. From the data acquired, we obtained switching probability (P_{sw}) vs perturb pulse amplitude (V_P) curves for a given pulse width (t_p). These P_{SW} distribution plots were obtained at t_p = 50ns, 100ns, and 500ns. P_{SW} measurements were taken in the P to AP switching direction ($V_P > 0$) for devices 1 through 12 and in the AP to P state switching direction ($V_P < 0$) for devices 13 through 24 (see Table 6.1). Figure 6.1(b) shows a sample of a P_{SW} distribution plot for t_p = 100 ns measured in the P to AP switching direction.

Both P_{SW} and DC current measurements were taken before TID exposure, after 300 krad (Si) exposure, and after 1 Mrad (Si) exposure (additional 700 krad (Si)). TID exposure was conducted utilizing a high dose rate Co-60 (1.332 and 1.173MeV) source at NASA's Jet Propulsion Laboratory Total Ionizing Dose Laboratory at dose rates of 16.5 rad(Si)/s, which was mapped and verified by GafChromic radiology film. The p-MTJs were unbiased during exposure with the Si substrate connected to ground. All experiments and TID exposure were performed at room temperature unless otherwise specified.

6.3 Results

6.3.1 DC current switching results

There are three important device properties that can be obtained from the DC current switching hysteresis plots. The first is the currents required to switch from P-to-AP states and from AP-to-P states ($I_{P\rightarrow AP}$ and $I_{AP\rightarrow P}$ respectively). The second and third are the TMR ratio when $V_{MTJ} = 0$ (TMR(0)) and the DC voltage required for the TMR ratio to be reduced to half of TMR(0) (V_H), which describes the shape of the STT hysteresis plots and can be

determined using Eq. 6.1 [153]. Each of these properties reveal information regarding the intrinsic device properties and any changes allow for a deeper inspection of the impact of TID exposure.

$$TMR(V_{MTJ}) = \frac{TMR(0)}{1 + (V_{MTJ}/V_H)^2} \tag{6.1}$$

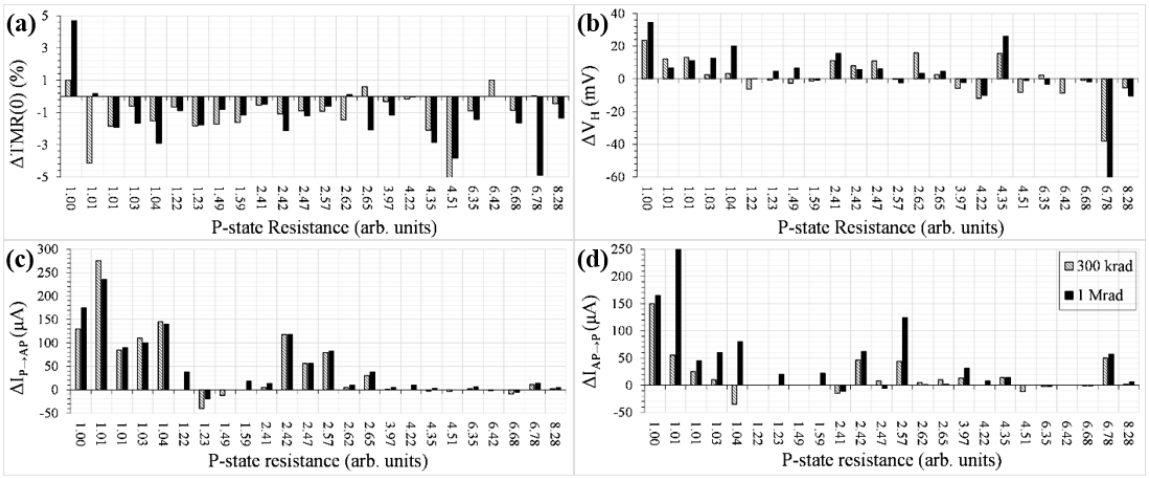


Fig. 6.2. Changes of **(a)** zero-bias TMR, **(b)** V_H, **(c)** $I_{P\rightarrow AP}$, and **(d)** $I_{AP\rightarrow P}$ before TID exposure, after 300 krad, and after 1 Mrad for every device tested. Images were extracted from [152].

The change in TMR(0), V_H, $I_{P\rightarrow AP}$, and $I_{AP\rightarrow P}$ after 300 krad (Si) and after 1 Mrad (Si) TID exposure for each device is shown in Fig. 6.2(a)-6.2(d). Figure 6.2(a) shows that TID exposure caused TMR(0) to decrease slightly in 21 of the 24 MTJs. However, most of these reductions were between 0.5% and 2%, which are small enough to be attributed to variations in contact resistance of the electrical probes rather than TID effects. Changes in TMR(0), ΔTMR(0), greater than 3% were observed in 3 p-MTJs and 1 p-MTJ showed a decrease in TMR(0) of nearly 5%. The results for ΔTMR(0) shown in Fig. 6.2(a) appear to be inconsistent between devices are consistent with the results presented in [147], where the authors attributed the increases in TMR(0) to local annealing induced by the x-ray irradiation via photoelectric effects and the decreases in TMR(0) to increased diffusion at

prolonged annealing.

Figure 6.2(b) shows that V_H increases after TID exposure for over half of the devices tested, which means that the rate at which the AP-state resistance drops with voltage will decrease. As with the TMR(0) measurements, many of the changes in V_H, ΔV_H, are less than 2% from the initial value and are therefore not large enough to be attributed to TID effects. These results show that ΔV_H, was increased by more than 10 mV for 7 p-MTJs tested and decreased by more than 10 mV for 3 p-MTJs.

Figures 6.2(c) and 6.2(d) shows that TID exposure caused both $I_{P\rightarrow AP}$ and $I_{AP\rightarrow P}$ to increase for most of the devices, however, many of these changes are less than 10 μA (less than a 5% change), which is not significant enough to be attributed to TID effects. Figure 6.2(c) shows that $\Delta I_{P\rightarrow AP} > 10$ μA for 10 p-MTJs and Fig. 6.2(d) shows that $\Delta I_{AP\rightarrow P} > 10$ μA for 13 p-MTJs. These figures also show that p-MTJs with larger D_{MTJ} had larger $\Delta I_{P\rightarrow AP}$ and $\Delta I_{AP\rightarrow P}$, where $\Delta I_{P\rightarrow AP} \geq 60\mu A$ and $\Delta I_{AP\rightarrow P} \geq 45\mu A$ for p-MTJs with nominal diameters of 210nm. On the other hand, for most of the p-MTJs with nominal diameters less than 100 nm, $\Delta I_{P\rightarrow AP}$ and $\Delta I_{AP\rightarrow P} < 10\mu A$, with only one exception.

Most of the changes shown in Fig. 6.2(a) – 6.2(d) are not significant enough to be attributed to TID effects. Furthermore, the changes that can be attributed to TID exposure are not consistent between devices and, in some cases, are not monotonic with TID, meaning that the magnitude of the changes decreases from 300 krad to 1 Mrad. The changes in $I_{P\rightarrow AP}$ and $I_{AP\rightarrow P}$ cannot be attributed to changes in TMR(0), since $|\Delta TMR(0)|$ never exceed 5%. The discrepancies between the 300 krad and 1 Mrad data sets seen in many of the devices suggests that TID exposure has multiple competing influences on the parameters investigated. It should be noted that the changes observed in Fig. 6.2 does not

necessarily influence STT-MRAM performance because the switching times for p-MTJs in STT-MRAM cells is several orders of magnitude lower than switching times during the DC current switching measurements. However, the results presented in Fig. 6.2 can be used to explain any changes observed in pulsed switching measurements.

6.3.2 Pulsed switching results

The switching probability distribution plots acquired from pulsed switching measurements were used to determine Δ, V_{C0}, and the write energy (E_{WRITE}). For these experiments, the pulse widths were > 10 ns, meaning that the primary mechanism for switching is thermal agitation [83], which means that our perturb pulses will be less than $0.8V_{C0}$. P_{SW} can be expressed in terms of t_p and τ using Eq. 6.2, where τ represents the average time required for the MTJ to switch states. From the thermal activation model [81], τ can be expressed in terms of V_P, Δ, V_{C0}, and τ_0 using Eq. 6.3. These equations show that when $\tau = t_p$, then $P_{SW} = 1 - e^{-1} \approx 0.63$, which occurs when V_P is equal to the critical switching voltage (V_C). Note that V_C is dependent on the pulse width whereas V_{C0} is an intrinsic property of the p-MTJ. By re-arranging Eq. 6.3 and setting $\tau = t_p$, V_C can be expressed in terms of t_p, Δ, and V_{C0}, as seen in Eq. 6.4.

$$P_{SW} = 1 - \exp\left(-\frac{t_p}{\tau}\right) \tag{6.2}$$

$$\tau = \tau_0 \exp\left(\Delta\left(1 - \frac{V_P}{V_{C0}}\right)\right) \tag{6.3}$$

$$V_C(t_p) = V_{C0}\left[1 - \frac{1}{\Delta}\ln\left(\frac{t_p}{\tau_0}\right)\right] \tag{6.4}$$

Typically, V_{C0} and Δ are determined by applying a linear fit to the $V_C(t_p)$ vs. $\ln(t_p/\tau_0)$ data [81]. However, this method is not sufficient for the data presented here for two reasons.

The first is that P_{SW} distribution curves were only obtained at three pulse widths, meaning that our Δ and V_{C0} calculations would be based on a linear fit of only 3 data points, thus resulting in high uncertainty. The second is that the $V_C(t_p)$ vs $\ln(t_p/\tau_0)$ plots typically show significant non-linearities for p-MTJs at $t_p < 300$ ns due to VCMA effects [81]. For these reason, the full P_{SW} distribution plots were used to calculate Δ and V_{C0} values at each pulse width separately. This was done by replacing t_p with τ and V_C with V_P in Eq. 6.4, where τ was calculated for each data point using Eq. 6.2. Then, a linear fit was applied to the V_P vs $\ln(\tau/\tau_0)$ data to determine Δ and V_{C0}.

For this study, changes in Vc were measured and used to calculate changes in the write energy (dE_{WRITE}) using Eq. 6.5 [85] for all devices. In this equation, R_{MTJ} is set to R_{AP} for AP to P switching and R_P for P to AP switching. Recall that changes in TMR(0) were only 1-2% for most of the devices tested and never exceeded 5% for any of the devices, meaning that neither R_P nor R_{AP} at low voltages changed significantly with TID exposure. However, TMR ratio is a function of voltage and V_H was modestly influenced by TID exposure, therefore, the R_{MTJ} values at large V_P may be influenced by TID exposure, even if V_C did not change significantly. Note that changes in R_{MTJ} are only considered for AP to P switching, since V_P does not have a strong influence on R_P (recall Fig. 6.1(a)).

$$E_{WRITE} = \frac{V_C^2 t_p}{R_{MTJ}} \tag{6.5}$$

The calculations for $d\Delta$, dV_{C0}, and dE_{WRITE} at t_p = 500ns are shown in Figs. 6.3(a)-6.3(b), Figs. 6.3(c)-6.3(d), and Figs. 6.3(e)-6.3(f), respectively, where Figs. 6.3(a), 6.3(c), and 6.3(e) shows the data for P to AP switching and Figs. 6.3(b), 6.3(d), and 6.3(f) shows the data for AP to P switching. Each plot shows results obtained after 300 krad and 1 Mrad

for every p-MTJ tested and the devices in each plot are organized by R_{MTJ}, where devices tested in the P to AP direction are organized by R_P and devices tested in the AP to P direction are organized by R_{AP}.

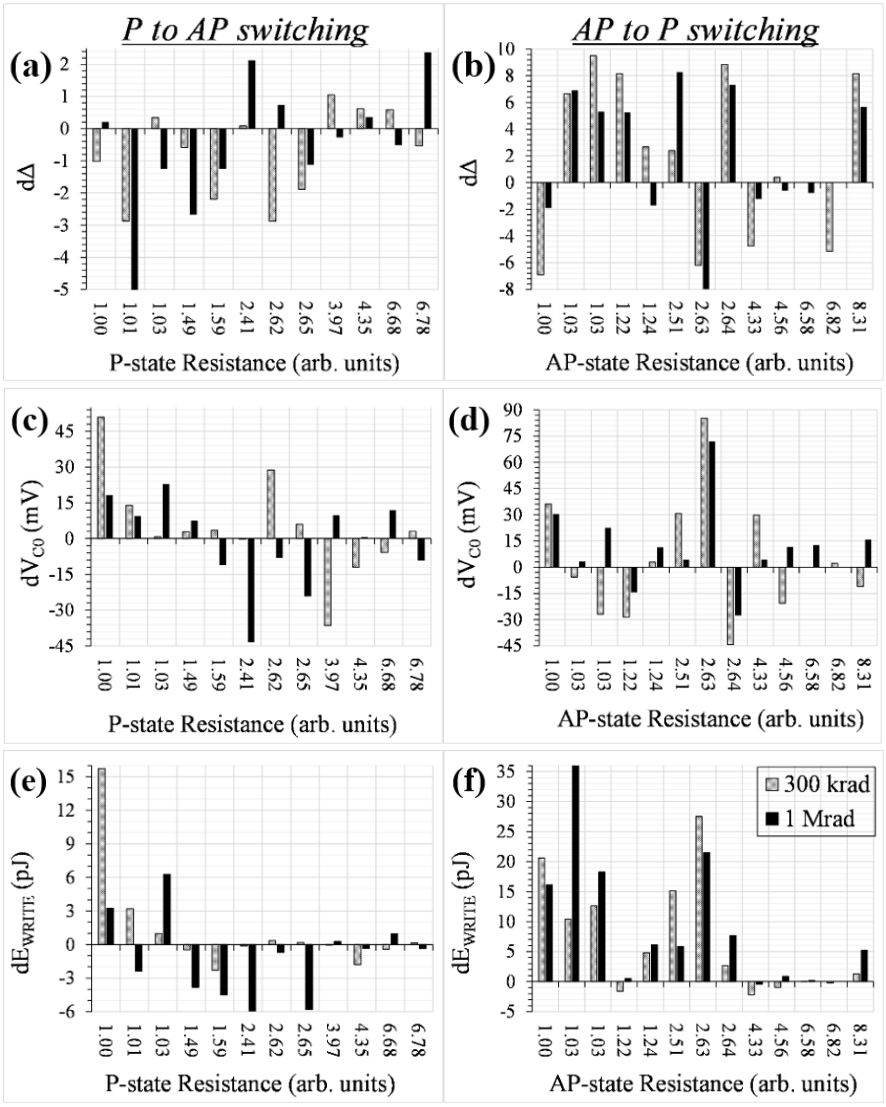


Fig. 6.3. Change in **(a-b)** thermal stability, **(c-d)** intrinsic critical switching voltage, and **(e-f)** write energy at pulse widths of 500ns in the **(a, c, e)** P to AP and **(b, d, f)** AP to P switching directions. Images were extracted from [152].

These plots show a few similarities in the ways that $d\Delta$, dV_{C0}, and dE_{WRITE} are affected by TID exposure. One similarity is that the magnitude of these changes are larger for AP

to P switching than for P to AP switching, which can be attributed to changes in V_H (recall Fig. 6.2(b)) since R_{AP} might be influenced at large V_P whereas R_P is not. Another is that larger changes were observed in p-MTJs with larger diameters. The final similarity is that the magnitude of these changes does not always increase, and may even decrease, from the 300 krad trials to the 1 Mrad trials. In some cases, the sign of $d\Delta$, dV_{C0}, and dE_{WRITE} switches from the 300 krad to the 1 Mrad trials. As with the results shown in Fig. 6.2(c) and 6.2(d), this data indicates that TID exposure has multiple effects on p-MTJs with competing influences on these parameters.

Figure 6.3(a) shows that Δ decreased slightly after TID exposure for P to AP switching for most of the devices. Furthermore, $d\Delta$ often fluctuated between 300 krad and 1 Mrad trials and even changed sign for some devices. Figure 6.3(b) shows us that Δ increased after TID exposure for AP to P switching for most devices, but also showed a decrease in Δ for a few other devices. Figures 6.4(a) – 6.4(f) illustrate the statistical variation $d\Delta$ at each pulse width, where Figs. 6.4(a), 6.4(b), and 6.4(c) shows data for P to AP switching at t_p = 500 ns, 100 ns, and 50 ns, respectively, and Figs. 6.4(d), 6.4(e), and 6.4(f) shows data for AP to P switching at t_p = 500 ns, 100 ns, and 50 ns, respectively. One key observation made from these plots is that even though some of the distribution curves shifted with TID exposure, their width does not significantly increase. This is important for STT-MRAM performance in rad-hard applications because it shows that the stability of the bits will not continue to change with continuing TID exposure. Furthermore, the sign of $d\Delta$ does not stay consistent with pulse width, for either switching direction. For example, $d\Delta$ is mostly skewed towards negative values for P to AP switching at t_p = 500ns but is approximately centered around 0 at t_p = 100ns and is significantly skewed towards positive

values at t_p = 50ns. Alternatively, for AP to P switching, dΔ is mostly skewed towards positive values at t_p = 500ns but is skewed towards negative values at t_p = 100ns and 50ns.

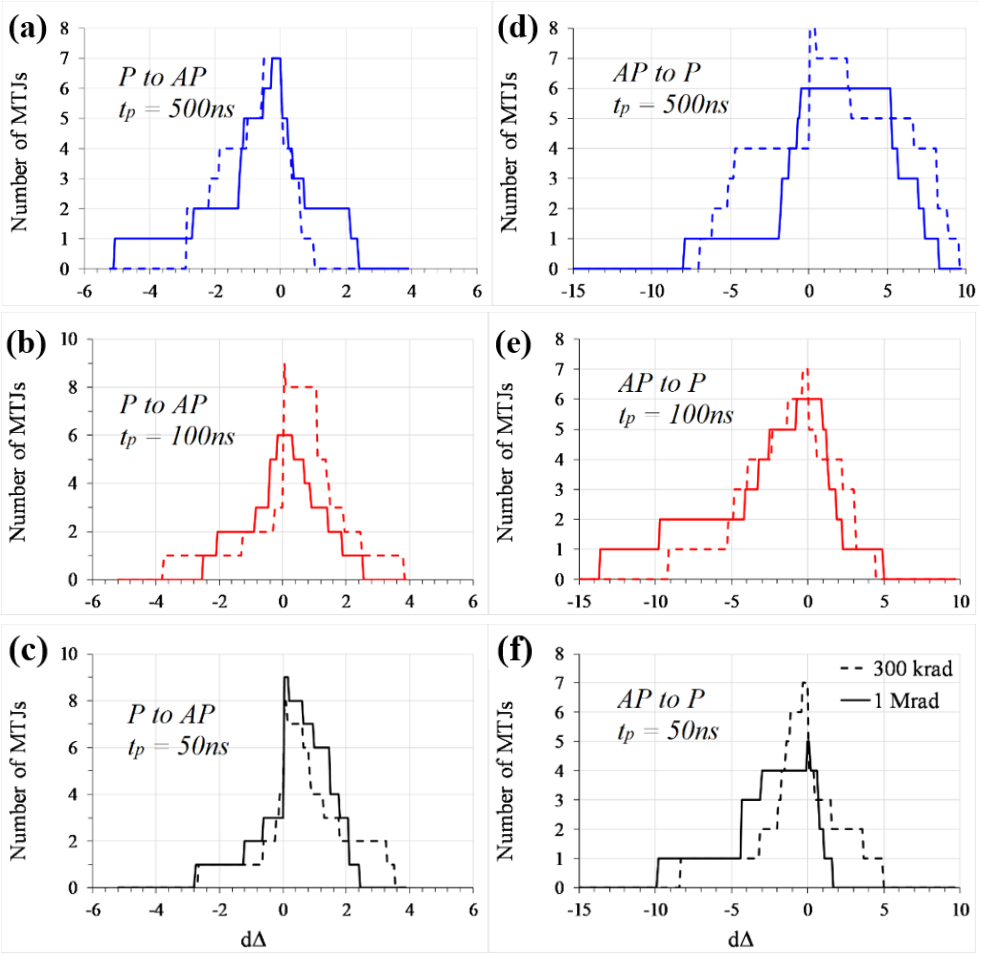


Fig. 6.4. Distribution of dΔ in the **(a-c)** P to AP and **(d-f)** AP to P switching directions at pulse widths of **(a, d)** 500ns, **(b, e)** 100ns, and **(c, f)** 50ns. Images were extracted from [152].

The influence of TID on Δ is consistent with the results in [148], which showed that TID caused variations in H_K for p-MTJs and their field switching plots showed that TID exposure had asymmetric influences between AP-to-P and P-to-AP switching. The results in [148] also suggest that there are multiple competing factors of TID effects on p-MTJs. The authors attributed the changes in Δ to Compton scattering as well as shifts in the exchange bias field in the fixed layer, which were the primary influences on H_K. Since H_K

$\propto \Delta$, these factors would also be applicable to the results presented here.

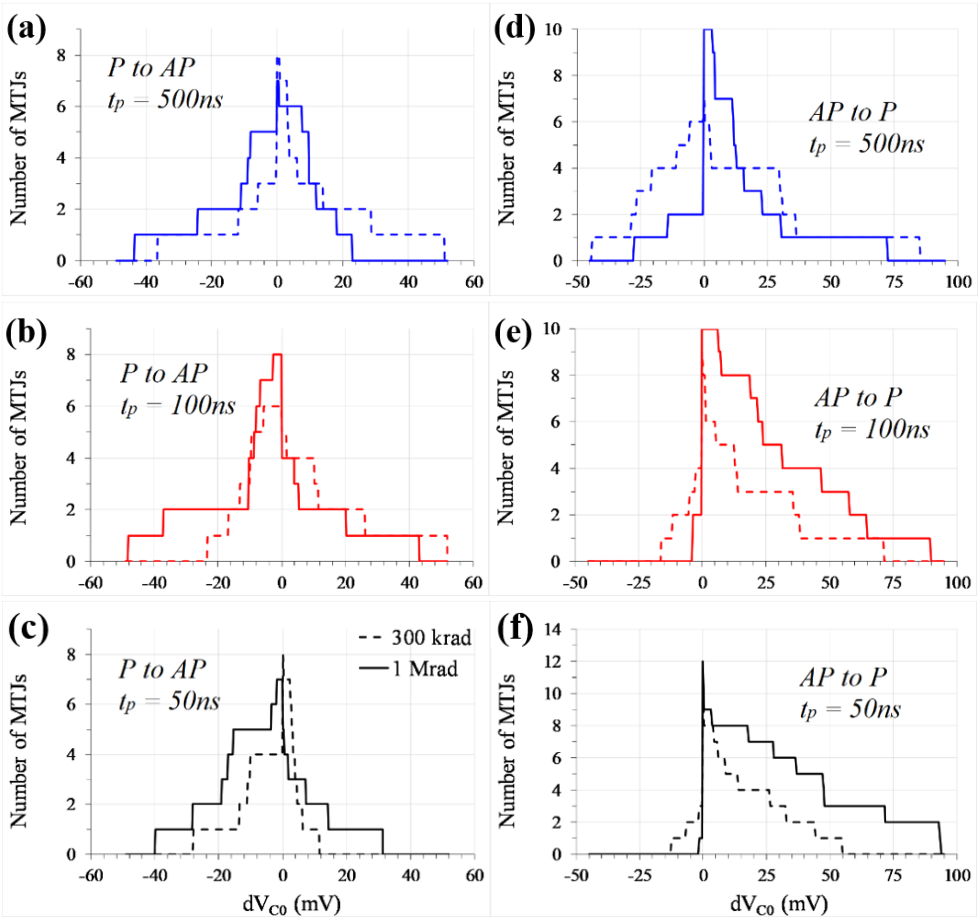


Fig. 6.5. Distribution of dV_{C0} in the **a-c.)** P to AP and **d-f.)** AP to P switching directions at pulse widths of **a, d.)** 500ns, **b, e.)** 100ns, **c, f.)** 50ns. Images were extracted from [152].

The explanations provided in [148] do not explain the dependence of $d\Delta$ on pulse width. One possible explanation for the dependence on pulse width is the influence of voltage on the interfacial anisotropy energy due to VCMA effects [154]. Since V_C increases as t_p decreases, it is possible that TID exposure affects the VCMA coefficient (ξ), which is related to the rate at which Δ changes with voltage, rather than the intrinsic Δ value. This would occur if IR exposure created dangling bonds at either the CoFeB(free)/MgO or CoFeB(fixed)/MgO interfaces, thus creating an electric field across the MgO tunneling barrier. Note that the sign of this IR induced electric field will depend on whether the

dangling bonds are created at the free layer/MgO interface or the fixed layer/MgO interface.

Figures 6.3(c) and 6.3(d) show that, at tp = 500 ns, dV_{C0} does not follow noticeable and well-defined trends. In both P to AP and AP to P switching directions, TID exposure caused V_{C0} to increase for about half of the devices tested and decrease for the other half. One observation seen in this data is that $|dV_{C0}|$ does not increase from 300 krad to 1 Mrad trials, meaning that increasing TID levels caused $|dV_{C0}|$ to saturate rather than continue to increase with higher doses, suggesting robustness in TID-rich environments.

Figures 6.5(a)-6.5(f) shows the distribution of dV_{C0} at each pulse width, where Fig. 6.5(a)-6.5(c) show the data for P to AP switching and Fig. 6.5(d)-6.5(f) show the data for AP to P switching. These plots reveal noticeable trends between TID level and dV_{C0} and how the trends differ between the two switching directions. When t_p = 100ns or 50ns, the distribution plots become slightly broader after 1 Mrad than after 300 krad for both switching directions, however, this did not occur when t_p = 500ns. In the P to AP direction, the dV_{C0} distribution is approximately centered at zero for t_p = 500ns, however, it becomes more skewed towards negative dV_{C0} as t_p decreases. Alternatively, for AP to P switching, dV_{C0} becomes more skewed towards positive values as t_p decreases. As with $d\Delta$, the dependence of dV_{C0} on pulse width can be attributed to changes in ξ.

The calculations for dE_{WRITE} for P to AP switching and AP to P switching at t_p = 500 ns are shown in Figs. 6.3(e) and 6.3(f) respectively. In the P to AP direction, TID exposure caused dE_{WRITE} to decrease for approximately 1/3 of the p-MTJs, increase for about 1/4 of the p-MTJs, and had no effect on the remaining devices. However, in the AP to P direction, TID exposures caused dE_{WRITE} to increase for most of the p-MTJs tested whereas any

decreases in dE_{WRITE} were relatively small and rare.

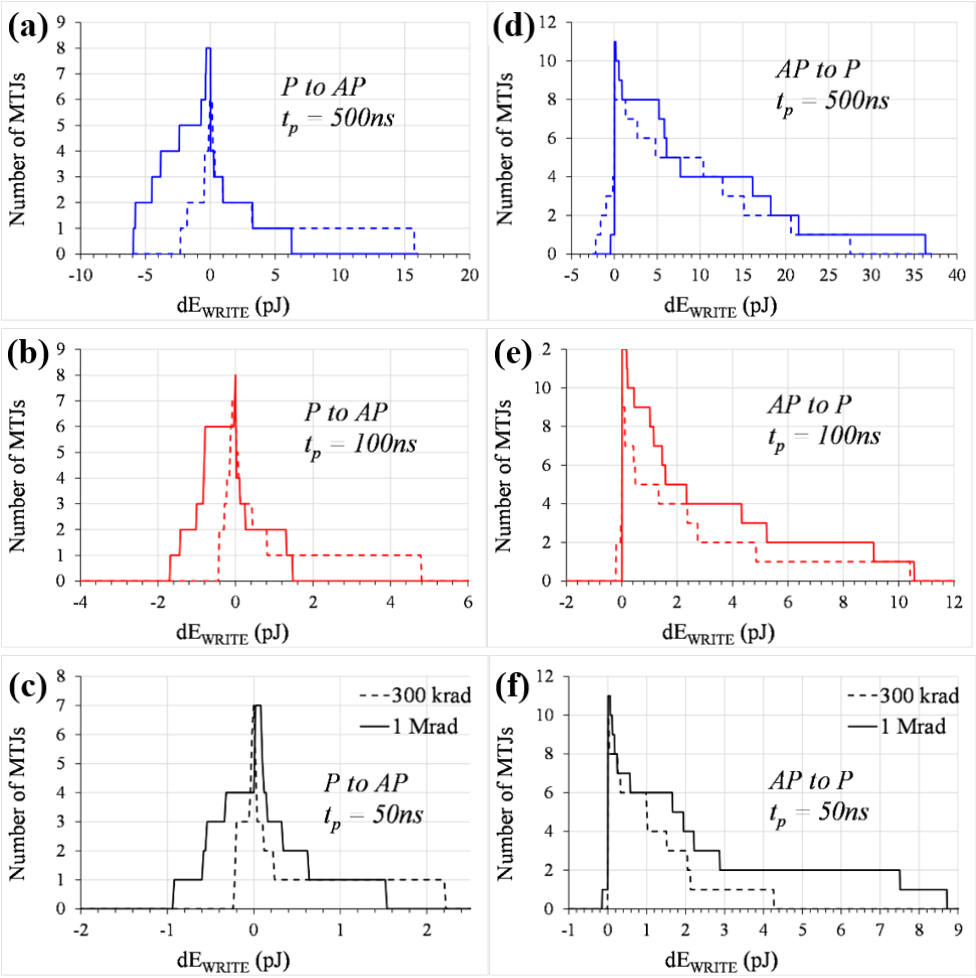


Fig. 6.6. dE_{WRITE} distribution in the **(a-c)** P to AP and **(d-f)** AP to P switching direction for pulse widths of **(a, d)** 500ns, **(b, e)** 100ns, and **(c, f)** 50ns. Images were extracted from [152].

The distribution plots of dE_{WRITE} for each pulse width tested are shown in Figs. 6.6(a)-6.6(f), where Figs. 6.6(a)-6.6(c) shows the data for P to AP switching and Figs. 6.6(d)-6.6(f) shows the data for AP to P switching. For P to AP switching, dE_{WRITE} tends to be negative and for AP to P switching, dE_{WRITE} tends to be positive. These plots also show that the distribution curves are much less skewed for P to AP switching, meaning that there is more overlap between positive and negative dE_{WRITE} for P to AP switching than for AP to P switching. Figures 6.3(c) and 6.3(f) suggested that both the magnitude and sign of

dE_{WRITE} fluctuates between 300 krad and 1 Mrad trials, however, Figs. 6.6(a) to 6.6(f) reveal that the overall dE_{WRITE} distribution becomes broader from 300 krad to 1 Mrad.

6.4 Discussion

The weighted average and a weighted standard deviation were calculated for the $d\Delta$, dV_{C0}, and dE_{WRITE} distribution curves, which represent the mean position and the width of the curves, respectively. The plots shown in Fig. 6.7(a) – 6.7(f) show these calculations with respect to pulse width. For each plot, the blue curves with circular marks are calculations for P to AP switching and the black curves with square marks are calculations for AP to P switching. Furthermore, the dotted lines and solid lines are calculations from data after 300 krad and 1 Mrad, respectively.

Figure 6.7(a) shows that Δ slightly increased for most of the devices tested when $t_p \leq$ 100ns and slightly decreased when t_p = 500ns for P to AP switching. On the other hand, Δ slightly decreased for most of the devices tested when $t_p \leq$ 100ns and slightly increased when t_p = 500ns for AP to P switching. The dependence that the average of $d\Delta$ has on pulse width suggests that TID exposure effects the interfacial anisotropy of the MTJ in such a way that it influences ξ. Figure 6.7(b) shows that the standard deviation of $d\Delta$ is larger than the average $d\Delta$ for both switching directions, which means that the changes in Δ with TID exposure varied significantly between individual p-MTJs and also varied between 300 krad and 1 Mrad trials.

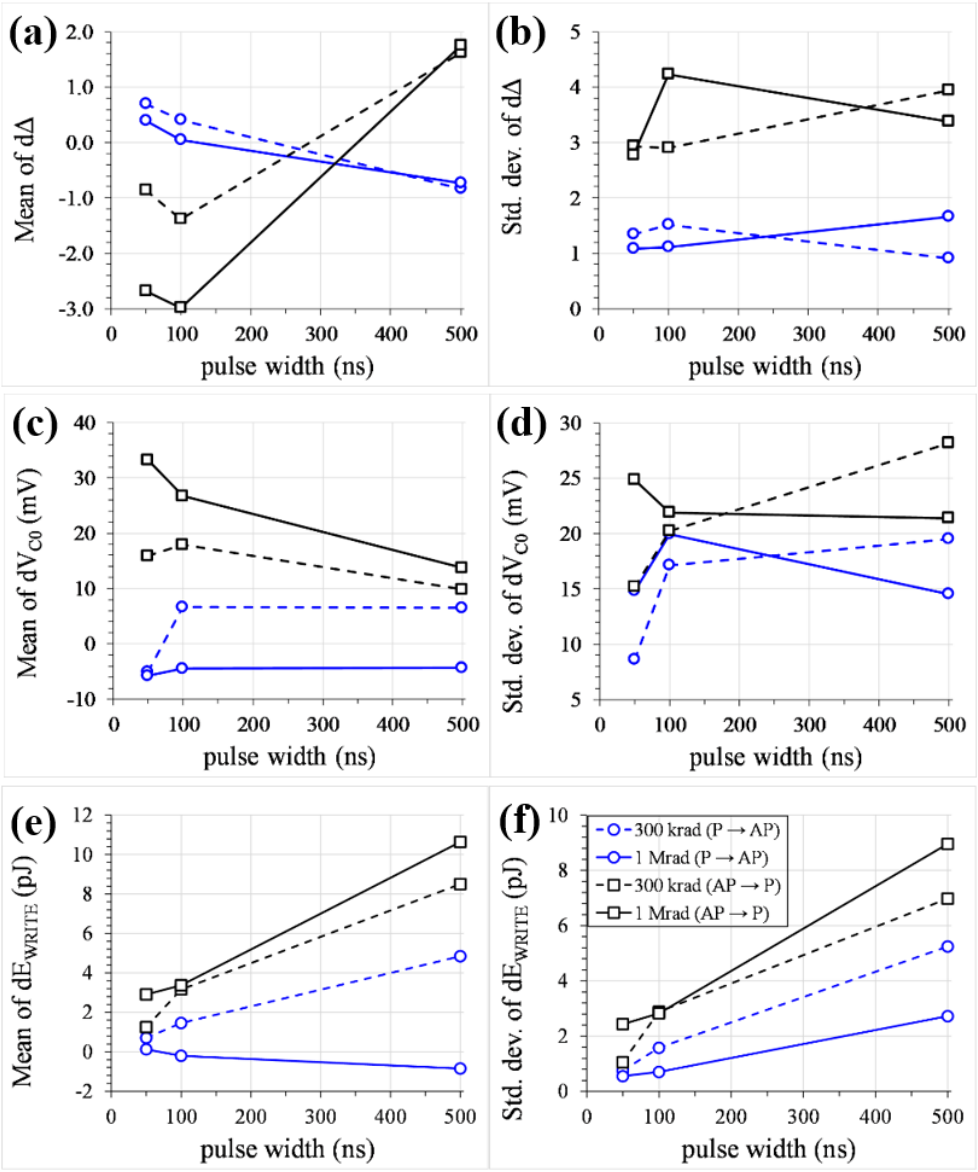


Fig. 6.7. Calculations for the mean and standard deviation of **(a-b)** dΔ, **(c-d)** dV_{C0}, and **(e-f)** dE_{WRITE} with respect to pulse width. Images were extracted from [152].

Figure 6.7(c) shows that the average $dV_{C0} > 0$ after 300 krad but $dV_{C0} < 0$ after 1 Mrad for P to AP switching. For AP to P switching, $dV_{C0} > 0$ after 300 krad and continues to increase after 1 Mrad. Figure 6.7(d) shows that the standard deviation of dV_{C0} is larger than the average of dV_{C0} for both switching directions. As with the dΔ results, the data shown in Fig. 6.7(c) and 6.7(d) indicates noticeable variation of dV_{C0} between devices as well as

between 300 krad and 1 Mrad trials.

A possible explanation for the behavior of dV_{C0} for AP to P switching is due to the change in resistance of the p-MTJ with voltage. Note that our DC measurements indicated that the resistance in either state did not change significantly, however, Fig. 6.2(b) showed that TID exposure caused V_H to increase, which means that TID exposure may cause R_{AP} to increase at higher voltages. Since V_{C0} is proportional to R_{MTJ}, meaning that the increase in R_{AP} will be most significant at smaller pulse widths, since larger V_P values are required for switching. This behavior of dV_{C0} for P to AP switching does not follow the same trends as for AP to P switching because R_P does not have as strong of a dependence on V_P than R_{AP} does.

Figure 6.7(e) shows that, for P to AP switching, the average E_{WRITE} increases after 300 krad, but then shows almost no change, or even a slight decrease, in E_{WRITE} after 1 Mrad. Recall that Fig. 6.3(e) shows that the sign of dE_{WRITE} varies between devices and even fluctuates between the 300 krad and 1 Mrad trials for some devices. The data shown in Fig. 6.7(e) suggests that TID exposure may cause temporary changes in E_{WRITE}, however, for P to AP switching, these changes may not be subject to long-term effects. Alternatively, for AP to P switching, Fig. 6.7(e) shows that, E_{WRITE} increases after 300 krad and continues to increase after 1 Mrad. Furthermore, Fig. 6.7(f) shows that the standard deviation of E_{WRITE} also continues to increase with increasing TID exposure for AP to P switching, but not for P to AP switching.

In both switching directions, the standard deviation of dE_{WRITE} is approximately equal to the average of dE_{WRITE}. For P to AP switching, the large variation in dE_{WRITE} is some of the p-MTJs had positive dE_{WRITE} whereas others had negative dE_{WRITE} with TID exposure.

However, in the AP to P direction, the large variation in dE_{WRITE} is due to fact that p-MTJs with diameters greater than 160 nm had much larger changes in E_{WRITE} with TID exposure than p-MTJs with diameters less than 100 nm, thus broadening the distribution curve in Fig. 6.6(f).

6.5 Summary

MTJs have been demonstrated to exhibit exceptional resilience to the effects of IR, which makes STT-MRAM a promising device candidate for rad-hard memory and computing applications. Previous experiments showed that MTJs had perfect state retention and maintained their TMR after being exposed to ionizing doses beyond 10 Mrad. Furthermore, MTJs that are irradiated at these high doses can still be reliably switched via field or STT switching.

The results presented in chapter not only confirm the rad-hard properties of MTJs since the MTJ resistance and TMR showed negligible change after 1 Mrad (Si) of TID exposure, but also these results provide a further in-depth analysis of the effects of TID exposure on key performance parameters for STT-MRAM. Subtle changes in the thermal stability factor, intrinsic critical switching voltage, and write energy were observed. These effects were larger as the diameter of the p-MTJ increased, which indicates that the changes induced by TID exposure would not limit device size scaling toward smaller sizes for higher density STT-MRAM cells. The distributions of these performance parameters show the changes do not always increase with increasing TID exposure. This means that changes in these parameters for the write operation in STT-MRAM may be somewhat significant initially, but do not necessarily increase at higher doses.

Additionally, some of the changes measured in Δ, V_{C0} and E_{WRITE} did not follow consistent trends between 300 krad and 1 Mrad trials. This indicates that there are multiple, competing influences on the parameters effected. Our results confirm that MTJs are very promising candidates for radiation tolerant memory applications, however, the subtle effects of TID demonstrated here should be accounted for in the design and fabrication of radiation tolerant devices with STT-MRAM.

Chapter 7: Dissertation conclusions

I presented my research on MTJ devices in new paradigms of novel computing and memory applications in this dissertation. In particular, I investigated the prospects of MTJs as probabilistic bits, tunable random number generators in stochastic computing, MRAM components with ultra-low switching current density, and memory cells in high radiation environments.

First, I presented a method of generating telegraphic switching signals with tunable stochasticity called ‘dual-biasing’. This method uses an external magnetic field and a DC bias voltage with certain specifications on their magnitudes and orientations. My results demonstrated that the signals generated using dual-biasing have two degrees of tunability, which means that the AP- and P-state dwell times can be tuned separately. This opens the door for three intriguing opportunities in probabilistic computing applications. One is that dual-biasing is a potential solution to overcome the effects of device-to-device variations through on-board corrections in large-scale p-bit networks. The second is neural signals in stochastic spiking networks can double its information capacity with dual-biasing since the two degrees of tunability feature means that the rate of spikes generated and the time-averaged output can be considered as independent variables. Lastly, telegraphic switching signals with bipolar representation are generated when applying the dual-biasing method to two MTJs connected in series. The two degrees of tunability feature ensures that, under certain bias fields, the average output becomes fixed at a 50% switching probability over a large range of bias voltages, thus making this method a promising solution for stochastic bit generators in stochastic computing schemes using bipolar encoding. Furthermore, I tested the dual-biasing method on 10 MTJs with varying thermal stabilities. My results

showed that the noise resilience and the maximum switching rate improved when dual-biasing on s-MTJs, however, the degree of separation in AP- and P-state tunability with bias voltage decreased, particularly at low bias fields. Nevertheless, the decrease in degree of separation between AP- and P-state tunability for s-MTJs was still large enough where two degrees of tunability was still achieved.

While the two degrees of tunability capability of dual-biasing is promising, this process still requires a tunable, local external magnetic field, which is detrimental for large-scale networks. In the following chapter, I examined the interplay between VCEC and SOT effects. I demonstrated that the SOT current reduced the coercivity and the IEC in p-MTJs with SAF free layers. Therefore, with the assistance of an SOT current, VCEC switching was achieved at current densities as low as 10^3 A/cm^2, which is nearly two orders of magnitude lower than the switching current densities for VCEC-only switching. Furthermore, I analyzed the influencing factors of SOT on the coercivity and the stray field and determined that Joule heating and the in-plane component of the SOT generated effective field were the primary factors. Most of my analysis was in the context of MRAM cells with ultra-low switching current density, however, I also presented results which indicate that SOT and VCEC can also be used as a dual-biasing method to generate tunable stochastic switching signals, thus eliminating the need for the local, tunable bias fields.

Next, I introduced SC-CRAM, which is a method of using the CRAM architecture to generate and compute stochastic bit-streams. In SC-CRAM, stochastic bit-streams are generated directly within the CRAM cells, meaning that the stochastic bit-stream generation and processing steps are embedded within the same circuit block. Furthermore, I showed that complex arithmetic functions can be performed in SC-CRAM with both the

circuit area and area-delay product reduced by nearly 100 times when compared to conventional CRAM. Since the information being processed in SC-CRAM is inherently stochastic, SC-CRAM is a promising method for applications involving approximations and detecting statistical anomalies on large data sets. For my analysis, I evaluated the performance metrics for SC-CRAM and conventional CRAM for three applications: local image thresholding, Bayesian inference for object location, and Bayesian belief network for heart disaster prediction. I demonstrated that SC-CRAM benefits from smaller subarray sizes and similar energy consumption compared to conventional CRAM.

Finally, I demonstrated that MTJs are promising candidates for on-board memory devices in high radiation applications, such as space exploration. I used synchronous switching probability measurements on 24 magnetic tunnel junctions to measure their thermal stabilities, critical switching voltages, and write energies, then repeated these measurements after exposure to 300 krad and 1 Mrad of γ-radiation. My results demonstrated that TID exposure caused these metrics to shift slightly. These shifts were dependent on pulse width, which suggests that TID exposure causes changes in the VCMA coefficient of the MTJs, which may be the result of dangling bonds at the MgO/CoFeB interface. However, my results also indicate that there are multiple competing influences, meaning that the changes in the three metrics examined will not continue to increase with long-term exposure. Furthermore, MTJs with larger diameters tended to be more affected by TID exposure meaning that future scaling towards smaller dimensions will not negatively influence the resilience against ionizing radiation.

Bibliography

[1] W. H. Butler, X.-G. Zhang, T. C. Shulthess, and J. M. MacLaren, *Spin-dependent tunneling conductance of Fe|MgO|Fe sandwiches*, Phys. Rev. B **63**, 054416 (2001).

[2] W. H. Butler, *Tunneling magnetoresistance from a symmetry filtering effect**, Sci. Technol. Adv. Mater. **9**, 014106 (2008).

[3] M. Julliere, *Tunneling between ferromagnetic films*, Phys. Lett. A **54**, 225 – 226 (1975).

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

[5] L. Lopez-Diaz, L. Torres, and E. Moro, *Transition from ferromagnetism to superparamagnetism on the nanosecond time scale*, Phys. Rev. B **65**, 224406 (2002).

[6] A. V. Khvalkovskiy, D. Apalkov, S. Watts, R. Chepulskii, R. S. Beach, A. Ong, X. Tang, A. Driskill-Smith, W. H. Butler, P. B. Visscher, D. Lottis, E. Chen, V. Nikitin, and M. Krounbi, *Basic principles of STT-MRAM cell operation in memory arrays*, J. Phys. D: Appl. Phys. **46**, 074001 (2013).

[7] J.-G. Zhu, *Magnetoresistive Random Access Memory: The Path to Competitiveness and Scalability*, Proc. IEEE **96**, 1786 – 1798 (2008).

[8] S. Bhatti, R. Sbiaa, A. Hirohata, H. Ohno, S. Fukami, and S. N. Piramanayagam, *Spintronics based random access memory: a review*, Materials Today **20**, 530 – 548 (2017).

[9] S. Amara, H. Bea, R. C. Sousa, and B. Dieny, *Barrier Breakdown Mechanisms in MgO-Based Magnetic Tunnel Junctions under Pulsed Conditions*, 2012 4th IEEE Int. Memory Workshop, pp. 1 – 4 (2012).

[10] L. Liu, C.-F. Pai, Y. Li, H. W. Tseng, D. C. Ralph, R. A. Buhrman, *Spin-torque switching with the giant spin Hall effect of tantalum*, Science **336**, 555 – 558 (2012).

[11] J. Sinova, S. O. Valenzuela, J. Wunderlich, C. H. Back, and T. Jungwirth, *Spin Hall effects*, Rev. Mod. Phys. **87**, 1213 – 1259 (2015).

[12] S. Fukami, T. Anekawa, C. Zhang, and H. Ohno, *A spin-orbit torque switching scheme with collinear magnetic easy axis and current configuration*, Nat. Nanotechnol. **11**, 621 – 625 (2016).

[13] R. Ramaswamy, J. M. Lee, K. Cai, and H. Yang, *Recent advances in spin-orbit torques: Moving towards device applications*, Appl. Phys. Rev. **5**, 031107 (2018).

[14] A. Manchon, J. Zelezny, I. M. Miron, T. Jungwirth, J. Sinova, A. Thiaville, K. Garello, and P. Gambardella, *Current-induced spin-orbit torques in ferromagnetic and antiferromagnetic systems*, Rev. Mod. Phys. **91**, 035004 (2019).

[15] J. Ryu, S. Lee, K.-J. Lee, and B.-G. Park, *Current-Induced Spin-Orbit Torques for Spintronic Applications*, Adv. Mater. **32**, 1907148 (2020).

[16] M. Wang, W. Cai, K. Cao, J. Zhou, J. Wrona, S. Peng, H. Yang, J. Wei, W. Kang, Y. Zhang, J. Langer, B. Ocker, A. Fert, and W. Zhao, *Current-induced magnetization switching in atom-thick tungsten engineered perpendicular magnetic tunnel junctions with large tunnel magnetoresistance*, Nat. Commun. **9**, 671 (2018).

[17] S. Shi, Y. Ou, S. V. Aradhya, D. C. Ralph, and R. A. Buhrman, *Fast Low-Current Spin-Orbit-Torque Switching of Magnetic Tunnel Junctions through Atomic Modifications of the Free-Layer Interfaces*, Phys. Rev. Appl. **9**, 011002 (2018).

[18] D. Zhang, M. Bapna, W. Jiang, D. Sousa, Y.-C. Liao, Z. Zhao, Y. Lv, P. Sahu, D. Lyu, A. Naeemi, T. Low, S. A. Majetich, and J.-P. Wang, *Bipolar Electric-Field Switching of Perpendicular Magnetic Tunnel Junctions through Voltage-Controlled Exchange Coupling*, Nano Lett. **22**, 622 – 629 (2022).

[19] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nat. Review **521**, 436 – 444 (2015).

[20] J. Grollier, D. Querlioz, and M. D. Stiles, *Spintronic Nanodevices for Bioinspired Computing*, Proc. IEEE **104**, 2024 – 2039 (2016).

[21] R. B. Stein, E. R. Gossen, and K. E. Jones, *Neuronal variability: Noise or part of the signal?*, Nat. Rev. Neurosci. **6**, 389 – 397 (2005).

[22] E. M. Izhikevich, *Which model to use for cortical spiking neurons?*, IEEE Trans. Neural Netw. **15**, 1063 – 1070 (2004).

[23] D. Bhowmik and M. Shanahan, *How well do oscillator models capture the behavior of biological neurons?*, 2012 Int. Joint Conf. Neural Netw. (IJCNN) pp. 1 – 8 (2012).

[24] D. Vodenicarevic, N. Locatelli, F. A. Araujo, J. Grollier, and D. Querlioz, *A Nanotechnology-Ready Computing Scheme based on a Weakly Coupled Oscillator Network*, Sci. Rep. **7**, 44772 (2017).

[25] G. Csaba, A. Raychowdhury, S. Datta, and W. Porod, *Computing with Coupled Oscillators: Theory, Devices, and Applications*, 2018 IEEE Int. Symp. Circuits and Syst. pp. 1 – 5 (2018).

[26] A. Sengupta, M. Parsa, B. Han, and K. Roy, *Probabilistic Deep Spiking Neural Network Systems Enabled by Magnetic Tunnel Junction*, IEEE Trans. Electron Devices **63**, 2963 – 2970 (2016).

[27] G. Srinivasan, A. Sengupta, and K. Roy, *Magnetic Tunnel Junction Based Long-Term Short-Term Stochastic Synapse for a Spiking Neural Network with On-Chip STDP Learning*, Sci. Rep. **6**, 29545 (2016).

[28] Z. Chowdhury, J. D. Harms, S. K. Khatamifard, M. Zabihi, Y. Lv, A. P. Lyle, S. S. Sapatnekar, U. R. Karpuzcu, and J.-P. Wang, *Efficient In-Memory Processing Using Spintronics*, IEEE Comp. Architect. Lett. **17**, 42 – 46 (2018).

[29] M. Zabihi, Z. I. Chowdhury, Z. Zhao, U. R. Karpuzcu, J.-P. Wang, and S. S. Sapatnekar, *In-Memory Processing on the Spintronic CRAM: From Hardware Design to Application Mapping*, IEEE Trans. Computers **68**, 1159 – 1173 (2019).

[30] M. Zabihi, Z. Zhao, M. DC, Z. I. Chowdhury, S. Resch, T. Peterson, U. R. Karpuzcu, J.-P. Wang, and S. S. Sapatnekar, *Using Spin-Hall MTJs to Build an Energy-Efficient In-memory Computation Platform*, 20th Int. Symp. Quality Electronic Design (ISQED) pp. 52 – 57 (2019).

[31] S. Resch, S. K. Khatamifard, Z. I. Chowdhury, M. Zabihi, Z. Zhao, J.-P. Wang, S. S. Sapatnekar, and U. R. Karpuzcu, *PIMBALL: Binary Neural Networks in Spintronic Memory*, ACM Trans. Architict. Code Optim., **16**, 41 (2019).

[32] H. Cilasun, S. Resch, Z. I. Chowdhury, E. Olson, M. Zabihi, Z. Zhao, T. Peterson, K. K. Parhi, J.-P. Wang, S. S. Sapatnekar, and U. R. Karpuzcu, *Spiking Neural Networks in Spintronic Computational RAM*, ACM Trans. Architict. Code Optim., **18**, 59 (2021).

[33] J. Jeddeloh and B. Keeth, *Hybrid memory cube new DRAM architecture increases density and performance*, Proc. IEEE Int. Symp. VLSI Technol., pp. 87 – 88 (2012).

[34] B. R. Gaines, *Stochastic Computing*, Proc. AFIPS Spring Joint Comput. Conf., pp. 149 – 156 (1967).

[35] W. J. Poppelbaum, C Afuso, and J. W. Esch, *Stochastic computing elements and systems*, Proc. AFIPS Spring Joint Comput. Conf., pp. 635 – 644 (1967).

[36] B. R. Gaines, *Stochastic computing systems*, Advances in Information Systems Science **2**, 37 – 172 (1969).

[37] A. Alaghi and J. P. Hayes, *Survey of Stochastic Computing*, ACM Trans. Embedded Comput. Syst., **12**, 92 (2013).

[38] W. Qian, X. Li, M. D. Riedel, K. Bazargan, and D. J. Lilja, *An Architecture for Fault-Tolerant Computation with Stochastic Logic*, IEEE Trans. Comput., **60**, 93 – 105 (2011).

[39] P. Li and D. J. Lilja, *A low power fault-tolerant architecture for the kernel density estimation based image segmentation algorithm*, Proc. IEEE Int. Conf. Appl.-Specific Syst. Archit. Process, pp. 161 – 168 (2011).

[40] J. M. de Aguiar and S. P. Khatri, *Exploring the viability of stochastic computing*, 2015 33rd IEEE Int. Conf. Comput. Design (ICCD), pp. 391 – 394 (2015).

[41] M. H. Najafi and M. E. Salehi, *A Fast Fault-Tolerant Architecture for Sauvola Local Image Thresholding Algorithm Using Stochastic Computing*, IEEE Trans. VLSI Syst., **24**, 808 – 812 (2016).

[42] A. Alaghi, W. Qian, and J. P. Hayes, *The Promise and Challenge of Stochastic Computing*, IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems, **37**, 1515 – 1531 (2018).

[43] P. Li, W. Qian, D. J. Lilja, K. Bazargan, and M. D. Riedel, *Case Studies of Logical Computation on Stochastic Bit Streams*, Integrated Circuit and Systems Design. Power and Timing Modeling, Optimization and Simulation (PATMOS). Lecture Notes in Computer Science, **7606**, 235 – 244 (2013).

[44] P. Li, D. J. Lilja, W. Qian, K. Bazargan, and M. D. Riedel, *Computation on Stochastic Bit Streams Digital Image Processing Case Studies*, IEEE Trans. VLSI Syst., **22**, 449 – 462 (2014).

[45] A. Alaghi, C. Li, and J. P. Hayes, *Stochastic circuits for real-time image-processing applications*, 2013 50th ACM/EDAC/IEEE Design Automat. Conf. (DAC), pp. 1 – 6 (2013).

[46] W. H. Choi, Y. Lv, J. Kim, A. Deshpande, G Kang, J.-P. Wang, and C. H. Kim, *A Magnetic Tunnel Junction Based True Random Number Generator With Conditional Perturb and Real-Time Output Probability Tracking*, IEEE Int. Electron Devices Meeting (IEDM), pp. 12.5.1 – 12.5.4 (2014).

[47] 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*, 13th IEEE Int. NEW Circuits and Syst. Conf. (NEWCAS), pp. 3 – 6 (2015).

[48] A. Mondal and A. Srivastava, *Energy-efficient Design of MTJ-based Neural Networks with Stochastic Computing*, ACM J. Emerging Technol. Computing Syst., **16**, 7 (2019).

[49] Y. Jiang, Y. Lv, M. Jamali, and J.-P. Wang, *Spin Analog-to-Digital Converter Using Magnetic Tunnel Junction and Spin Hall Effect*, IEEE Electron Device Lett., **36**, 511 – 513 (2015).

[50] W. H. Choi, Y. Lv, H. Kim, J.-P. Wang, and C. H. Kim, *An 8-bit Analog-to-Digital Converter based on the voltage-dependent switching probability of a Magnetic Tunnel Junction*, 2015 Symp. VLSI Technol., pp. T162 – T163 (2015).

[51] L. A. de Barros, H. Cai, Y. Wang, W. Zhao, and A. Ben Dhia, *Stochastic computation with spin torque transfer magnetic tunnel junction*, 2015 IEEE 13th Int. New Circuits and Syst. Conf. (NEWCAS), pp. 1 – 4 (2015).

[52] N. Onizawa, D. Katagiri, W. J. Gross, and T. Hanyu, *Analog-to-stochastic converter using magnetic tunnel junction devices for vision chips*, IEEE Trans. Nanotechnol., **15**, 705 – 714 (2016).

[53] X. Jia, J. Yang, Z. Wang, Y. Chen, H. Li, and W. Zhao, *Spintronics based stochastic computing for efficient Bayesian inference system*, 2018 23rd Asia and South Pacific Design Automat. Conf. (ASP-DAC), pp. 580 – 585 (2018).

[54] Y. Lv and J.-P. Wang, *A single magnetic-tunnel-junction stochastic computing unit*, IEEE Int. Electron Devices Meeting (IEDM), pp. 36.2.1 – 36.2.4 (2017).

[55] D. Zhang, L. Zeng, F. Gong, T. Gao, S. Gao, Y. Zhang, and W. Zhao, *Realization of Neural Coding by Stochastic Switching of Magnetic Tunnel Junction*, 2015 15th Non-Volatile Memory Technology Symp. (NVMTS), pp. 1 – 4 (2015).

[56] D. Zhang, L. Zeng, Y. Zhang, W. Zhao, and J. O. Klein, *Stochastic spintronic device based synapses and spiking neurons for neuromorphic computation*, 2016 IEEE/ACM Int. Symp. Nanoscale Archit. (NANOARCH), pp. 173 – 178 (2016).

[57] C. M. Liyanagedera, A. Sengupta, A. Jaiswal, and K. Roy, *Stochastic Spiking Networks Enabled by Magnetic Tunnel Junctions: From Nontelegraphic to Telegraphic Switching Regimes*, Phys. Rev. Appl. **8**, 064017 (2017).

[58] P. Krzysteczko, J. Müchenberger, M. Schäfers, G. Reiss, and A. Thomas, *The Memristive Magnetic Tunnel Junction as a Nanoscopic Synapse-Neuron System*, Adv. Mater. **24**, 762 – 766 (2012).

[59] D. I. Suh, G. Y. Bae, H. S. Oh, and W. Park, *Neural coding using telegraphic switching of magnetic tunnel junction*, J. Appl. Phys. **117**, 17D714 (2015).

[60] A. Sengupta, P. Panda, P. Wijesinghe, Y. Kim, and K. Roy, *Magnetic Tunnel Junction Mimics Stochastic Cortical Spiking Neurons*, Sci. Rep. **6**, 30039 (2016).

[61] G. Srinivasan, A. Sengupta, and K. Roy, *Magnetic tunnel junction enabled all-spin stochastic spiking neural network*, Design, Automation & Test in Europe Conf Exhibition (DATE), pp. 530 – 535 (2017).

[62] A. Sengupta, G. Srinivasan, D. Roy, and K. Roy, *Stochastic Inference and Learning Enabled by Magnetic Tunnel Junctions*, 2018 IEEE Int. Electron Devices Meeting (IEDM), pp. 15.6.1 – 15.6.4 (2018).

[63] K. Yang and A. Sengupta, *Stochastic magnetoelectric neuron for temporal information encoding*, Appl. Phys. Lett., **116**, 043701 (2020).

[64] K. Y. Camsari, R. Faria, B. M. Sutton, and S. Datta, *Stochastic p-Bits for Invertible Logic*, Phys. Rev. X, **7**, 031014 (2017).

[65] B. Behin-Aein, V. Diep, and S. Datta, *A building block for hardware belief networks*, Sci. Rep., **6**, 29893 (2016).

[66] B. Sutton, K. Y. Camsari, B. Behin-Aein, and S. Datta, *Intrinsic optimization using stochastic nanomagnets*, Sci. Rep., **7**, 44370 (2017).

[67] K. Y. Camsari, B. M. Sutton, and S. Datta, *p-bits for probabilistic spin logic*, Appl. Phys. Rev., **6**, 011305 (2019).

[68] R. Faria, K. Y. Camsari, and S. Datta, *Implementing Bayesian networks with embedded stochastic MRAM*, AIP Advances, **8**, 045101 (2018).

[69] R. Zand, K. Y. Camsari, S. D. Pyle, I. Ahmed, C. H. Kim, and R. F. DeMara, *Low-Energy Deep Belief Networks Using Intrinsic Sigmoidal Spintronic-based Probabilistic Neurons*, Proc. 2018 Great Lakes Symp. VLSI (GLSVLSI), pp. 15 – 20 (2018).

[70] O. Hassan, R. Faria, K. Y. Camsari, J. Z. Sun, and S. Datta, *Low-Barrier Magnet Design for Efficient Hardware Binary Stochastic Neurons*, IEEE Magn. Lett., **10**, 4502805 (2019).

[71] V. Ostwal and J. Appenzeller, *Spin-Orbit Torque-Controlled Magnetic Tunnel Junction With Low Thermal Stability for Tunable Random Number Generation*, IEEE Magn. Lett., **10**, 4503305 (2019).

[72] P. Debashis, R. Faria, K. Y. Camsari, J. Appenzeller, S. Datta, and Z. Chen, *Experimental demonstration of nanomagnet networks as hardware for Ising computing*, 2016 IEEE Int. Electron Devices Meeting (IEDM), pp. 34.3.1 – 34.3.4 (2016).

[73] R. Faria, K. Y. Camsari, and S. Datta, *Low-Barrier Nanomagnets as p-Bits for Spin Logic*, IEEE Magn. Lett., **8**, 4105305 (2017).

[74] K. Y. Camsari, S. Salahuddin, and S. Datta, *Implementing p-bits With Embedded MTJ*, IEEE Electon Device Lett., **38**, 1767 – 1770 (2017).

[75] P. Debashis, R. Faria, K. Y. Camsari, and Z. Chen, *Design of Stochastic Nanomagnets for Probabilistic Spin Logic*, IEEE Magn. Lett., **9**, 4305205 (2018).

[76] V. Ostwal, P. Debashis, R. Faria, Z. Chen, and J. Appenzeller, *Spin-torque devices with hard axis initialization as Stochastic Binary Neurons*, Sci. Rep., **8**, 16689 (2018).

[77] B. R. Zink, Y. Lv, and J.-P. Wang, *Telegraphic switching signals by magnetic tunnel junctions for neural spiking signals with high information capacity*, J. Appl. Phys., **124**, 152121 (2018).

[78] B. R. Zink, Y. Lv, and J.-P. Wang, *Independent Control of Antiparallel- and Parallel-State Thermal Stability Factors in Magnetic Tunnel Junctions for Telegraphic Signals With Two Degrees of Tunability*, IEEE Trans. Electron Devices, **66**, 5353 – 5359 (2019).

[79] B. R. Zink and J.-P. Wang, *Influence of Intrinsic Thermal Stability on Switching Rate and Tunability of Dual-Biased Magnetic Tunnel Junctions for Probabilistic Bits*, IEEE Magn. Lett., **12**, 4501405 (2021).

[80] Y. Lv, R. P. Bloom, and J.-P. Wang, *Experimental Demonstration of Probabilistic Spin Logic by Magnetic Tunnel Junctions*, IEEE Magn. Lett., **10**, 4510905 (2019).

[81] R. Heindl, W. H. Rippard, S. E. Russek, M. R. Pufall, and A. B. Kos, *Validity of the thermal activation model for spin-transfer torque switching in magnetic tunnel junctions*, J. Appl. Phys., **109**, 073910 (2011).

[82] H. Zhao, Y. Zhang, P. K. Amiri, J. A. Katine, J. Langer, H. Jiang, I. N. Krivorotov, K. L. Wang, and J.-P. Wang, *Spin-Torque Driven Switching Probability Density Function Asymmetry*, IEEE Trans. Magn., **48**, 3818 – 3820 (2012).

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

[84] T. Aoki, Y. Ando, M. Oogane, and H. Naganuma, *Reproducible trajectory on subnanosecond spin-torque magnetization switching under a zero-bias field for MgO-based ferromagnetic tunnel junctions*, Appl. Phys. Lett., **96**, 142502 (2010).

[85] H. Zhao, A. Lyle, Y. Zhang, P. K. Amiri, G. Rowlands, Z. Zeng, J. Katine, H. Jiang, K. Galatsis, K. L. Wang, I. N. Krivorotov, and J.-P. Wang, *Low writing energy and sub nanosecond spin torque transfer switching of in-plane magnetic tunnel junction for spin torque transfer random access memory*, J. Appl. Phys., **109**, 07C720 (2011).

[86] M. R. Pufall, W. H. Rippard, S. Kaka, S. E. Russek, and T. J. Silva, *Large-angle, gigahertz-rate random telegraphic switching induced by spin-momentum transfer*, Phys. Rev. B, **69**, 214409 (2004).

[87] G. D. Fuchs, J. A. Katine, S. I. Kiselev, D. Mauri, K. S. Wooley, D. C. Ralph, and R. A. Buhrman, *Spin Torque, Tunnel-Current Spin Polarization, and Magnetoresistance in MgO Magnetic Tunnel Junctions*, Phys. Rev. Lett., **95**, 186603 (2006).

[88] W. Rippard, R. Heindl, M. Pufall, S. Russek, and A. Kos, *Thermal relaxation rates of magnetic nanoparticles in the presence of magnetic fields and spin-transfer effects*, Phys. Rev. B, **84**, 064439 (2011).

[89] S. K. Piotrowski, M. Bapna, S. D. Oberdick, and S. A. Majetich, *Size and voltage dependence of effective anisotropy in sub-100-nm perpendicular magnetic tunnel junctions*, Phys. Rev. B, **94**, 014404 (2016).

[90] M. Bapna and S. A. Majetich, *Current control of time-averaged magnetization in superparamagnetic tunnel junctions*, Appl. Phys. Lett., **111**, 243107 (2017).

[91] B. Parks, A. Abdelgawad, T. Wong, R. F. L. Evans, and S. A. Majetich, *Magnetoresistance Dynamics in Superparamagnetic Co-Fe-B Nanodots*, Phys. Rev. Appl., **13**, 014063 (2020).

[92] M. W. Daniels, A. Madhavan, P. Talatchian, A. Mizrahi, and M. D. Stiles, *Energy-Efficient Stochastic Computing with Superparamagnetic Tunnel Junctions*, Phys. Rev. Appl., **13**, 034016 (2020).

[93] M. A. Abeed and S. Bandyopadhyay, *Sensitivity of the Power Spectra of Thermal Magnetization Fluctuations in Low Barrier Nanomagnets Proposed for Stochastic Computing to In-Plane Barrier Height Variations and Structural Defects*, Spin, **10**, 2050001 (2020).

[94] Y. Lv, B. R. Zink, and J.-P. Wang, *Bipolar Random Spike and Bipolar Random Number Generation by Two Magnetic Tunnel Junctions*, IEEE Trans. Electron Devices, **69**, 1582 – 1587 (2022).

[95] X. Yang, J. Lin, W. Zheng, J. Zhao, M. Ji, Y. Lei, and Z. Chai, *Research on learning mechanism designing for equilibrated bipolar spiking neural networks*, Artif. Intell. Rev., **53**, 5189 – 5215 (2020).

[96] M. A. Mansor and S. Sathasivam, *Activation function comparison in neural-symbolic integration*, AIP Conf. Proc., **1750**, 020013 (2016).

[97] L. E. Bassham, A. L. Rukhin, J. Soto, J. R. Nechvatal, M. E. Smid, S. D. Leigh, M. Levenson, M. Vangel, N. A. Heckert, and D. L. Banks, *A statistical test suite for random and pseudo-random number generators for cryptographic applications*, NIST Special Publication, NIST, Gaithersburg, MD, USA, Tech. Rep. 800-22 Rev 1a (2010) [Online]. Available: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=906762.

[98] C. Tannous and J. Gieraltowski, *The Stoner-Wohlfarth model of ferromagnetism*, Eur. J. Phys., **29**, 475 – 487 (2008).

[99] Z. Li and S. Zhang, *Thermally assisted magnetization reversal in the presence of a spin-transfer torque*, Phys. Rev. B, **69**, 134416 (2004).

[100] Y. Shiota, F. Bonell, S. Miwa, N. Mizuochi, T. Shinjo, and Y. Suzuki, *Opposite signs of voltage-induced perpendicular magnetic anisotropy change in CoFeB|MgO junctions with different underlayers*, Appl. Phys. Lett., **103**, 082410 (2013).

[101] D. Apalkov, B. Dieny, and J. M. Slaughter, *Magnetoresistive Random Access Memory*, Proc. IEEE, **104**, 1796 – 1830 (2016).

[102] B. Dieny, I. L. Prejbeanu, K. Garello, P. Gambardella, P. Freitas, R. Lehndorff, W. Raberg, U. Ebels, S. O. Demokritov, J. Akerman, A. Deac, P. Pirro, C. Abelmann, A. Anane, A. V. Chumak, A. Hiroata, S. Mangin, M. Cengiz Onbaşli, M. d'Aquino, G. Prenat, G. Finocchio, L. Lopez Diaz, R. Chantrell, O. Chubykalo-Fesenko, and P. Bortolotti, *Opportunities and challenges for spintronics in the microelectronic industry*, Nat. Electronics, **3**, 446 – 459 (2016).

[103] S. Ikegawa, F. B. Mancoff, J. Janesky, and S. Aggarwal, *Magnetoresistive Random Access Memory: Present and Future*, IEEE Trans. Electron Devices, **67**, 1407 – 1419 (2020).

[104] S. Ikeda, J. Hayakawa, Y. Min Lee, F. Matsukura, Y. Ohno, T. Hanyu, and H. Ohno, *Magnetic Tunnel Junctions for Spintronic Memories and Beyond*, IEEE Trans. Electron Devices, **54**, 991 – 1002 (2007).

[105] T. Endoh and H. Honjo, *A Recent Progress of Spintronics Devices for Integrated Circuit Applications*, J. Low Power Electron. Appl., **8**, 44 (2018).

[106] Y.-C. Liao, C. Pan, and A. Naeemi, *Benchmarking and Optimization of Spintronic Memory Arrays*, IEEE J. Explor. Solid-State Computat. Devices and Circuits, **6**, 9 – 17 (2020).

[107] K. Garello, C. Onur Avci, I. Mihai Miron, M. Baumgartner, A. Ghosh, S. Auffret, O. Boulle, G. Gaudin, and P. Gambardella, *Ultrafast magnetization switching by spin-orbit torques*, Appl. Phys. Lett., **105**, 212402 (2014).

[108] M. Baumgartner, K. Garello, J. Mendil, C. Onur Avci, E. Grimaldi, C. Murer, J. Feng, M. Gabureac, C. Stamm, Y. Acremann, S. Finizio, S. Wintz, J. Raabe, and

P. Gambardella, *Spatially and time-resolved magnetization dynamics driven by spin-orbit torques*, Nat. Nanotech., **12**, 980 – 986 (2017).

[109] M. Cubukcu, O. Boulle, N. Mikuszeit, C. Hamelin, T. Brächer, N. Lamard, M.-C. Cyrille, L. Buda-Prejbeanu, K. Garello, I. Mihai Miron, O. Klein, G. de Loubens, V. V. Naletov, J. Langer, B. Ocker, P. Gambardella, and G. Gaudin, *Ultra-Fast Perpendicular Spin-Orbit Torque MRAM*, IEEE Trans. Magn., **54**, 9300204 (2018).

[110] K. Jhuria, J. Hohfeld, A. Pattabi, E. Martin, A. Ygnacio Arriola Córdova, X. Shi, R. Lo Conte, S. Petit-Watelot, J. C. Rojas-Sanchez, G. Malinowski, S. Mangin, A. Lemaître, M. Hehn, J. Bokor, R. B. Wilson, and J. Gorchon, *Spin-orbit torque switching of a ferromagnet with picosecond electrical pulses*, Nat. Electronics., **3**, 680 – 686 (2020).

[111] G. Yu, P. Upadhyaya, Y. Fan, J. G. Alzate, W. Jiang, K. L. Wong, S. Takei, S. A. Bender, L.-T. Chang, Y. Jiang, M. Lang, J. Tang, Y. Wang, Y. Tserkovnyak, P. Khalili Amiri, and K. L. Wang, *Switching of perpendicular magnetization by spin-orbit torques in the absence of external magnetic fields*, Nat. Nanotech., **9**, 548 – 554 (2014).

[112] S. Fukami, C. Zhang, S. DuttaGupta, A. Kurenkov, and H. Ohno, *Magnetization switching by spin-orbit torque in an antiferromagnet-ferromagnet bilayer system*, Nat. Mater., **15**, 535 – 542 (2016).

[113] Q. Ma, Y. Li, D. B. Gopman, Y. P. Kabanov, R. D. Shull, and C. L. Chien, *Switching a Perpendicular Ferromagnetic Layer by Competing Spin Currents*, Phys. Rev. Lett., **120**, 117703 (2018).

[114] N. Sato, F. Xue, R. M. White, C. Bi, and S. X. Wang, *Two-terminal spin-orbit torque magnetoresistive random access memory*, Nat. Electronics, **1**, 508 – 511 (2018).

[115] C.-Y. You and S. D. Bader, *Prediction of switching/rotation of the magnetization direction with applied voltage in a controllable interlayer exchange coupled system*, J. Magn. Magn. Mater., **195**, 488 – 500 (1999).

[116] S. Yakata, H. Kubota, T. Sugano, T. Seki, K. Yakushiji, A. Fukushima, S. Yuasa, and K. Ando, *Thermal stability and spin-transfer switchings in MgO-based magnetic tunnel junctions with ferromagnetically and antiferromagnetically coupled synthetic free layers*, Appl. Phys. Lett., **95**, 242504 (2009).

[117] M. Wang, W. Cai, D. Zhu, Z. Wang, J. Kan, Z. Zhao, K. Cao, Z. Wang, Y. Zhang, T. Zhang, C. Park, J.-P. Wang, A. Fert, and W. Zhao, *Field-free switching of a perpendicular magnetic tunnel junction through the interplay of spin-orbit and spin-transfer torques*, Nat. Electronics, **1**, 582 – 588 (2018).

[118] B. Zeinali, J. K. Madsen, P. Raghavan, and F. Moradi, *Ultra-Fast SOT-MRAM Cell with STT Current for Deterministic Switching*, 2017 IEEE Int. Conf. Computer Design (ICCD), pp. 463 – 468 (2017).

[119] E. Grimaldi, V. Krizakova, G. Sala, F. Yasin, S. Couet, G. Sankar Kar, K. Garello, and P. Gambardella, *Single-shot dynamics of spin-orbit torque and spin transfer torque switching in three-terminal magnetic tunnel junctions*, Nat. Nanotech., **15**, 111 – 117 (2020).

[120] A. van den Brink, S. Cosemans, S. Cornelissen, M. Manfrini, A. Vaysset, W. Van Roy, T. Min, H. J. M. Swagten, and B. Koopmans, *Spin-Hall-assisted magnetic random access memory*, Appl. Phys. Lett., **104**, 012403 (2014).

[121] S. Pathak, C. Youm, and J. Hong, *Impact of Spin-Orbit Torque on Spin-Transfer Torque Switching in Magnetic Tunnel Junctions*, Sci. Rep., **10**, 2799 (2020).

[122] M. DC, R. Grassi, J.-Y. Chen, M. Jamali, D. Reifsnyder Hickey, D. Zhang, Z. Zhao, H. Li, P. Quarterman, Y. Lv, M. Li, A. Manchon, K. A. Mkhoyan, T. Low, and J.-P. Wang, *Room-temperature high spin-orbit torque due to quantum confinement in sputtered $Bi_xS_{(1-x)}$ films*, Nat. Mater., **17**, 800 – 807 (2018).

[123] B. R. Zink, D. Zhang, H. Li, O. J. Benally, Y. Lv, D. Lyu, and J.-P. Wang, *Ultralow Current Switching of Synthetic-Antiferromagnetic Magnetic Tunnel Junctions Via Electric-Field Assisted by Spin-Orbit Torque*, Adv. Electron. Mater., 2200382 (2022).

[124] A. Rajanikanth, T. Hauet, F. Montaigne, S. Mangin, and S. Andrieu, *Magnetic anisotropy modified by electric field in V/Fe/MgO(001)/Fe epitaxial magnetic tunnel junction*, Appl. Phys. Lett., **103**, 062402 (2013).

[125] T. Nozaki, A. Koziol-Rachwal, W. Skowroński, V. Zayets, Y. Shiota, S. Tamaru, H. Kubota, A. Fukushima, S. Yuasa, and Y. Suzuki, *Large Voltage-Induced Changes in the Perpendicular Magnetic Anisotropy of an MgO-Based Tunnel Junction with an Ultrathin Fe Layer*, Phys. Rev. Appl., **5**, 044006 (2016).

[126] G. Mihajlović, N. Smith, T. Santos, J. Li, M. Tran, M. Carey, B. D. Terris, and J. A. Katine, *Origin of the Resistance-Area-Product Dependence of Spin-Transfer-Torque Switching in Perpendicular Magnetic Random-Access Memory Cells*, Phys. Rev. Appl., **13**, 024004 (2020).

[127] V. Krizakova, E. Grimaldi, K. Garello, G. Sala, S. Couet, G. Sankar Kar, and P. Gambardella, *Interplay of Voltage Control of Magnetic Anisotropy, Spin-Transfer Torque, and Heat in the Spin-Orbit-Torque Switching of Three-Terminal Magnetic Tunnel Junctions*, Phys. Rev. Appl., **15**, 054055 (2021).

[128] Z. Zhao, A. K. Smith, M. Jamali, and J.-P. Wang, *External-Field-Free Spin Hall Switching of Perpendicular Magnetic Nanopillar with a Dipole-Coupled Composite Structure*, Adv. Electron. Mater., **6**, 1901368 (2020).

[129] M. P. Sharrock, *Time-dependent magnetic phenomena and particle-size effects in recording media*, IEEE Trans. Magn., **26**, 193 – 197 (1990).

[130] L. Liu, O. J. Lee, T. J. Gudmundsen, D. C. Ralph, and R. A. Buhrman, *Current-Induced Switching of Perpendicularly Magnetized Magnetic Layers Using Spin Torque from the Spin Hall Effect*, Phys. Rev. Lett., **109**, 096602 (2012).

[131] Y. Li, X. Jin, P. Pan, F. Nan Tan, W. Siang Lew, and F. Ma, *Temperature-dependent interlayer exchange coupling strength in synthetic antiferromagnetic $[Pt/Co]_2/Ru/[Co/Pt]_4$ multilayers*, Chin. Phys. B, **27**, 127502 (2018).

[132] Y. Henry and K. Ounadjela, *Evidence for a Direct d-d Hybridization Mechanism for the Interlayer Exchange Coupling in Epitaxial Co/Mn Multilayers*, Phys. Rev. Lett., **76**, 1944 – 1947 (1996).

[133] C. D. Shunman, T. E. Potok, R. M. Patton, J. D. Birdwell. M. E. Dean, G. S. Rose, and J. S. Plank, *A Survey of Neuromorphic Computing and Neural Networks in Hardware*, ArXiv 1705.06963 (2017).

[134] C. D. Shunman, S. R. Kulkarni, M. Parsa, J. P. Mitchell, P. Date, and B. Kay, *Opportunities for neuromorphic computing algorithms and applications*, Nat. Comput. Sci. **2**, 10 – 19 (2022).

[135] G. Indiveri and S. C. Liu, *Memory and information processing in neuromorphic systems*, Proc. IEEE **103**, 1379 – 1397 (2015).

[136] D. Querlioz, O. Bichler, A. F. Vincent, and C. Gamrat, *Bioinspired programming of memory devices for implementing an inference engine*, Proc. IEEE **103**, 1398 – 1416 (2015).

[137] Y. Liu and K. K. Parhi, *Computing hyperbolic tangent and sigmoid using stochastic logic*, 50th Aislomar Conf. Signals, Syst., Computers, pp. 1580 – 1585 (2016).

[138] H. Hsiao, J. Anderson, and Y. Hara-Azumi, *Generating Stochastic Bitstreams*, in Stochastic Computing: Techniques and Applications, Springer, pp. 137 – 152 (2019).

[139] M. Zabihi, A. K. Sharma, M. G. Mankalale, Z. I. Chowdhury, Z. Zhao, S. Resch, U. R. Karpuzcu, J.-P. Wang, and S. S. Sapatnekar, *Analyzing the Effects of Interconnect Parasitics in the STT CRAM In-Memory Computational Platform*, IEEE J. Explor. Solid-State Comput. Devices and Circuits, **6**, 71 – 79 (2020).

[140] A. Keys, J. Adams, J. Cressler, M. Johnson, and M. Patrick, *A review of NASA's radiation-hardened electronics for space environments project*, in Proc. AIAA SPACE Conf. Expo., pp. 576 – 582 (2008).

[141] T. R. Oldham and F. B. Mclean, *Total ionizing dose effects in MOS oxides and devices*, IEEE Trans. Nucl. Sci., **50**, 483 – 499 (2003).

[142] G. Cellere and A. Paccagnella, *A review of ionizing radiation effects in floating gate memories*, IEEE Trans. Device Mater. Rel., **4**, 359 – 370 (2004).

[143] R. H. Maurer, M. E. Fraeman, M. N. Martin, and D. R. Roth, *Harsh Environments: Space radiation environment, effects, and mitigation*, Johns Hopkins APL Tech. Dig., **28**, 17 – 28 (2008).

[144] F.-X. Yu, J.-R. Liu, Z.-L. Huang, H. Luo, and Z.-M. Lu, *Overview of radiation hardening techniques for IC design*, Inf. Technol. J., **9**, 1068 – 1080 (2010).

[145] F. Ren, A. Jander, P. Dhagat, and C. Nordman, *Radiation tolerance of magnetic tunnel junctions with MgO tunnel barriers*, IEEE Trans. Nucl. Sci., **59**, 3034 – 3038 (2012).

[146] H. Hughes, K. Bussmann, P. J. McMarr, S.-F. Cheng, R. Shull, A. P. Chen, S. Schafer, T. Mewes, A. Ong, E. Chen, M. H. Mendenhall, and R. A. Reed, *Radiation studies of spin-transfer torque materials and devices*, IEEE Trans. Nucl. Sci., **59**, 3027 – 3033 (2012).

[147] N. Anuniwat, *Radiation effects on magnet tunnel junctions and novel magnetic thin films*, PhD Dissertation, Dept. Phys., Univ. Virginia (2015).

[148] B. Wang, Z. Wang, K. Cao, X. Bi, Y. Zhao, Y. Zhang, and W. Zhao, *Effects of gamma irradiation on magnetic properties of double-interface CoFeB/MgO multifilms*, IEEE Trans. Nucl. Sci., **66**, 77 – 81 (2019).

[149] M. Bagatin, S. Geradin, A. Paccagnella, S. Beltrami, A. Costantino, M. Muschitiello, A. Zadeh, and V. Ferlet-Cavrois, *Total ionizing dose effects in 3-D NAND flash memories*, IEEE Trans. Nucl. Sci., **66**, 48 – 53 (2019).

[150] R. Fang, Y. G. Velo, W. Chen, K. E. Holbert, M. N. Kozicki, H. Barnaby, and S. Yu, *Total ionizing dose effect of γ-ray radiation on the switching characteristics and filament stability of HfOx resistive random access memory*, Appl. Phys Lett., **104**, 183507 (2014).

[151] P. C. Adell, S. Moro, L. Gouyet, C. Chatry, and B. Vermeire, *Single event effect assessment of a 1-Mbit commercial magneto-resistive random access memory (MRAM)*, in Proc. IEEE Radiat. Effects Data Workshop (REDW), pp. 178 – 182 (2017).

[152] B. R. Zink, J. Yang-Scharlotta, F. Mancoff, J. Sun, M. Han, and J.-P. Wang, *Influence of Ionizing Dose on Magnetic Tunnel Junctions With Perpendicular Anisotropy*, IEEE Trans. Nucl. Sci., **68**, 748 – 755 (2021).

[153] A. Vatankhahghadim, S. Huda, and A. Sheikholeslami, *A survey on circuit modeling of spin-transfer-torque magnetic tunnel junctions*, IEEE Trans. Circuits Syst. I, **61**, 2634 – 2643 (2014).

[154] J. Song, *Evaluation of operating margin and switching probability of voltage-controlled magnetic anisotropy magnetic tunnel junctions*, IEEE J. Explor. Solid-State Comput. Devices Circuits, **4**, 76 – 84 (2018).

Appendix A: Source Code of SC-CRAM arithmetic functions

In this section, I show the MATLAB source code for the arithmetic functions in SC-CRAM described in sections 5.2.1 and 5.2.2.

Part 1. AND multiplication

This part contains the source code for performing multiplication using AND logic in SC-CRAM, results of which are shown in 5.10(a).

```
%Test multiplication in SC-CRAM

%Desired input probabilities, bit-stream lengths, and pulse widths
figNum=1;
pA_Des=[0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9];
pB_Des=[0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9];
bitRes=8;
Npts=length(pA_Des);
Nbits=round(2^bitRes);
Ntrials=100;
tP=2e-9;
tW=5e-9;
Plogic=0.99;

%Set variation in delta, Vco and resistances
sigR=0.1;
sigDel=sigR;
sigAv=0.00;
sigVco=sigR*0.1;

%Intrinsic MTJ properties
Vco_Des=0.65;
Rp_Des=1000;
Rap_Des=2500;
del_Des=60;
Av_Des=2.1e9;
tau0=1e-9;

%Calculate Vlogic
```

```
tau_logic=-tW/log(1-Plogic);
if tW>=5e-9
   Vswitch=Vco_Des*(1-(1/del_Des)*log(tau_logic/tau0));
else
   Vswitch=(1/(tau_logic*Av_Des))+Vco_Des;
end
Rswitch=((Rp_Des*Rap_Des)/(Rp_Des+Rap_Des));
Ilogic=Vswitch/Rap_Des;
Vlogic=Ilogic*(Rswitch+Rap_Des);


Ymeas=zeros(Ntrials,Npts);
Ycalc_Des=zeros(1,Npts);
for n=1:Ntrials
   for j=1:Npts
      %Find real values of Vco, Rap, Rp, and del
      Vco=zeros(3,1);
      Rp=zeros(3,1);
      Rap=zeros(3,1);
      del=zeros(3,1);
      Av=zeros(3,1);
      for i=1:3
         varVco=(2*rand-1)*sigVco*Vco_Des;
         varRap=(2*rand-1)*sigR*Rap_Des;
         varRp=(2*rand-1)*sigR*Rp_Des;
         varDel=(2*rand-1)*sigDel*del_Des;
         varAv=(2*rand-1)*sigAv*Av_Des;
         Vco(i)=Vco_Des+varVco;
         Rap(i)=Rap_Des+varRp;
         Rp(i)=Rp_Des+varRp;
         del(i)=del_Des+varDel;
         Av(i)=Av_Des+varAv;
      end

      %Calculate perturb voltages based on input probabilities
      if tP>=5e-9
         VA=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pA_Des(j)))));
         VB=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pB_Des(j)))));
         tauA=tau0*exp(del(1)*(1-(VA/Vco(1))));
         tauB=tau0*exp(del(2)*(1-(VB/Vco(2))));
      else
         VA=Vco_Des-(1/(tP*Av_Des))*log(1-pA_Des(j));
         VB=Vco_Des-(1/(tP*Av_Des))*log(1-pB_Des(j));
         tauA=1/(Av(1)*(VA-Vco(1)));
         tauB=1/(Av(2)*(VB-Vco(2)));
      end
```

```
pA=1-exp(-tP/tauA);
pB=1-exp(-tP/tauB);

%Initialized bit-streams
A=zeros(Nbits,1);
B=zeros(Nbits,1);
Y=ones(Nbits,1);
Ycalc_Des(1,j)=pA_Des(j)*pB_Des(j);
Ycalc=pA*pB;
for k=1:Nbits
   rA=rand;
   rB=rand;
   if rA<pA
      A(k)=1;
   end

   if rB<pB
      B(k)=1;
   end

   if A(k)==0 && B(k)==0
      Rin=(Rp(1)*Rp(2))/(Rp(1)+Rp(2));
   elseif A(k)==1 && B(k)==0
      Rin=(Rap(1)*Rp(2))/(Rap(1)+Rp(2));
   elseif A(k)==0 && B(k)==1
      Rin=(Rp(1)*Rap(2))/(Rp(1)+Rap(2));
   else
      Rin=(Rap(1)*Rap(2))/(Rap(1)+Rap(2));
   end

   Req=Rin+Rap(3);
   IW=Vlogic/Req;
   VY=IW*Rap(3);
   if tW>=5e-9
      tauY=tau0*exp(del(3)*(1-(VY/Vco(3))));
   else
      tauY=1/(Av(3)*(VY-Vco(3)));
   end
   pY=1-exp(-tW/tauY);
   rY=rand;
   if rY<pY
      Y(k)=0;
   end
end
Ymeas(n,j)=mean(Y);
Ameas=mean(A);
```

```
      Bmeas=mean(B);
   end
end

Yavg=zeros(1,Npts);
Ydev=zeros(1,Npts);
for j=1:Npts
   Yavg(1,j)=mean(Ymeas(:,j));
   Ydev(1,j)=std(Ymeas(:,j))/sqrt(Ntrials);
end
outData=[transpose(pA_Des),transpose(Ycalc_Des),transpose(Yavg),transpose(Ydev)];

if tP>=5e-9
   Vcrit=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-0.63))));
else
   Vcrit=Vco_Des-(1/(tP*Av_Des))*log(1-0.63);
end

figure(figNum)
plot(pA_Des,Yavg,'ko-',pA_Des,Ycalc_Des,'r--')
xlabel('pA')
ylabel('Y')
legend('measured','expected')
```

Part 2. Scaled addition using MAJ3 logic

This part contains the source code for performing scaled addition using MAJ3 logic in SC-CRAM, results of which are shown in 5.10(b).

```
%Test scaled addition using MAJ3 logic in SC-CRAM

%Desired input probabilities, bit-stream lenghts, and pulse widths
figNum=1;
pA_Des=[0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9];
pB_Des=[0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9];
pS_Des=0.5;
bitRes=10;
Npts=length(pA_Des);
Nbits=round(2^bitRes);
```

```
Ntrials=100;
tP=2e-9;
tW=5e-9;
Plogic=0.99;

%Set variation in Vco and resistances
sigR=0.0;
sigDel=sigR;
sigAv=0.00;
sigVc0=sigR*0.1;

%Intrinsic MTJ properties
Vco_Des=0.65;
Rp_Des=1000;
Rap_Des=3000;
del_Des=60;
Av_Des=2.1e9;
tau0=1e-9;

%Calculate Vlogic
tau_logic=-tW/log(1-Plogic);
if tW>=5e-9
   Vswitch=Vco_Des*(1-(1/del_Des)*log(tau_logic/tau0));
else
   Vswitch=(1/(tau_logic*Av_Des))+Vco_Des;
end
Rswitch=((Rp_Des*Rap_Des)/(Rp_Des+2*Rap_Des));
Ilogic=Vswitch/Rap_Des;
Vlogic=Ilogic*(Rswitch+Rap_Des);

Ymeas=zeros(Ntrials,Npts);
Ycalc_Des=zeros(1,Npts);
for n=1:Ntrials
   for j=1:Npts
      %Find real values of Vco, Rap, Rp, and del
      Vco=zeros(4,1);
      Rp=zeros(4,1);
      Rap=zeros(4,1);
      del=zeros(4,1);
      Av=zeros(4,1);
      for i=1:4
         varVco=(2*rand-1)*sigVco*Vco_Des;
         varRap=(2*rand-1)*sigR*Rap_Des;
         varRp=(2*rand-1)*sigR*Rp_Des;
         varDel=(2*rand-1)*sigDel*del_Des;
         varAv=(2*rand-1)*sigAv*Av_Des;
```

```
    Vco(i)=Vco_Des+varVco;
    Rap(i)=Rap_Des+varRp;
    Rp(i)=Rp_Des+varRp;
    del(i)=del_Des+varDel;
    Av(i)=Av_Des+varAv;
end

%Calculate perturb voltages based on input probabilities
if tP>=5e-9
    VA=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pA_Des(j)))));
    VB=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pB_Des(j)))));
    VS=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pS_Des))));
    tauA=tau0*exp(del(1)*(1-(VA/Vco(1))));
    tauB=tau0*exp(del(2)*(1-(VB/Vco(2))));
    tauS=tau0*exp(del(3)*(1-(VS/Vco(3))));
else
    VA=Vco_Des-(1/(tP*Av_Des))*log(1-pA_Des(j));
    VB=Vco_Des-(1/(tP*Av_Des))*log(1-pB_Des(j));
    VS=Vco_Des-(1/(tP*Av_Des))*log(1-pS_Des);
    tauA=1/(Av(1)*(VA-Vco(1)));
    tauB=1/(Av(2)*(VB-Vco(2)));
    tauS=1/(Av(3)*(VB-Vco(3)));
end
pA=1-exp(-tP/tauA);
pB=1-exp(-tP/tauB);
pS=1-exp(-tP/tauS);

%Initialized bit-streams
A=zeros(Nbits,1);
B=zeros(Nbits,1);
S=zeros(Nbits,1);
Y=ones(Nbits,1);
Ycalc_Des(1,j)=pS_Des*(pA_Des(j)+pB_Des(j));
Ycalc=pA*pB;
for k=1:Nbits
    rA=rand;
    rB=rand;
    rS=rand;
    if rA<pA
        A(k)=1;
    end

    if rB<pB
        B(k)=1;
    end
```

```
            if rS<pS
               S(k)=1;
            end

            if A(k)==0 && B(k)==0 && S(k)==0
               Rin=(Rp(1)*Rp(2)*Rp(3))/(Rp(2)*Rp(3)+Rp(1)*Rp(3)+Rp(1)*Rp(2));
            elseif A(k)==0 && B(k)==0 && S(k)==1
               Rin=(Rp(1)*Rp(2)*Rap(3))/(Rp(2)*Rap(3)+Rp(1)*Rap(3)+Rp(1)*Rp(2));
            elseif A(k)==0 && B(k)==1 && S(k)==0
               Rin=(Rp(1)*Rap(2)*Rp(3))/(Rap(2)*Rp(3)+Rp(1)*Rp(3)+Rp(1)*Rap(2));
            elseif A(k)==1 && B(k)==0 && S(k)==0
               Rin=(Rap(1)*Rp(2)*Rp(3))/(Rp(2)*Rp(3)+Rap(1)*Rp(3)+Rap(1)*Rp(2));
            elseif A(k)==1 && B(k)==0 && S(k)==1
               Rin=(Rap(1)*Rp(2)*Rap(3))/(Rp(2)*Rap(3)+Rap(1)*Rap(3)+Rap(1)*Rp(2));
            elseif A(k)==1 && B(k)==1 && S(k)==0
               Rin=(Rap(1)*Rap(2)*Rp(3))/(Rap(2)*Rp(3)+Rap(1)*Rp(3)+Rap(1)*Rap(2));
            elseif A(k)==1 && B(k)==1 && S(k)==1

Rin=(Rap(1)*Rap(2)*Rap(3))/(Rap(2)*Rap(3)+Rap(1)*Rap(3)+Rap(1)*Rap(2));
            end

            Req=Rin+Rap(3);
            IW=Vlogic/Req;
            VY=IW*Rap(3);
            if tW>=5e-9
               tauY=tau0*exp(del(4)*(1-(VY/Vco(4))));
            else
               tauY=1/(Av(4)*(VY-Vco(4)));
            end
            pY=1-exp(-tW/tauY);
            rY=rand;
            if rY<pY
               Y(k)=0;
            end
         end
         Ymeas(n,j)=mean(Y);
         Ameas=mean(A);
         Bmeas=mean(B);
      end
end

Yavg=zeros(1,Npts);
Ydev=zeros(1,Npts);
for j=1:Npts
   Yavg(1,j)=mean(Ymeas(:,j));
   Ydev(1,j)=std(Ymeas(:,j))/sqrt(Ntrials);
```

```
end
outData=[transpose(pA_Des),transpose(Ycalc_Des),transpose(Yavg),transpose(Ydev)];

figure(figNum)
plot(pA_Des,Yavg,'ko-',pA_Des,Ycalc_Des,'r--')
xlabel('pA')
ylabel('Y')
legend('measured','expected')
```

Part 3. Scaled addition using MUX

This part contains the source code for performing scaled addition using MUX in SC-CRAM, results of which are shown in 5.10(c).

```
%Test scaled addition using a MUX in SC-CRAM

%Desired input probabilities, bit-stream lenghts, and pulse widths
figNum=1;
pA_Des=[0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9];
pB_Des=[0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9];
pS_Des=0.5;
bitRes=8;
Ntrials=100;
Nbits=round(2^bitRes);
Npts=length(pA_Des);
tP=2e-9;
tW=5e-9;
P_AND=0.99;
P_NAND=0.99;
P_OR=0.95;

%Set variation in Vco and resistances
sigR=0.00;
sigDel=sigR;
sigAv=0.00;
sigVco=sigR*0.1;

%Intrinsic MTJ properties
Vco_Des=0.65;
Rp_Des=1000;
```

```
Rap_Des=2500;
del_Des=60;
tau0=1e-9;

%Calculate Vlogic
tau_AND=-tW/log(1-P_AND);
if tW>=5e-9
   Vswitch_AND=Vco_Des*(1-(1/del_Des)*log(tau_AND/tau0));
else
   Vswitch_AND=(1/(tau_AND*Av_Des))+Vco_Des;
end
Rswitch_AND=((Rp_Des*Rap_Des)/(Rp_Des+Rap_Des));
Ilogic_AND=Vswitch_AND/Rap_Des;
V_AND=Ilogic_AND*(Rswitch_AND+Rap_Des);

tau_NAND=-tW/log(1-P_NAND);
if tW>=5e-9
   Vswitch_NAND=Vco_Des*(1-(1/del_Des)*log(tau_NAND/tau0));
else
   Vswitch_NAND=(1/(tau_NAND*Av_Des))+Vco_Des;
end
Rswitch_NAND=((Rp_Des*Rap_Des)/(Rp_Des+Rap_Des));
Ilogic_NAND=Vswitch_NAND/Rp_Des;
V_NAND=Ilogic_NAND*(Rswitch_NAND+Rp_Des);

tau_OR=-tW/log(1-P_OR);
if tW>=5e-9
   Vswitch_OR=Vco_Des*(1-(1/del_Des)*log(tau_OR/tau0));
else
   Vswitch_OR=(1/(tau_OR*Av_Des))+Vco_Des;
end
Rswitch_OR=((Rp_Des*Rp_Des)/(Rp_Des+Rp_Des));
Ilogic_OR=Vswitch_OR/Rap_Des;
V_OR=Ilogic_OR*(Rswitch_OR+Rap_Des);

Ymeas=zeros(Ntrials,Npts);
Ycalc_Des=zeros(1,Npts);
for n=1:Ntrials
   for j=1:Npts
      %Find real values of Vco, Rap, Rp, and del
      Vco=zeros(7,1);
      Rp=zeros(7,1);
      Rap=zeros(7,1);
      del=zeros(7,1);
      for i=1:7
         varVco=(2*rand-1)*sigVco*Vco_Des;
```

```
    varRap=(2*rand-1)*sigR*Rap_Des;
    varRp=(2*rand-1)*sigR*Rp_Des;
    varDel=(2*rand-1)*sigDel*del_Des;
    Vco(i)=Vco_Des+varVco;
    Rap(i)=Rap_Des+varRp;
    Rp(i)=Rp_Des+varRp;
    del(i)=del_Des+varDel;
end

%Calculate perturb voltages based on input probabilities
VA=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pA_Des(j)))));
VB=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pB_Des(j)))));
VS=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pS_Des))));
tauA=tau0*exp(del(1)*(1-(VA/Vco(1))));
tauB=tau0*exp(del(2)*(1-(VB/Vco(2))));
tauS=tau0*exp(del(3)*(1-(VS/Vco(3))));
pA=1-exp(-tP/tauA);
pB=1-exp(-tP/tauB);
pS=1-exp(-tP/tauS);

%Initialized bit-streams
A=zeros(Nbits,1);
B=zeros(Nbits,1);
S=zeros(Nbits,2);
M=ones(Nbits,2);
Y=ones(Nbits,1);
Ycalc=pS*pA+(1-pS)*pB;
Ycalc_Des(1,j)=pS_Des*pA_Des(j)+(1-pS_Des)*pB_Des(j);
SerrCount=0;
for k=1:Nbits
    rA=rand;
    rB=rand;
    rS=rand;
    if rA<pA
        A(k)=1;
    end

    if rB<pB
        B(k)=1;
    end

    %NOT operation (using NAND logic with either input A and B, which were
initialized to 0)
    if rS<pS
        S(k,1)=1;
        Rin_Sn=(Rap(3)*Rap(5))/(Rap(3)+Rap(5));
```

```
else
   Rin_Sn=(Rp(3)*Rap(5))/(Rp(3)+Rap(5));
end

Req_Sn=Rin_Sn+Rp(4);
IW_Sn=V_NAND/Req_Sn;
VSn=IW_Sn*Rp(4);
tau_Sn=tau0*exp(del(4)*(1-(VSn/Vco(4))));
pSn=1-exp(-tW/tau_Sn);
rSn=rand;
if rSn<pSn
   S(k,2)=1;
end

%AND operation of A and S
if A(k)==0 && S(k,1)==0
   Rin_M1=(Rp(3)*Rp(1))/(Rp(3)+Rp(1));
elseif A(k)==1 && S(k,1)==0
   Rin_M1=(Rp(3)*Rap(1))/(Rp(3)+Rap(1));
elseif A(k)==0 && S(k,1)==1
   Rin_M1=(Rap(3)*Rp(1))/(Rap(3)+Rp(1));
else
   Rin_M1=(Rap(3)*Rap(1))/(Rap(3)+Rap(1));
end

Req_M1=Rin_M1+Rap(5);
IW_M1=V_AND/Req_M1;
VM1=IW_M1*Rap(5);
tau_M1=tau0*exp(del(5)*(1-(VM1/Vco(5))));
pM1=1-exp(-tW/tau_M1);
rM1=rand;
if rM1<pM1
   M(k,1)=0;
end

%AND operation of B and not(S)
if B(k)==0 && S(k,2)==0
   Rin_M2=(Rp(2)*Rp(4))/(Rp(2)+Rp(4));
elseif B(k)==1 && S(k,2)==0
   Rin_M2=(Rap(2)*Rp(4))/(Rap(2)+Rp(4));
elseif B(k)==0 && S(k,2)==1
   Rin_M2=(Rp(2)*Rap(4))/(Rp(2)+Rap(4));
else
   Rin_M2=(Rap(2)*Rap(4))/(Rap(2)+Rap(4));
end
```

```
        Req_M2=Rin_M2+Rap(6);
        IW_M2=V_AND/Req_M2;
        VM2=IW_M2*Rap(6);
        tau_M2=tau0*exp(del(6)*(1-(VM2/Vco(6))));
        pM2=1-exp(-tW/tau_M2);
        rM2=rand;
        if rM2<pM2
           M(k,2)=0;
        end

        %OR operation of M1 and M2
        if M(k,1)==0 && M(k,2)==0
           Rin_Y=(Rp(5)*Rp(6))/(Rp(5)+Rp(6));
        elseif M(k,1)==1 && M(k,2)==0
           Rin_Y=(Rap(5)*Rp(6))/(Rap(5)+Rp(6));
        elseif M(k,1)==0 && M(k,2)==1
           Rin_Y=(Rp(5)*Rap(6))/(Rp(5)+Rap(6));
        else
           Rin_Y=(Rap(5)*Rap(6))/(Rap(5)+Rap(6));
        end

        Req_Y=Rin_Y+Rap(7);
        IW_Y=V_OR/Req_Y;
        VY=IW_Y*Rap(7);
        tau_Y=tau0*exp(del(7)*(1-(VY/Vco(7))));
        pY=1-exp(-tW/tau_Y);
        rY=rand;
        if rY<pY
           Y(k)=0;
        end

        if S(k,1)==S(k,2)
           SerrCount=SerrCount+1;
        end
      end
      Ymeas(n,j)=mean(Y);
      Ameas=mean(A);
      Bmeas=mean(B);
      Smeas=mean(S(:,1));
      Snmeas=mean(S(:,2));
      Serr=SerrCount/Nbits;
   end
end

Yavg=zeros(1,Npts);
Ydev=zeros(1,Npts);
```

```
for j=1:Npts
   Yavg(1,j)=mean(Ymeas(:,j));
   Ydev(1,j)=std(Ymeas(:,j))/sqrt(Ntrials);
end
outData=[transpose(pA_Des),transpose(Ycalc_Des),transpose(Yavg),transpose(Ydev)];

figure(figNum)
plot(pA_Des,Yavg,'ko-',pA_Des,Ycalc_Des,'r--')
xlabel('pA')
ylabel('Y')
legend('measured','expected')
```

Part 4. Absolute value subtraction using XOR logic

This part contains the source code for performing absolute value subtraction ($Y = |X-X^2|$) in SC-CRAM, results of which are shown in 5.10(d).

```
%Test absolute valued subtraction with XOR in SC-CRAM
%Note that this code test the function Y=|X-X^2|

%Desired input probabilities, bit-stream lenghts, and pulse widths
figNum=1;
pX_Des=[0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9];
bitRes=10;
Npts=length(pX_Des);
Nbits=round(2^bitRes);
Ntrials=100;
tP=2e-9;
tW=5e-9;
P_AND=0.99;
P_NAND=0.995;
P_OR=0.995;

%Set variation in Vco and resistances
sigR=0.1;
sigDel=sigR;
sigAv=0.00;
sigVco=sigR*0.1;

%Intrinsic MTJ properties
```

```
Vco_Des=0.65;
Rp_Des=1000;
Rap_Des=2500;
del_Des=60;
Av_Des=2.1e9;
tau0=1e-9;

%Calculate Vlogic
tau_AND=-tW/log(1-P_AND);
if tW>=5e-9
   Vswitch_AND=Vco_Des*(1-(1/del_Des)*log(tau_AND/tau0));
else
   Vswitch_AND=(1/(tau_AND*Av_Des))+Vco_Des;
end
Rswitch_AND=((Rp_Des*Rap_Des)/(Rp_Des+Rap_Des));
I_AND=Vswitch_AND/Rap_Des;
V_AND=I_AND*(Rswitch_AND+Rap_Des);

tau_OR=-tW/log(1-P_OR);
if tW>=5e-9
   Vswitch_OR=Vco_Des*(1-(1/del_Des)*log(tau_OR/tau0));
else
   Vswitch_OR=(1/(tau_OR*Av_Des))+Vco_Des;
end
Rswitch_OR=((Rp_Des*Rp_Des)/(Rp_Des+Rp_Des));
I_OR=Vswitch_OR/Rap_Des;
V_OR=I_OR*(Rswitch_OR+Rap_Des);

tau_NAND=-tW/log(1-P_NAND);
if tW>=5e-9
   Vswitch_NAND=Vco_Des*(1-(1/del_Des)*log(tau_NAND/tau0));
else
   Vswitch_NAND=(1/(tau_NAND*Av_Des))+Vco_Des;
end
Rswitch_NAND=((Rp_Des*Rap_Des)/(Rp_Des+Rap_Des));
I_NAND=Vswitch_NAND/Rp_Des;
V_NAND=I_NAND*(Rswitch_NAND+Rp_Des);


Ymeas=zeros(Ntrials,Npts);
XSmeas=zeros(2,Npts);
Ycalc_Des=zeros(1,Npts);
for n=1:Ntrials
   for j=1:Npts
      %Find real values of Vco, Rap, Rp, and del
      Vco=zeros(6,1);
```

```
Rp=zeros(6,1);
Rap=zeros(6,1);
del=zeros(6,1);
Av=zeros(6,1);
for i=1:6
   varVco=(2*rand-1)*sigVco*Vco_Des;
   varRap=(2*rand-1)*sigR*Rap_Des;
   varRp=(2*rand-1)*sigR*Rp_Des;
   varDel=(2*rand-1)*sigDel*del_Des;
   varAv=(2*rand-1)*sigAv*Av_Des;
   Vco(i)=Vco_Des+varVco;
   Rap(i)=Rap_Des+varRp;
   Rp(i)=Rp_Des+varRp;
   del(i)=del_Des+varDel;
   Av(i)=Av_Des+varAv;
end

%Calculate perturb voltages based on input probabilities
if tP>=5e-9
   VX=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pX_Des(j)))));
   tauX1=tau0*exp(del(1)*(1-(VX/Vco(1))));
   tauX2=tau0*exp(del(2)*(1-(VX/Vco(2))));
else
   VX=Vco_Des-(1/(tP*Av_Des))*log(1-pX_Des(j));
   tauX1=1/(Av(1)*(VX-Vco(1)));
   tauX2=1/(Av(2)*(VX-Vco(2)));
end
pX1=1-exp(-tP/tauX1);
pX2=1-exp(-tP/tauX2);

%Initialized bit-streams
X=zeros(Nbits,2);
XS=ones(Nbits,1);
M1=zeros(Nbits,1);
M2=ones(Nbits,1);
Y=ones(Nbits,1);
Ycalc_Des(1,j)=abs(pX_Des(j)-(pX_Des(j)^2));
Ycalc=abs(pX1-(pX1^2));
for k=1:Nbits
   rX1=rand;
   rX2=rand;
   if rX1<pX1
      X(k,1)=1;
   end

   if rX2<pX2
```

```
    X(k,2)=1;
end

%Find XS
if X(k,1)==0 && X(k,2)==0
    Rin_XS=(Rp(1)*Rp(2))/(Rp(1)+Rp(2));
elseif X(k,1)==1 && X(k,2)==0
    Rin_XS=(Rap(1)*Rp(2))/(Rap(1)+Rp(2));
elseif X(k,1)==0 && X(k,2)==1
    Rin_XS=(Rp(1)*Rap(2))/(Rp(1)+Rap(2));
else
    Rin_XS=(Rap(1)*Rap(2))/(Rap(1)+Rap(2));
end

Req_XS=Rin_XS+Rap(3);
IXS=V_AND/Req_XS;
VXS=IXS*Rap(3);
if tW>=5e-9
    tauXS=tau0*exp(del(3)*(1-(VXS/Vco(3))));
else
    tauXS=1/(Av(3)*(VXS-Vco(3)));
end
pXS=1-exp(-tW/tauXS);
rXS=rand;
if rXS<pXS
    XS(k)=0;
end

%Find M1
if X(k,1)==0 && XS(k)==0
    Rin_M1=(Rp(1)*Rp(3))/(Rp(1)+Rp(3));
elseif X(k,1)==1 && XS(k)==0
    Rin_M1=(Rap(1)*Rp(3))/(Rap(1)+Rp(3));
elseif X(k,1)==0 && XS(k)==1
    Rin_M1=(Rp(1)*Rap(3))/(Rp(1)+Rap(3));
else
    Rin_M1=(Rap(1)*Rap(3))/(Rap(1)+Rap(3));
end

Req_M1=Rin_M1+Rp(4);
IM1=V_NAND/Req_M1;
VM1=IM1*Rp(4);
if tW>=5e-9
    tauM1=tau0*exp(del(4)*(1-(VM1/Vco(4))));
else
    tauM1=1/(Av(4)*(VM1-Vco(4)));
```

```
end
pM1=1-exp(-tW/tauM1);
rM1=rand;
if rM1<pM1
   M1(k)=1;
end

%Find M2
if X(k,1)==0 && XS(k)==0
   Rin_M2=(Rp(1)*Rp(3))/(Rp(1)+Rp(3));
elseif X(k,1)==1 && XS(k)==0
   Rin_M2=(Rap(1)*Rp(3))/(Rap(1)+Rp(3));
elseif X(k,1)==0 && XS(k)==1
   Rin_M2=(Rp(1)*Rap(3))/(Rp(1)+Rap(3));
else
   Rin_M2=(Rap(1)*Rap(3))/(Rap(1)+Rap(3));
end

Req_M2=Rin_M2+Rap(5);
IM2=V_OR/Req_M2;
VM2=IM2*Rap(5);
if tW>=5e-9
   tauM2=tau0*exp(del(5)*(1-(VM2/Vco(5))));
else
   tauM2=1/(Av(5)*(VM2-Vco(5)));
end
pM2=1-exp(-tW/tauM2);
rM2=rand;
if rM2<pM2
   M2(k)=0;
end

%Find Y
if M1(k)==0 && M2(k)==0
   Rin_Y=(Rp(4)*Rp(5))/(Rp(4)+Rp(5));
elseif M1(k)==1 && M2(k)==0
   Rin_Y=(Rap(4)*Rp(5))/(Rap(4)+Rp(5));
elseif M1(k)==0 && M2(k)==1
   Rin_Y=(Rp(4)*Rap(5))/(Rp(4)+Rap(5));
else
   Rin_Y=(Rap(4)*Rap(5))/(Rap(4)+Rap(5));
end

Req_Y=Rin_Y+Rap(6);
IY=V_AND/Req_Y;
VY=IY*Rap(6);
```

```
        if tW>=5e-9
            tauY=tau0*exp(del(6)*(1-(VY/Vco(6))));
        else
            tauY=1/(Av(6)*(VY-Vco(6)));
        end
        pY=1-exp(-tW/tauY);
        rY=rand;
        if rY<pY
            Y(k)=0;
        end
      end
      Ymeas(n,j)=mean(Y);
      XSmeas(1,j)=mean(XS);
      XSmeas(2,j)=pX_Des(j)^2;
   end
end

Yavg=zeros(1,Npts);
Ydev=zeros(1,Npts);
for j=1:Npts
   Yavg(1,j)=mean(Ymeas(:,j));
   Ydev(1,j)=std(Ymeas(:,j))/sqrt(Ntrials);
end
outData=[transpose(pX_Des),transpose(Ycalc_Des),transpose(Yavg),transpose(Ydev)];

figure(figNum)
plot(pX_Des,Yavg,'ko-',pX_Des,Ycalc_Des,'r--')
xlabel('pX')
ylabel('Y')
legend('measured','expected')
```

Part 5. Square root using AND-OR-OR logic

This part contains the source code for performing a square root function in SC-CRAM, results of which are shown in 5.10(e).

```
%Test square root in SC-CRAM
%Note that this code test the function Y=sqrt(X)

%Desired input probabilities, bit-stream lenghts, and pulse widths
```

```
figNum=1;
pX_Des=[0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9];
pC_Des=[0.67,0.18];
bitRes=8;
Npts=length(pX_Des);
Nbits=round(2^bitRes);
Ntrials=100;
tP=2e-9;
tW=5e-9;
P_AND=0.99;
P_OR=0.96;

%Set variation in Vco and resistances
sigR=0.05;
sigDel=sigR;
sigAv=0.00;
sigVco=sigR*0.1;

%Intrinsic MTJ properties
Vco_Des=0.65;
Rp_Des=1000;
Rap_Des=2500;
del_Des=60;
Av_Des=2.1e9;
tau0=1e-9;

%Calculate Vlogic
tau_AND=-tW/log(1-P_AND);
if tW>=5e-9
   Vswitch_AND=Vco_Des*(1-(1/del_Des)*log(tau_AND/tau0));
else
   Vswitch_AND=(1/(tau_AND*Av_Des))+Vco_Des;
end
Rswitch_AND=((Rp_Des*Rap_Des)/(Rp_Des+Rap_Des));
I_AND=Vswitch_AND/Rap_Des;
V_AND=I_AND*(Rswitch_AND+Rap_Des);

tau_OR=-tW/log(1-P_OR);
if tW>=5e-9
   Vswitch_OR=Vco_Des*(1-(1/del_Des)*log(tau_OR/tau0));
else
   Vswitch_OR=(1/(tau_OR*Av_Des))+Vco_Des;
end
Rswitch_OR=((Rp_Des*Rp_Des)/(Rp_Des+Rp_Des));
I_OR=Vswitch_OR/Rap_Des;
V_OR=I_OR*(Rswitch_OR+Rap_Des);
```

```
Ymeas=zeros(Ntrials,Npts);
XSmeas=zeros(2,Npts);
Ycalc_Des=zeros(1,Npts);
for n=1:Ntrials
   for j=1:Npts
      %Find real values of Vco, Rap, Rp, and del
      Vco=zeros(7,1);
      Rp=zeros(7,1);
      Rap=zeros(7,1);
      del=zeros(7,1);
      Av=zeros(7,1);
      for i=1:7
         varVco=(2*rand-1)*sigVco*Vco_Des;
         varRap=(2*rand-1)*sigR*Rap_Des;
         varRp=(2*rand-1)*sigR*Rp_Des;
         varDel=(2*rand-1)*sigDel*del_Des;
         varAv=(2*rand-1)*sigAv*Av_Des;
         Vco(i)=Vco_Des+varVco;
         Rap(i)=Rap_Des+varRp;
         Rp(i)=Rp_Des+varRp;
         del(i)=del_Des+varDel;
         Av(i)=Av_Des+varAv;
      end

      %Calculate perturb voltages based on input probabilities
      if tP>=5e-9
         VX=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pX_Des(j)))));
         tauX1=tau0*exp(del(1)*(1-(VX/Vco(1))));
         tauX2=tau0*exp(del(2)*(1-(VX/Vco(2))));
      else
         VX=Vco_Des-(1/(tP*Av_Des))*log(1-pX_Des(j));
         tauX1=1/(Av(1)*(VX-Vco(1)));
         tauX2=1/(Av(2)*(VX-Vco(2)));
      end
      pX1=1-exp(-tP/tauX1);
      pX2=1-exp(-tP/tauX2);

      %Initialized bit-streams
      X=zeros(Nbits,2);
      C=zeros(Nbits,2);
      M1=ones(Nbits,1);
      M2=ones(Nbits,1);
      Y=ones(Nbits,1);
      Ycalc_Des(1,j)=sqrt(pX_Des(j));
      Ycalc=sqrt(pX1);
```

```
for k=1:Nbits
   rX1=rand;
   rX2=rand;
   rC1=rand;
   rC2=rand;
   if rX1<pX1
      X(k,1)=1;
   end

   if rX2<pX2
      X(k,2)=1;
   end

   if rC1<pC_Des(1)
      C(k,1)=1;
   end

   if rC2<pC_Des(2)
      C(k,2)=1;
   end

   %Find M1
   if X(k,1)==0 && C(k,1)==0
      Rin_M1=(Rp(1)*Rp(3))/(Rp(1)+Rp(3));
   elseif X(k,1)==1 && C(k,1)==0
      Rin_M1=(Rap(1)*Rp(3))/(Rap(1)+Rp(3));
   elseif X(k,1)==0 && C(k,1)==1
      Rin_M1=(Rp(1)*Rap(3))/(Rp(1)+Rap(3));
   else
      Rin_M1=(Rap(1)*Rap(3))/(Rap(1)+Rap(3));
   end

   Req_M1=Rin_M1+Rap(5);
   IM1=V_AND/Req_M1;
   VM1=IM1*Rap(5);
   if tW>=5e-9
      tauM1=tau0*exp(del(5)*(1-(VM1/Vco(5))));
   else
      tauM1=1/(Av(5)*(VM1-Vco(5)));
   end
   pM1=1-exp(-tW/tauM1);
   rM1=rand;
   if rM1<pM1
      M1(k)=0;
   end
```

```
%Find M2
if X(k,2)==0 && M1(k)==0
   Rin_M2=(Rp(2)*Rp(5))/(Rp(2)+Rp(5));
elseif X(k,2)==1 && M1(k)==0
   Rin_M2=(Rap(2)*Rp(5))/(Rap(2)+Rp(5));
elseif X(k,2)==0 && M1(k)==1
   Rin_M2=(Rp(2)*Rap(5))/(Rp(2)+Rap(5));
else
   Rin_M2=(Rap(2)*Rap(5))/(Rap(2)+Rap(5));
end

Req_M2=Rin_M2+Rap(6);
IM2=V_OR/Req_M2;
VM2=IM2*Rap(6);
if tW>=5e-9
   tauM2=tau0*exp(del(6)*(1-(VM2/Vco(6))));
else
   tauM2=1/(Av(6)*(VM2-Vco(6)));
end
pM2=1-exp(-tW/tauM2);
rM2=rand;
if rM2<pM2
   M2(k)=0;
end

%Find Y
if C(k,2)==0 && M2(k)==0
   Rin_Y=(Rp(4)*Rp(6))/(Rp(4)+Rp(6));
elseif C(k,2)==1 && M2(k)==0
   Rin_Y=(Rap(4)*Rp(6))/(Rap(4)+Rp(6));
elseif C(k,2)==0 && M2(k)==1
   Rin_Y=(Rp(4)*Rap(6))/(Rp(4)+Rap(6));
else
   Rin_Y=(Rap(4)*Rap(6))/(Rap(4)+Rap(6));
end

Req_Y=Rin_Y+Rap(7);
IY=V_OR/Req_Y;
VY=IY*Rap(7);
if tW>=5e-9
   tauY=tau0*exp(del(7)*(1-(VY/Vco(7))));
else
   tauY=1/(Av(7)*(VY-Vco(7)));
end
pY=1-exp(-tW/tauY);
rY=rand;
```

```
        if rY<pY
            Y(k)=0;
        end
    end
    Ymeas(n,j)=mean(Y);
  end
end

Yavg=zeros(1,Npts);
Ydev=zeros(1,Npts);
for j=1:Npts
   Yavg(1,j)=mean(Ymeas(:,j));
   Ydev(1,j)=std(Ymeas(:,j))/sqrt(Ntrials);
end
outData=[transpose(pX_Des),transpose(Ycalc_Des),transpose(Yavg),transpose(Ydev)];


figure(figNum)
plot(pX_Des,Yavg,'ko-',pX_Des,Ycalc_Des,'r--')
xlabel('pX')
ylabel('Y')
legend('measured','expected')
```

Part 6. Scaled division using JK flip flop

This part contains the source code for performing scaled division (Y=A/(A+B)) in SC-CRAM, results of which are shown in 5.10(f).

```
%Test scaled division using a JK flip flop in SC-CRAM
%1->A, 2->B, 3->Q, 4->Qn, 5->J, 6->K1, 7->K2, 8->Y

%Desired input probabilities, bit-stream lenghts, and pulse widths
figNum=1;
pA_Des=[0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9];
pB_Des=[0.9,0.8,0.7,0.6,0.5,0.4,0.3,0.2,0.1];
bitRes=12;
Npts=length(pA_Des);
Nbits=round(2^bitRes);
Ntrials=100;
tP=2e-9;
```

```
tW=5e-9;
P_AND=0.99;
P_NAND=0.98;

%Set variation in Vco and resistances
sigR=0.01;
sigDel=sigR;
sigAv=0.00;
sigVco=sigR*0.1;

%Intrinsic MTJ properties
Vco_Des=0.65;
Rp_Des=1000;
Rap_Des=2500;
del_Des=60;
Av_Des=2.1e9;
tau0=1e-9;

%Calculate Vlogic
tau_AND=-tW/log(1-P_AND);
if tW>=5e-9
   Vswitch_AND=Vco_Des*(1-(1/del_Des)*log(tau_AND/tau0));
else
   Vswitch_AND=(1/(tau_AND*Av_Des))+Vco_Des;
end
Rswitch_AND=((Rp_Des*Rap_Des)/(Rp_Des+Rap_Des));
I_AND=Vswitch_AND/Rap_Des;
V_AND=I_AND*(Rswitch_AND+Rap_Des);

tau_NAND=-tW/log(1-P_NAND);
if tW>=5e-9
   Vswitch_NAND=Vco_Des*(1-(1/del_Des)*log(tau_NAND/tau0));
else
   Vswitch_NAND=(1/(tau_NAND*Av_Des))+Vco_Des;
end
Rswitch_NAND=((Rp_Des*Rap_Des)/(Rp_Des+Rap_Des));
I_NAND=Vswitch_NAND/Rp_Des;
V_NAND=I_NAND*(Rswitch_NAND+Rp_Des);

Ymeas=zeros(Ntrials,Npts);
XSmeas=zeros(2,Npts);
Ycalc_Des=zeros(1,Npts);
for n=1:Ntrials
   for j=1:Npts
      %Find real values of Vco, Rap, Rp, and del
      Vco=zeros(9,1);
```

```
Rp=zeros(9,1);
Rap=zeros(9,1);
del=zeros(9,1);
Av=zeros(9,1);
for i=1:9
   varVco=(2*rand-1)*sigVco*Vco_Des;
   varRap=(2*rand-1)*sigR*Rap_Des;
   varRp=(2*rand-1)*sigR*Rp_Des;
   varDel=(2*rand-1)*sigDel*del_Des;
   varAv=(2*rand-1)*sigAv*Av_Des;
   Vco(i)=Vco_Des+varVco;
   Rap(i)=Rap_Des+varRp;
   Rp(i)=Rp_Des+varRp;
   del(i)=del_Des+varDel;
   Av(i)=Av_Des+varAv;
end

%Calculate perturb voltages based on input probabilities
if tP>=5e-9
   VA=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pA_Des(j)))));
   VB=Vco_Des*(1-(1/del_Des)*log(-tP/(tau0*log(1-pB_Des(j)))));
   tauA=tau0*exp(del(1)*(1-(VA/Vco(1))));
   tauB=tau0*exp(del(2)*(1-(VB/Vco(2))));
else
   VA=Vco_Des-(1/(tP*Av_Des))*log(1-pA_Des(j));
   VB=Vco_Des-(1/(tP*Av_Des))*log(1-pB_Des(j));
   tauA=1/(Av(1)*(VA-Vco(1)));
   tauB=1/(Av(2)*(VB-Vco(2)));
end
pA=1-exp(-tP/tauA);
pB=1-exp(-tP/tauB);

%Initialized bit-streams
A=zeros(Nbits,1);
B=zeros(Nbits,1);
Q=ones(Nbits+1,1);
Qn=zeros(Nbits,1);
J=zeros(Nbits,1);
K1=zeros(Nbits,1);
K2=zeros(Nbits,1);
Y=zeros(Nbits,1);
C=ones(Nbits+1,1);
Ycalc_Des(1,j)=pA_Des(j)/(pA_Des(j)+pB_Des(j));
Ycalc=pA/(pA+pB);
%Q(1)=0;
for k=1:Nbits
```

```
rA=rand;
rB=rand;
if rA<pA
   A(k)=1;
end

if rB<pB
   B(k)=1;
end

%Find Qn
if C(k)==0 && Q(k)==0
   Rin_Qn=(Rp(9)*Rp(3))/(Rp(9)+Rp(3));
elseif C(k)==1 && Q(k)==0
   Rin_Qn=(Rap(9)*Rp(3))/(Rap(9)+Rp(3));
elseif C(k)==0 && Q(k)==1
   Rin_Qn=(Rp(9)*Rap(3))/(Rp(9)+Rap(3));
else
   Rin_Qn=(Rap(9)*Rap(3))/(Rap(9)+Rap(3));
end

Req_Qn=Rin_Qn+Rp(4);
IQn=V_NAND/Req_Qn;
VQn=IQn*Rp(4);
if tW>=5e-9
   tauQn=tau0*exp(del(4)*(1-(VQn/Vco(4))));
else
   tauQn=1/(Av(4)*(VQn-Vco(4)));
end
pQn=1-exp(-tW/tauQn);
rQn=rand;
if rQn<pQn
   Qn(k)=1;
end

%Find J
if A(k)==0 && Qn(k)==0
   Rin_J=(Rp(1)*Rp(4))/(Rp(1)+Rp(4));
elseif A(k)==1 && Qn(k)==0
   Rin_J=(Rap(1)*Rp(4))/(Rap(1)+Rp(4));
elseif A(k)==0 && Qn(k)==1
   Rin_J=(Rp(1)*Rap(4))/(Rp(1)+Rap(4));
else
   Rin_J=(Rap(1)*Rap(4))/(Rap(1)+Rap(4));
end
```

```
Req_J=Rin_J+Rp(5);
IJ=V_NAND/Req_J;
VJ=IJ*Rp(5);
if tW>=5e-9
   tauJ=tau0*exp(del(5)*(1-(VJ/Vco(5))));
else
   tauJ=1/(Av(5)*(VJ-Vco(5)));
end
pJ=1-exp(-tW/tauJ);
rJ=rand;
if rJ<pJ
   J(k)=1;
end

%Find K1
if B(k)==0 && Q(k)==0
   Rin_K1=(Rp(2)*Rp(3))/(Rp(2)+Rp(3));
elseif B(k)==1 && Q(k)==0
   Rin_K1=(Rap(2)*Rp(3))/(Rap(2)+Rp(3));
elseif B(k)==0 && Q(k)==1
   Rin_K1=(Rp(2)*Rap(3))/(Rp(2)+Rap(3));
else
   Rin_K1=(Rap(2)*Rap(3))/(Rap(2)+Rap(3));
end

Req_K1=Rin_K1+Rp(6);
IK1=V_NAND/Req_K1;
VK1=IK1*Rp(6);
if tW>=5e-9
   tauK1=tau0*exp(del(6)*(1-(VK1/Vco(6))));
else
   tauK1=1/(Av(6)*(VK1-Vco(6)));
end
pK1=1-exp(-tW/tauK1);
rK1=rand;
if rK1<pK1
   K1(k)=1;
end

%Find K2
if Q(k)==0 && K1(k)==0
   Rin_K2=(Rp(3)*Rp(6))/(Rp(3)+Rp(6));
elseif Q(k)==1 && K1(k)==0
   Rin_K2=(Rap(3)*Rp(6))/(Rap(3)+Rp(6));
elseif Q(k)==0 && K1(k)==1
   Rin_K2=(Rp(3)*Rap(6))/(Rp(3)+Rap(6));
```

```
else
   Rin_K2=(Rap(3)*Rap(6))/(Rap(3)+Rap(6));
end

Req_K2=Rin_K2+Rp(7);
IK2=V_NAND/Req_K2;
VK2=IK2*Rp(7);
if tW>=5e-9
   tauK2=tau0*exp(del(7)*(1-(VK2/Vco(7))));
else
   tauK2=1/(Av(7)*(VK2-Vco(7)));
end
pK2=1-exp(-tW/tauK2);
rK2=rand;
if rK2<pK2
   K2(k)=1;
end

%Find Y
if J(k)==0 && K2(k)==0
   Rin_Y=(Rp(5)*Rp(7))/(Rp(5)+Rp(7));
elseif J(k)==1 && K2(k)==0
   Rin_Y=(Rap(5)*Rp(7))/(Rap(5)+Rp(7));
elseif J(k)==0 && K2(k)==1
   Rin_Y=(Rp(5)*Rap(7))/(Rp(5)+Rap(7));
else
   Rin_Y=(Rap(5)*Rap(7))/(Rap(5)+Rap(7));
end

Req_Y=Rin_Y+Rp(8);
IY=V_NAND/Req_Y;
VY=IY*Rp(8);
if tW>=5e-9
   tauY=tau0*exp(del(8)*(1-(VY/Vco(8))));
else
   tauY=1/(Av(8)*(VY-Vco(8)));
end
pY=1-exp(-tW/tauY);
rY=rand;
if rY<pY
   Y(k)=1;
end

%Find Q(k+1)
if Y(k)==0 && C(k)==0
   Rin_Q=(Rp(8)*Rp(9))/(Rp(8)+Rp(9));
```

```
        elseif Y(k)==1 && C(k)==0
            Rin_Q=(Rap(8)*Rp(9))/(Rap(8)+Rp(9));
        elseif Y(k)==0 && C(k)==1
            Rin_Q=(Rp(8)*Rap(9))/(Rp(8)+Rap(9));
        else
            Rin_Q=(Rap(8)*Rap(9))/(Rap(8)+Rap(9));
        end

        Req_Q=Rin_Q+Rap(3);
        IQ=V_AND/Req_Q;
        VQ=IQ*Rap(3);
        if tW>=5e-9
            tauQ=tau0*exp(del(3)*(1-(VY/Vco(3))));
        else
            tauQ=1/(Av(3)*(VY-Vco(3)));
        end
        pQ=1-exp(-tW/tauQ);
        rQ=rand;
        if rQ<pQ
            Q(k+1)=0;
        end
      end
      Ymeas(n,j)=mean(Y);
      XSmeas(1,j)=mean(B);
      XSmeas(2,j)=pA_Des(j)^2;
   end
end

Yavg=zeros(1,Npts);
Ydev=zeros(1,Npts);
for j=1:Npts
   Yavg(1,j)=mean(Ymeas(:,j));
   Ydev(1,j)=std(Ymeas(:,j))/sqrt(Ntrials);
end
outData=[transpose(pA_Des),transpose(Ycalc_Des),transpose(Yavg),transpose(Ydev)];

figure(figNum)
plot(pA_Des,Yavg,'ko-',pA_Des,Ycalc_Des,'r--')
xlabel('pX')
ylabel('Y')
legend('measured','expected')
```

Appendix B: Source Code of SC-CRAM applications

In this section, I show the MATLAB source code for analyzing the accuracy for the applications in SC-CRAM described in sections 5.3.1 to 5.3.3 and results presented in section 5.4.1.

Part 1. Local image thresholding

This part contains the source code for performing local image thresholding in SC-CRAM, results of which are shown in 5.13(a).

```
%Perform local image thresholding on a grayscale image (data in excel).
%Assumes 9 x 9 window and uses MAJ3 gates for averages

figNum=1;
filePath='C:\Users\zinkb\OneDrive\Documents\2022 Research
updates\Manuscripts\CRAM patent\';
readFile=strcat(filePath,'Digit pixel data.xlsx');
readSheet='Sheet4';
readCells='C3:DG33';
expYCells='C45:DG75';
bitRes=8;
windDim=9;
nRate=0.1;
Ntrials=10;

cFact=[0.67,0.18];
pFact1=0.52;

digData=xlsread(readFile,readSheet,readCells);
expY=xlsread(readFile,readSheet,expYCells);
Nx=length(digData(1,:));
Ny=length(digData(:,1));
xPix=1:1:Nx;
yPix=1:1:Ny;

Nbits=round(2^bitRes);
Nwind=round(windDim^2);
```

```
Ncyc=round((Nx-8)*(Ny-8));
Npts=round(Nx*Ny);

Cnst1=ones(Nbits,1);
yErr=zeros(Ntrials,1);

for t=1:Ntrials
   TVals=zeros(Ny,Nx);
   Y=zeros(Ny,Nx);
   for n=1:Ncyc
      yPos=floor((n-1)/101)+5;
      xPos=105-mod(n,101);

      A=zeros(Nbits,Nwind,2);
      P=zeros(Nbits,Nwind);

      MA1=zeros(Nbits,41);
      MA2=zeros(Nbits,20);
      MA3=zeros(Nbits,10);
      MA4=zeros(Nbits,5);
      MA5=zeros(Nbits,3);
      MA6=zeros(Nbits,2);
      MA=zeros(Nbits,1);

      MP1=zeros(Nbits,41);
      MP2=zeros(Nbits,20);
      MP3=zeros(Nbits,10);
      MP4=zeros(Nbits,5);
      MP5=zeros(Nbits,3);
      MP6=zeros(Nbits,2);
      MP=zeros(Nbits,1);

      Sel1=zeros(Nbits,40);
      Sel2=zeros(Nbits,20);
      Sel3=zeros(Nbits,10);
      Sel4=zeros(Nbits,5);
      Sel5=zeros(Nbits,3);
      Sel6=zeros(Nbits,1);
      Sel=zeros(Nbits,1);
      Amult1=zeros(Nbits,1);
      Amult6=zeros(Nbits,1);

      MS=zeros(Nbits,1);
      D=zeros(Nbits,2);
      DF=zeros(Nbits,1);
```

```
C=zeros(Nbits,2);
R=zeros(Nbits,2);
SDV=zeros(Nbits,1);
MT=zeros(Nbits,3);
SelT=zeros(Nbits,1);
T=zeros(Nbits,1);

for w=1:Nwind
   yIndAdd=floor((w-1)/9)-4;
   yInd=floor((n-1)/101)+5+yIndAdd;
   xIndAdd=-1*mod(w,9)+4;
   xInd=105-mod(n,101)+xIndAdd;
   for k=1:Nbits
      rA=[rand,rand];
      rNA=[rand,rand];
      for a=1:2
         if rA(a)<digData(yInd,xInd)
            A(k,w,a)=1;
         end

         if rNA(a)<nRate
            A(k,w,a)=not(A(k,w,a));
         end
      end
      P(k,w)=and(A(k,w,1),A(k,w,2));
   end
end

for k=1:Nbits
   %MAJ3 layer 1
   rS1=zeros(40,1);
   rSN1=zeros(40,1);
   for m=1:40
      rS1(m)=rand;
      rSN1(m)=rand;
      if rS1(m)<0.5
         Sel1(k,m)=1;
      end

      if rSN1(m)<nRate
         Sel1(k,m)=not(Sel1(k,m));
      end

      indA=2*m-1;
      if sum([A(k,indA,1),A(k,indA+1,1),Sel1(k,m)])>1
         MA1(k,m)=1;
```

```
    end
    if sum([P(k,indA),P(k,indA+1),Sel1(k,m)])>1
        MP1(k,m)=1;
    end
end

rAm1=rand;
rAmN1=rand;
if rAm1<0.03125
    Amult1(k,1)=1;
end

if rAmN1<nRate
    Amult1(k,1)=not(Amult1(k,1));
end

MA1(k,41)=and(A(k,81,1),Amult1(k,1));
MP1(k,41)=and(P(k,81,1),Amult1(k,1));

%MAJ3 layer 2
rS2=zeros(20,1);
rSN2=zeros(20,1);
for m=1:20
    rS2(m)=rand;
    rSN2(m)=rand;
    if rS2(m)<0.5
        Sel2(k,m)=1;
    end

    if rSN2(m)<nRate
        Sel2(k,m)=not(Sel2(k,m));
    end

    indA=2*m-1;
    if sum([MA1(k,indA),MA1(k,indA+1),Sel2(k,m)])>1
        MA2(k,m)=1;
    end
    if sum([MP1(k,indA),MP1(k,indA+1),Sel2(k,m)])>1
        MP2(k,m)=1;
    end
end

%MAJ3 layer 3
rS3=zeros(10,1);
rSN3=zeros(10,1);
for m=1:10
```

```
    rS3(m)=rand;
    rSN3(m)=rand;
    if rS3(m)<0.5
        Sel3(k,m)=1;
    end

    if rSN3(m)<nRate
        Sel3(k,m)=not(Sel3(k,m));
    end

    indA=2*m-1;
    if sum([MA2(k,indA),MA2(k,indA+1),Sel3(k,m)])>1
        MA3(k,m)=1;
    end
    if sum([MP2(k,indA),MP2(k,indA+1),Sel3(k,m)])>1
        MP3(k,m)=1;
    end
end

%MAJ3 layer 4
rS4=zeros(5,1);
rSN4=zeros(5,1);
for m=1:5
    rS4(m)=rand;
    rSN4(m)=rand;
    if rS4(m)<0.5
        Sel4(k,m)=1;
    end

    if rSN4(m)<nRate
        Sel4(k,m)=not(Sel4(k,m));
    end

    indA=2*m-1;
    if sum([MA3(k,indA),MA3(k,indA+1),Sel4(k,m)])>1
        MA4(k,m)=1;
    end
    if sum([MP3(k,indA),MP3(k,indA+1),Sel4(k,m)])>1
        MP4(k,m)=1;
    end
end

%MAJ3 layer 5
rS5=zeros(3,1);
rSN5=zeros(3,1);
for m=1:3
```

```
    rS5(m)=rand;
    rSN5(m)=rand;
    if rS5(m)<0.5
        Sel5(k,m)=1;
    end

    if rSN5(m)<nRate
        Sel5(k,m)=not(Sel5(k,m));
    end

    if m<3
        indA=2*m-1;
        if sum([MA4(k,indA),MA4(k,indA+1),Sel5(k,m)])>1
            MA5(k,m)=1;
        end
        if sum([MP4(k,indA),MP4(k,indA+1),Sel5(k,m)])>1
            MP5(k,m)=1;
        end
    end
end

if sum([MA4(k,5),MA1(k,41),Sel5(k,3)])>1
    MA5(k,3)=1;
end
if sum([MP4(k,5),MP1(k,41),Sel5(k,3)])>1
    MP5(k,3)=1;
end

%MAJ3 layer 6
rS6=rand;
rAm6=rand;
rSN6=rand;
rAmN6=rand;
if rS6<0.5
    Sel6(k,1)=1;
end
if rAm6<0.5
    Amult6(k,1)=1;
end

if rSN6<nRate
    Sel6(k,1)=not(Sel6(k,1));
end
if rAmN6<nRate
    Amult6(k,1)=not(Amult6(k,1));
end
```

```
if sum([MA5(k,1),MA5(k,2),Sel6(k,1)])>1
   MA6(k,1)=1;
end
if sum([MP5(k,1),MP5(k,2),Sel6(k,1)])>1
   MP6(k,1)=1;
end

MA6(k,2)=and(MA5(k,3),Amult6(k,1));
MP6(k,2)=and(MP5(k,3),Amult6(k,1));

%MAJ3 final layer
rS=rand;
rSN=rand;
if rS<0.5
   Sel(k,1)=1;
end

if rSN<nRate
   Sel(k,1)=not(Sel(k,1));
end

if sum([MA6(k,1),MA6(k,2),Sel(k,1)])>1
   MA(k,1)=1;
end
if sum([MP6(k,1),MP6(k,2),Sel(k,1)])>1
   MP(k,1)=1;
end

if k>1
   MS(k,1)=and(MA(k,1),MA(k-1,1));
end
D(k,1)=not(and(MS(k,1),MP(k,1)));
D(k,2)=or(MS(k,1),MP(k,1));
DF(k,1)=and(D(k,1),D(k,1));

rC=[rand,rand];
for c=1:2
   if rC(c)<cFact(c)
      C(k,c)=1;
   end
end
R(k,1)=and(DF(k,1),C(k,1));
if k>1
   R(k,2)=or(R(k,1),DF(k-1,1));
else
```

```
            R(k,2)=R(k,1);
        end
        SDV(k,1)=or(R(k,2),C(k,2));

        rST=rand;
        if rST<0.5
            SelT(k,1)=1;
        end
        MT(k,1)=and(Cnst1(k,1),SelT(k,1));
        MT(k,2)=and(SDV(k,1),not(SelT(k,1)));
        MT(k,3)=or(MT(k,1),MT(k,2));
        T(k,1)=and(MA(k,1),MT(k,3));
    end
    TVals(yPos,xPos)=mean(T(:,1));
    if TVals(yPos,xPos)<=pFact1*digData(yPos,xPos)
        Y(yPos,xPos)=1;
    end
  end

  yErrCount=0;
  for i=1:Ny
    for j=1:Nx
      yErrCount=abs(Y(i,j)-expY(i,j))+yErrCount;
    end
  end
  yErr(t,1)=yErrCount/Npts;

  progress=strcat('trial number =',num2str(t))
end

yErr_Avg=mean(yErr(:,1));
yErr_Dev=std(yErr(:,1))/sqrt(Ntrials);

for i=1:Ny
  for j=1:Nx
    digData(i,j)=1-digData(i,j);
    Y(i,j)=1-(Y(i,j));
  end
end

yPix=flip(yPix);
set(0,'FixedWidthFontName','Times New Roman');


figure(figNum+1)
surf(xPix,yPix,digData)
```

```
xlabel('x pixels','FontName','FixedWidth','FontWeight','bold','FontSize',14)
ylabel('y pixels','FontName','FixedWidth','FontWeight','bold','FontSize',14)
colormap gray

figure(figNum)
surf(xPix,yPix,Y)
title(strcat('nRate =',num2str(nRate*100),'%, Error=',num2str(yErr_Avg*100),'% +/-
',num2str(yErr_Dev*100),'%'))
xlabel('x pixels','FontName','FixedWidth','FontWeight','bold','FontSize',14)
ylabel('y pixels','FontName','FixedWidth','FontWeight','bold','FontSize',14)
colormap gray
```

Part 2. Bayesian Inference system

This part contains the source code for performing object location via Bayesian inference system in SC-CRAM, results of which are shown in 5.13(b).

```
%Test Bayesian Inference system in SC-CRAM with noise injection

figNum=1;
bitRes=8;
Ngrid=64;
nRate=[0,1,2,4,8,12,16,20,25,30,35,40,45,50];
Ntrials=10;

objPos=[28,29];
sensPos=[0,0;0,32;32,0];
Dcalc=sqrt((objPos(1)^2)+(objPos(2)^2));

Nsets=length(nRate);
Nbits=round(2^bitRes);
Npts=round(Ngrid^2);
xPix=1:1:Ngrid;
yPix=1:1:Ngrid;

Dexp=zeros(3,1);
Bexp=zeros(3,1);
SBval=0.2454;
dxObj=zeros(3,1);
```

```
dyObj=zeros(3,1);

for p=1:3
   dxObj(p)=objPos(1)-sensPos(p,1);
   dyObj(p)=objPos(2)-sensPos(p,2);
   Dexp(p,1)=sqrt((dxObj(p)^2)+(dyObj(p)^2));
   Bexp(p,1)=atan(dyObj(p)/dxObj(p));
end

dx=zeros(Ngrid,Ngrid,3);
dy=zeros(Ngrid,Ngrid,3);
Dvals=zeros(Ngrid,Ngrid,3);
Bvals=zeros(Ngrid,Ngrid,3);
SDvals=zeros(Ngrid,Ngrid,3);
PDvals=zeros(Ngrid,Ngrid,3);
PBvals=zeros(Ngrid,Ngrid,3);

for i=1:Ngrid
   for j=1:Ngrid
      for p=1:3
         dx(j,i,p)=(i-1)-sensPos(p,1);
         dy(j,i,p)=(j-1)-sensPos(p,2);
         Dvals(j,i,p)=sqrt((dx(j,i,p)^2)+(dy(j,i,p)^2));
         SDvals(j,i,p)=5+(Dvals(j,i,p)/10);
         Bvals(j,i,p)=atan(dy(j,i,p)/dx(j,i,p));
         PDvals(j,i,p)=exp(-0.5*(((Dvals(j,i,p)-Dexp(p,1))/sqrt(SDvals(j,i,p)))^2));
         PBvals(j,i,p)=exp(-0.5*(((Bvals(j,i,p)-Bexp(p,1))/sqrt(SBval))^2));
      end
   end
end

Dmeas=zeros(Nsets,Ntrials);
Derr=zeros(Nsets,Ntrials);
xWidth=zeros(Nsets,Ntrials);
yWidth=zeros(Nsets,Ntrials);
Derr_Avg=zeros(Nsets,1);
Derr_Dev=zeros(Nsets,1);
xWidth_Avg=zeros(Nsets,1);
xWidth_Dev=zeros(Nsets,1);
yWidth_Avg=zeros(Nsets,1);
yWidth_Dev=zeros(Nsets,1);
for n=1:Nsets
   for t=1:Ntrials
      Ymeas=zeros(Ngrid,Ngrid);
      for i=1:Ngrid
         for j=1:Ngrid
```

```
        PB=zeros(Nbits,3);
        PD=zeros(Nbits,3);
        M=zeros(Nbits,4);
        Y=zeros(Nbits,1);

        for k=1:Nbits
           rD=[rand,rand,rand];
           rB=[rand,rand,rand];
           rND=[rand,rand,rand];
           rNB=[rand,rand,rand];

           for p=1:3
              if rD(p)<PDvals(j,i,p)
                 PD(k,p)=1;
              end
              if rB(p)<PBvals(j,i,p)
                 PB(k,p)=1;
              end

              if rND(p)<(nRate(n)/100)
                 PD(k,p)=not(PD(k,p));
              end
              if rNB(p)<(nRate(n)/100)
                 PB(k,p)=not(PB(k,p));
              end
           end

           M(k,1)=and(PB(k,1),PD(k,1));
           M(k,2)=and(M(k,1),PB(k,2));
           M(k,3)=and(M(k,2),PD(k,2));
           M(k,4)=and(M(k,3),PB(k,3));
           Y(k,1)=and(M(k,4),PD(k,3));
        end
        Ymeas(j,i)=mean(Y(:,1));
     end
  end

  yMaxArray=zeros(Ngrid,1);
  xIndArray=zeros(Ngrid,1);
  for j=1:Ngrid
     [yMaxArray(j,1),xIndArray(j,1)]=max(Ymeas(j,:));
  end

  [yMax,yInd]=max(yMaxArray);
  xInd=xIndArray(yInd,1);
  maxInd=[xInd,yInd];
```

```
        gXFit=fit(transpose(xPix),transpose(Ymeas(yInd,:)),'gauss1');
        gXCoeff=coeffvalues(gXFit);
        xWidth(n,t)=gXCoeff(3);
        xPos=gXCoeff(2)-1;

        gYFit=fit(transpose(yPix),Ymeas(:,xInd),'gauss1');
        gYCoeff=coeffvalues(gYFit);
        yWidth(n,t)=gYCoeff(3);
        yPos=gYCoeff(2)-1;

        Dmeas(n,t)=sqrt((xPos^2)+(yPos^2));
        Derr(n,t)=abs(Dmeas(n,t)-Dcalc);
    end
    Derr_Avg(n,1)=mean(Derr(n,:));
    Derr_Dev(n,1)=std(Derr(n,:))/sqrt(Ntrials);
    xWidth_Avg(n,1)=mean(xWidth(n,:));
    xWidth_Dev(n,1)=std(xWidth(n,:))/sqrt(Ntrials);
    yWidth_Avg(n,1)=mean(yWidth(n,:));
    yWidth_Dev(n,1)=std(yWidth(n,:))/sqrt(Ntrials);

    progress=strcat('set number=',num2str(n),'/',num2str(Nsets))
end

figure(figNum)
errorbar(nRate,Derr_Avg,Derr_Dev,'ko-')
title('BIS output error')
ylabel('Error')
xlabel('noise injection rate (%)')

nRate=transpose(nRate);
outData=[nRate,Derr_Avg,Derr_Dev,xWidth_Avg,xWidth_Dev,yWidth_Avg,yWidth_D
ev];
```

Part 3. Bayesian belief network

This part contains the source code for heart disaster prediction via Bayesian belief network in SC-CRAM, results of which are shown in 5.13(c).

```
%Test Bayesian Belief network in SC-CRAM with noise injection
```

```
figNum=2;
bitRes=8;
nRate=[0,1,2,4,8,12,16,20,25,30];
Ntrials=100;

P_EDvals=[0.25,0.45,0.55,0.75];
PDval=0.25;
PEval=0.7;
P_BPvals=[0.85,0.2];
P_CPvals=[0.74,0.3];

Nbits=round(2^bitRes);
Nsets=length(nRate);

PHD_ED_Calc=((P_EDvals(1)*PDval+P_EDvals(2)*(1-
PDval))*PEval)+((P_EDvals(3)*PDval+P_EDvals(4)*(1-PDval))*(1-PEval));
PHD_Calc=(P_BPvals(1)*P_CPvals(1)*PHD_ED_Calc)/((P_BPvals(1)*P_CPvals(1)*P
HD_ED_Calc)+(P_BPvals(2)*P_CPvals(2)*(1-PHD_ED_Calc)));
PHD_Meas=zeros(Nsets,Ntrials);
PHD_Err=zeros(Nsets,Ntrials);
PHD_Avg=zeros(1,Nsets);
PHD_Dev=zeros(1,Nsets);

for i=1:Nsets
   for t=1:Ntrials
      P_ED=zeros(Nbits,4);
      PD=zeros(Nbits,2);
      PE=zeros(Nbits,1);
      P_BP=zeros(Nbits,2);
      P_CP=zeros(Nbits,2);

      M_ED=zeros(Nbits,4,3);
      PHD_ED=zeros(Nbits,1);
      M_BP=zeros(Nbits,2,2);

      Q=zeros(Nbits,1);
      J=zeros(Nbits,1);
      K=zeros(Nbits,2);
      PHD=zeros(Nbits,1);

      for k=1:Nbits
         rED=[rand,rand,rand,rand];
         rD=[rand,rand];
         rE=rand;
         rBP=[rand,rand];
```

```
rCP=[rand,rand];
rN=zeros(11,1);

for n=1:11
   rN(n)=rand;
end

for p=1:4
   if rED(p)<P_EDvals(p)
      P_ED(k,p)=1;
   end

   if rN(p)<(nRate(i)/100)
      P_ED(k,p)=not(P_ED(k,p));
   end
end

for p=1:2
   if rBP(p)<P_BPvals(p)
      P_BP(k,p)=1;
   end

   if rN(p+4)<(nRate(i)/100)
      P_BP(k,p)=not(P_BP(k,p));
   end

   if rCP(p)<P_CPvals(p)
      P_CP(k,p)=1;
   end

   if rN(p+6)<(nRate(i)/100)
      P_CP(k,p)=not(P_CP(k,p));
   end

   if rD(p)<PDval
      PD(k,p)=1;
   end

   if rN(p+8)<(nRate(i)/100)
      PD(k,p)=not(PD(k,p));
   end
end

if rE<PEval
   PE(k)=1;
end
```

```
        if rN(11)<(nRate(i)/100)
           PE(k)=not(PE(k));
        end

        M_ED(k,1,1)=and(P_ED(k,1),PD(k,1));
        M_ED(k,2,1)=and(P_ED(k,2),not(PD(k,1)));
        M_ED(k,3,1)=and(P_ED(k,3),PD(k,2));
        M_ED(k,4,1)=and(P_ED(k,4),not(PD(k,2)));

        M_ED(k,1,2)=or(M_ED(k,1,1),M_ED(k,2,1));
        M_ED(k,2,2)=or(M_ED(k,3,1),M_ED(k,4,1));

        M_ED(k,1,3)=and(M_ED(k,1,2),PE(k));
        M_ED(k,2,3)=and(M_ED(k,2,2),not(PE(k)));
        PHD_ED(k)=or(M_ED(k,1,3),M_ED(k,2,3));

        M_BP(k,1,1)=and(P_BP(k,1),P_CP(k,1));
        M_BP(k,2,1)=and(P_BP(k,2),P_CP(k,2));
        M_BP(k,1,2)=and(M_BP(k,1,1),PHD_ED(k));
        M_BP(k,2,2)=and(M_BP(k,2,1),not(PHD_ED(k)));

        if k>1
           Q(k)=PHD(k-1);
        end

        J(k,1)=not(and(M_BP(k,1,2),not(Q(k))));
        K(k,1)=not(and(M_BP(k,2,2),Q(k)));
        K(k,2)=not(and(K(k,1),Q(k)));
        PHD(k)=not(and(K(k,2),J(k,1)));
      end

      PHD_Meas(i,t)=mean(PHD(:,1));
      PHD_Err(i,t)=abs(PHD_Calc-PHD_Meas(i,t));
   end
   PHD_Avg(1,i)=mean(PHD_Err(i,:));
   PHD_Dev(1,i)=std(PHD_Err(i,:))/sqrt(Ntrials);
end

figure(figNum)
errorbar(nRate,PHD_Avg,PHD_Dev,'ko-')
title('BBN output error')
ylabel('Error')
xlabel('noise injection rate (%)')

nRate=transpose(nRate);
```

```
PHD_Avg=transpose(PHD_Avg);
PHD_Dev=transpose(PHD_Dev);
outData=[nRate,PHD_Avg,PHD_Dev];
```

Appendix C: List of Publications

Research articles:

[1] H. Yun, D. Lyu, Y. Lv, **B. R. Zink**, P. Khanal, B. Zhou, W. Wang, J.-P. Wang, and K. A. Mkhoyan, *In-situ TEM and Spectroscopy Studies of Nanoscale Perpendicular Magnetic Tunnel Junction*, Microscopy and Microanalysis, **28**, 838 – 840 (2022).

[2] H. Cilasun, S. Resch, Z. I. Chowdhury, M. Zabihi, Y. Lv, **B. R. Zink**, J.-P. Wang, S. S. Sapatnekar, and U. R. Karpuzcu, *Error Detection and Correction for Processing in Memory (PiM)*, ArXiv: 2207.13261 (2022).

[3] **B. R. Zink**, D. Zhang, H. Li, O. J. Benally, Y. Lv, D. Lyu, and J.-P. Wang, *Ultralow Current Switching of Synthetic-Antiferromagnetic Magnetic Tunnel Junctions Via Electric-Field Assisted by Spin-Orbit Torque*, Adv. Electron. Mater., 2200382 (2022).

[4] D. Lyu, P. Khanal, Y. Lv, B. Zhou, H. Yun, Q. Jia, **B. R. Zink**, Y. Fan, K. A. Mkhoyan, W. Wang, and J.-P. Wang, *Sub-ns Switching and Cryogenic-Temperature Performance of Mo-Based Perpendicular Magnetic Tunnel Junctions*, IEEE Electron Device Lett., **43**, 1215 – 1218 (2022).

[5] T. J. Peterson, A. Hurben, W. Jiang, D. Zhang, **B. R. Zink**, Y.-C. Chen, Y. Fan, T. Low, and J.-P. Wang, *Enhancement of voltage controlled magnetic anisotropy (VCMA) through electron depletion*, IEEE Electron Device Lett., **131**, 153904 (2022).

[6] Y. Lv, **B. R. Zink**, and J.-P. Wang, *Bipolar Random Spike and Bipolar Random Number Generation by Two Magnetic Tunnel Junctions*, IEEE Trans. Electron Devices, **69**, 1582 – 1587 (2022).

[7] **B. R. Zink** and J.-P. Wang, *Influence of Intrinsic Thermal Stability on Switching Rate and Tunability of Dual-Biased Magnetic Tunnel Junctions for Probabilistic Bits*, IEEE Magn. Lett., **12**, 4501405 (2021).

[8] **B. R. Zink**, Y. Lv, and J.-P. Wang, *Influence of size and shape anisotropy on key performance metrics in spin-torque oscillators*, AIP Adv., **11**, 025215 (2021).

[9] **B. R. Zink**, J. Yang-Scharlotta, F. Mancoff, J. Sun, M. Han, and J.-P. Wang, *Influence of Ionizing Dose on Magnetic Tunnel Junctions With Perpendicular Anisotropy*, IEEE Trans. Nucl. Sci., **68**, 748 – 755 (2021).

[10] **B. R. Zink**, Y. Lv, and J.-P. Wang, *Independent Control of Antiparallel- and Parallel-State Thermal Stability Factors in Magnetic Tunnel Junctions for Telegraphic Signals With Two Degrees of Tunability*, IEEE Trans. Electron Devices, **66**, 5353 – 5359 (2019).

[11] **B. R. Zink**, Y. Lv, and J.-P. Wang, *Telegraphic switching signals by magnetic tunnel junctions for neural spiking signals with high information capacity*, J. Appl. Phys., **124**, 152121 (2018).

Patents:

[1] **B. R. Zink**, Y. Lv, and J.-P. Wang, *Stochastic Computing Using Logic-Memory Cells*, US Patent, Submitted Aug. 24th, 2022.