Multilevel Resistance Programming in

Conductive Bridge Resistive Memory

by

Debayan Mahalanabis

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2015 by the
Graduate Supervisory Committee:

Hugh Barnaby, Chair
Michael Kozicki
Sarma Vrudhula
Shimeng Yu

ARIZONA STATE UNIVERSITY

December 2015

ABSTRACT

This work focuses on the existence of multiple resistance states in a type of emerging non-volatile resistive memory device known commonly as Programmable Metallization Cell (PMC) or Conductive Bridge Random Access Memory (CBRAM), which can be important for applications such as multi-bit memory as well as non-volatile logic and neuromorphic computing. First, experimental data from small signal, quasi-static and pulsed mode electrical characterization of such devices are presented which clearly demonstrate the inherent multi-level resistance programmability property in CBRAM devices. A physics based analytical CBRAM compact model is then presented which simulates the ion-transport dynamics and filamentary growth mechanism that causes resistance change in such devices. Simulation results from the model are fitted to experimental dynamic resistance switching characteristics. The model designed using Verilog-a language is computation-efficient and can be integrated with industry standard circuit simulation tools for design and analysis of hybrid circuits involving both CMOS and CBRAM devices. Three main circuit applications for CBRAM devices are explored in this work. Firstly, the susceptibility of CBRAM memory arrays to single event induced upsets is analyzed via compact model simulation and experimental heavy ion testing data that show possibility of both high resistance to low resistance and low resistance to high resistance transitions due to ion strikes. Next, a non-volatile sense amplifier based flip-flop architecture is proposed which can help make leakage power consumption negligible by allowing complete shutdown of power supply while retaining its output data in CBRAM devices. Reliability and energy consumption of the flip-flop circuit for different CBRAM low resistance levels and supply voltage values are analyzed and compared to CMOS

designs. Possible extension of this architecture for threshold logic function computation using the CBRAM devices as re-configurable resistive weights is also discussed. Lastly, Spike timing dependent plasticity (STDP) based gradual resistance change behavior in CBRAM device fabricated in back-end-of-line on a CMOS die containing integrate and fire CMOS neuron circuits is demonstrated for the first time which indicates the feasibility of using CBRAM devices as electronic synapses in spiking neural network hardware implementations for non-Boolean neuromorphic computing.

DEDICATION

To my parents Saumen Mahalanabis and Bhaswati Mahalanabis, and to my brother Satyaki Mahalanabis

ACKNOWLEDGMENTS

I would first like to express my deepest gratitude and appreciation to my committee members - my advisor Dr. Hugh J. Barnaby for his constant encouragement, guidance and support of my work, and Dr. Michael N. Kozicki, Dr. Sarma Vrudhula and Dr. Shimeng Yu for providing invaluable guidance and key insights.

Research is always collaborative in nature and I would like to thank my research colleagues Adnan Mahmud, Wenhao Chen and Pradeep Dandamudi for their technical collaboration and for fabricating the devices whose characterization formed a key part of my work. I would also like to thank and acknowledge Dr. Yago Gonzalez Velo, Vineeth Bharadwaj, Mahraj Sivaraj, Saba Rajabi, Jinghua Yang, Mehdi Saremi and Jennifer Taggart for their collaboration and helpful technical discussions. I also want to thank my other research colleagues Weijie Yu, Mehmet Balaban, Ninad chamele and Runchen Fang.

I would also like to thank the following sponsors for funding my doctoral research - National Science Foundation (NSF) through the Partnerships for Innovation-Building Innovation Capacity under Award 1237856, through the Fundamental Research Program under Award 1230401 and through the NSF Office of the Director under Grant 0856090, the Defense Threat Reduction Agency (DTRA) under Grant HDTRA1-11-1-0055, and the U.S. Air Force Research Laboratory (AFRL) , Arlington, VA, USA, under Grant FA9453-13-1-0288. Last but not least, I would like to thank my parents and my brother for their love and support throughout my academic career and all my friends for their encouragement and companionship.

TABLE OF CONTENTS

Page

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER

1 CONDUCTIVE BRIDGE RESISTIVE MEMORY TECHNOLOGY 1

1.1 Overview of Emerging Non-Volatile Memories..1

1.2 Conductive Bridge Resistive Memory: Theory 5

1.3 CBRAM Programming Dynamics 8

2 EXPERIMENTAL DEVICE CHARACTERIZATION.. 12

2.1 Impedance Spectroscopy Analysis 12

2.2 Quasi-static Resistance Switching Characterization 24

2.3 Pulsed Mode (Transient) Characterization 28

3 CBRAM COMPACT MODEL FOR CIRCUIT SIMULATIONS........................ 33

3.1 Resistance Estimation from Filament Dimensions 33

3.2 Compact Model Transient Simulation...34

3.3 Model Fitting to Experimental Data 38

4 SINGLE EVENT EFFECTS IN CBRAM MEMORY ARRAYS 45

4.1 Single Event Effects Overview..45

4.2 Radiation Effects In CBRAM Memory..49

4.3 CBRAM Heavy Ion Testing: Experimental Results...................................53

4.4 Voltage Dependent CBRAM Resistance Change.......................................58

CHAPTER Page

4.5 Single Event Effects Simulation: CBRAM 1T-1R Memory 60

4.6 Single Event Effects Simulation: CBRAM Crossbar Array 71

4.7 Conclusions & Future Work 75

5 NON-VOLATILE FLIP-FLOP ARCHITECTURE USING CBRAMS……......... 77

5.1 Background & Motivation 77

5.2 Non-Volatile Sense Amplifier Flip-flop Circuit Architecture 81

5.3 Flip-Flop Energy & Reliability Estimation 88

5.4 Potential Application for Reconfigurable Threshold Logic Gate 93

6 SPIKE TIMING DEPENDENT PLASTICITY IN CBRAM DEVICES…….......104

6.1 Neuromorphic computing: Overview 104

6.2 Synaptic Plasticity in Neural Networks 106

6.3 Spike Timing Dependent Plasticity (STDP) of Synapses 107

6.4 Resistive Memory as Electronic Synapses: Previous Work 110

6.5 STDP & Associative Memory Simulation Using CBRAM Model..........111

6.6 Neuron Hardware Implementation 116

6.7 Silicon Neuron Post-Processing for CBRAM Synapse Integration 118

6.8 Post-Silicon Validation & On-Chip STDP Demonstration 119

6.9 Conclusions & Future Work 123

7 SUMMARY & CONCLUSIONS……………………………………...…....... 125

REFERENCES...128

APPENDIX Page

A VERILOG-A CBRAM COMPACT MODEL 141

BIOGRAPHICAL SKETCH……………………………...………… ……………..…... 146

LIST OF TABLES

Table Page

3.1 CBRAM Compact Model Simulation Parameters .... ………….44

5.1 Truth Table for Threshold Function $f = ab + bc + ca$ 100

LIST OF FIGURES

Figure Page

1.1 (a) CBRAM DC I-V Characteristics With a 10 µA Compliance Current Which Shows Resistive Switching Between High Resistance State (HRS) and Low Resistance State (LRS). (b) Experimentally Obtained Inverse Power Law Dependence of LRS CBRAM Resistance on Compliance (Programming) Current Used During DC Characterization...…………………………………………………………….... 7

1.2 (a) Illustration of CBRAM Device Cross-section With Ag Anode, Ni Cathode and Chalcogenide ($Ge_{30}Se_{70}$) Via Layer. (b) Micro-photograph Showing Top View of a CBRAM Device With Area 500 µm x 500 µm ,….…………….………... 8

2.1 Cole-Cole Plot Representation of Complex Impedance Spectra of a Parallel RC Element (R_1-C_1)………………………………………………………… 13

2.2 (a) TEM Image of a CBRAM Device (HRS Condition) Cross-section Using High Angular Dark Field. (b) Magnified View of the Same Cross-section Near the Chalcogenide Via- Ni Cathode Interface, Indicating Presence of Two Layers (One Lighter and Another Darker) Within the Chalcogenide (ChG) More Clearly. ……………………………………………………………... 16

2.3 (a) Illustration of Atomic Profiles of Ge (Red), Se(Green), Ag(Dark Blue) and Ni(Light Blue) Along the CBRAM Device, Showing the Sharp Change in Ag Count Near the Cathode Side, Indicating Existence of a Ag Gradient As

Figure Page

Modeled in the Impedance Spectra Equivalent Circuit. Ag- poor Chalcogenide Layer Thickness Can Be estimated to Be Approximately 5 nm Based on the Obtained Profile. …………………………………………………………. 17

2.4 (a) HRS Impedance Spectra of a CBRAM Device (100 µm by 100 µm Diameter) Showing Both Measured Data and Fitted Characteristics. (b) Imaginary Component vs. Real Component of Electric Modulus (M (ω)) Derived from Complex Impedance Data in (a). ………………………………. 18

2.5 (a) Photo-doped HRS CBRAM RC Equivalent Circuit With Two Regions With Distinct RC Characteristics. (b) Idealized Complex Plane Representation of Real and Imaginary Components of Impedance (Z(ω)) (top) and Electric Modulus (M(ω)) (bottom) Associated With circuit With Two Parallel R-C Elements…………………………………………………………………..….... 19

2.6 Cole-Cole Plots Showing Scaling of HRS CBRAM Resistance Parameters R_1 and R_2 With Via Area. Both R_1 and R_2 Vary Inversely With Via Area…..… 20

2.7 Scaling of HRS CBRAM Capacitance Parameters Showing Non-linear Scaling of HRS CBRAM Capacitance Parameter C_1 (a) With Area and Linear Scaling of HRS CBRAM Capacitance Parameter C_2 (b) With Area……............... 21

2.8 (a) LRS CBRAM Equivalent Circuit. (b) LRS Impedance Spectra of a CBRAM Device (100 µm Via Diameter) Showing Both Measured Data and Fitted Characteristics, Indicating Single RC Element.. 22

Figure Page

2.9 (a) LRS Resistance Parameter R_{on} vs. Area. (b) LRS Capacitance Parameter C_{on} vs. Area…......………………………………….…………. ……………. 23

2.10 (a) Capacitance vs. Area Showing Close Match Between C_{on} and Series Combination of C_1 and C_2. ……………………………….....…………….. 24

2.11 (a) Measured DC I-V Characteristics of a CBRAM Device With a 10 µA Compliance Current, With the Corresponding Resistance vs. Voltage Plot (Semi-log Scale) Shown As Inset. (b) Another Typical I-V Characteristics With Some Key Parameters Identified Such As On Switching threshold, Set Programming Current (Compliance Current), Peak Reset Current, High Resistance State (HRS) Region and Low Resistance State (LRS) Region.………………………………………….....………….....…. 25

2.12 (a) Experimental Quasi-static I-V Switching Characteristics at 100µA Compliance Current For 20 Cycles, (b) I-V Characteristics For Positive Voltage Double Sweep Between 0 V and 1 V at Different Compliance Currents (CC). Arrow Indicates Direction in Which LRS Decreases. (c).Experimental Dependence of LRS Resistance on Compliance Current (Log-log Plot) During DC Voltage Sweep. (d) Experimental Dependence of Peak Erase Current on Compliance Current (Log-Log plot) During DC Voltage sweep…..……...…………………………………………..………. 26

2.13 Incremental Resistance Change in CBRAM Under DC Bias: (a) Resistance Decrease Under Positive DC Bias With Repeated Sweeps While Increasing

Figure Page

Stop Voltage (V_s) From 65 mV to 120 mV. Arrow Indicates Direction in Which Resistance Gradually Decreases. (b) Resistance Increase Under Negative Bias With Repeated Sweeps While Increasing Stop Voltage (Vs) From -0.5 V to -1.5 V. Arrow Indicates Direction in Which Resistance Gradually Increases ……………………………………………………………… 27

2.14 Transient Voltage Pulse Based PMC Programming:(a) Experimental Test Setup (Measured PMC Resistance During Read Operation = ($V_1 – V_2$)/I); (b) Typical PMC Write and Read Voltage Transients Showing Applied Voltage and Measured Output Voltage; (c) Typical PMC Erase and Read Voltage Transients Showing Applied Voltage and Measured Output Voltage. ……………………………………………………………………………… 29

2.15 (a) Transient Write-Erase Operation Performed on CBRAM in Series With Current Limiting Resistor (Experimental Setup Shown in Inset). Each Write and Erase Pulse is Followed By a Low Magnitude Read Pulse. (b) Experimental Data Showing Multiple LRS For 4 Different Applied Voltage Pulse Magnitudes. (c) LRS Obtained From Data of (b) vs. Applied Pulsed Voltage Amplitude. …………………………………………………… 30

2.16 Incremental Resistance Programming of a CBRAM Device With 100 μs, 1.5 V Write Pulses and 200 μs, -1.5 V Erase Pulses, Each Pulse followed By 100 mV Read Pulses………………………………………………………………… 31

2.17 Incremental Change in Resistance Depends on the Width of Applied Voltage

Figure Page

Pulses. (a) Applied Sequence of Programming Voltage Pulses of Width tP.
(b) Resistance vs. Applied Voltage Pulse Magnitude For tP = 0.1 ms, 0.5 ms
and 1 ms. . ………………………………………………………………... 32

3.1 Simulated Voltage Double Sweep (from -1 V to +1V and Back to -1V) to
Generate I-V Characteristics of CBRAM Model Showing Applied Voltage
and Corresponding Change in the Filament Height and Radius…...................… 36

3.2 Simulated Pulsed Voltage Write-Read-Erase-Read Sequence For a 1T-1R
Simulation Using CBRAM Compact Model. During Write, Vertical Filament
Growth Occurs, Followed By Increase in Filament Radius, LRS Value Being
Determined Based on the Radius. No Change in Filament Dimensions Occur
During Read. During Erase, Filament Radius Decreases First Followed By
Decrease in Filament Height, Thus Causing LRS to HRS Transition……..… 38

3.3 Comparison of Simulation and Experimental Results- (a) I-V Characteristics
at I_{comp} = 100μA, (b) LRS Dependence on Compliance Current, (c) Peak
Reset Current Dependence on Compliance Current and (d) LRS Dependence
on Applied Voltage Pulse During Transient Characterization and the
Simulated Filament Radius……………….................................…………… … 39

3.4 1T-1R CBRAM Device With Ag Anode, GeS_2 Via Layer and W Cathode,
(b) Simulated I-V Characteristics During Anode Voltage Sweep For the 1T-1R
CBRAM With Source Grounded For Different Fixed Gate Voltages.........….. 40

Figure | Page

3.5 Experimental Resistance Cycling of 1T-1R Element During 100 Write-Erase Operations Performed By Sweeping the Anode Voltage……….…................ 41

3.6 Comparison of CBRAM Compact Model Simulation Results and Corresponding Experimental Data: (a) Dependence of On Switching Threshold Voltage (During HRS to LRS Transition) on the Applied Voltage Ramp Rate, (b) Multilevel LRS in 1T-1R CBRAM During Anode Voltage Sweep For Different NMOS Gate Voltages, (c) Programming Currents Corresponding to the Different Applied Gate Voltages For Each LRS Level Shown in (b), and (d) Peak Erase Currents in 1T-1R CBRAM During Erase Operation Voltage Sweep After Programming at Different Gate Voltages. ………………………………………………………………….. 42

3.7 Simulated and Experimentally Obtained Anode Voltage and Current Transients of a Two Terminal CBRAM Device During a Typical Pulsed Write Erase Operation With Cathode Grounded. …………...…...…...……….. 43

4.1 (a) Schematic Illustration of Processes Typically Associated With Single Event Effects Using Example of a CMS Inverter. (b) SEE Induced Current Transient Showing Contributions From the Various Charge Collection Processes………………………………………………………..…… 49

4.2 Endurance of CBRAM Devices Exposed to Co-60 Gamma-rays. HRS and LRS Resistance As a Function of the Number of Write-Read-Erase-Read Cycles For a Control CBRAM and a CBRAM Exposed to 4.62 MRad …...…. 51

Figure Page

4.3 (a) 1T-1R Array Architecture With Word Lines (WL) Driving Access NMOS Gates in the Same Row and Elements in Same Column Sharing the Same Bit and Select Lines. (b) Bias Conditions in an Inactive 1T-1R Element (Circled in (a)) Susceptible to Single Event Induced Upsets During Write, Read and Erase Operation Being Performed on Another Cell in the Same Column………………………………………………………………………. 52

4.4 Experimental Results Showing Mean Resistance of 1T-1R Test Array Elements After Ion Strike vs. Incident Ion LET. HRS CBRAMs Undergo Decrease in Resistance with Increasing LET. Also Shown is the Corresponding SEU Error Cross-section vs. LET. It Can Be Seen that LRS CBRAMs Subjected to Ion Strike Didnot Revert to HRS ....……...…... 55

4.5 (a) 1T-1R Array Used for Heavy Ion Testing. (b) Smartboard Designed at Aerospace For Electrical Characterization and Biased Heavy Ion Testing..…... 56

4.6 Plot Showing Resistance Ratio Parameter Defined by (4.2) For (a) Each HRS Programmed Exposed Device, and (b) Each HRS Programmed Control (Unexposed) Device. Devices With a Ratio of 10 or More are Considered to Have Experienced Upset in the Form of HRS to LRS Transition …………...... 57

4.7 Plots Showing Resistance Ratio Parameter Defined by (4.3) for (a) Each LRS Programmed Exposed Device, and (b) Each LRS Programmed Control (Unexposed) Device. Devices with a Ratio of 10 or More are Considered to Have Experienced upset in the Form of LRS to HRS Transition. ………......….58

Figure Page

4.8 Contour Plot Showing Simulated Change in CBRAM Resistance (ΔR_{CF}) vs. Magnitude and Duration of a Voltage Pulse Applied Across the CBRAM. Simulation was Performed on a CBRAM Initially in HRS. A Higher Value of ΔR_{CF} Increases Likelihood of a Transition to LRS. ………..…………….. 59

4.9 (a) Ion Strike Induced Single Event on a HRS Element (Word Line 2) in 1T-1R Array While Another Element in same row (Word Line 0) is Being Written, and (b) Equivalent Circuit For the HRS Cell (word line 2) to Model the Single event Induced Resistance Change. ……………………………..…... 63

4.10 (a) A Simulated Voltage Transient Generated Across 1T-1R CBRAM and Corresponding Change in CBRAM Resistance For LET = 8MeV.cm^2/mg and (b) Simulated Results Showing 1T-1R CBRAM Resistance After Ion Strike Event vs. Incident Ion LET, With Threshold LET ≈ 7.5 MeV.cm^2/mg For HRS to LRS Transition Indicated, Assuming a Minimum Acceptable HRS to LRS Ratio of 10. ………………………………....…………...….. 64

4.11 (a) Ion Strike Induced Single Event on a LRS Element (Word Line 2) in 1T-1R Array While Another Element in Same Row (Word Line 0) is Being Erased and (b) Equivalent Circuit For the Encircled LRS Cell to Model the Single Event Induced Resistance Change. ……………………………………. 65

4.12 Simulated TCAD Structure with Fixed Resistance Representing CBRAM,

Figure Page

and Corresponding Circuit Simulation Using Photocurrent (I_{ph}) From TCAD..………………………………………………………………...….. 66

4.13 (a) Generated Photocurrent Transients For Cases I and II For the Same Incident Ion LET (25 MeV.cm^2/mg). (b) Comparison of Photocurrents Generated Under Write and Read Bias (Case I) For the Same LET (25 MeV.cm^2/mg). (c) 1T-1R Circuit Simulation Using CBRAM Compact Model With the Photocurrents From (b). CBRAM Undergoes HRS to LRS Transition under Write Bias But Not Under Read Bias. ….....….. 68

4.14 (a) Generated Photocurrent Transients For Cases I and II For the Same Incident Ion LET (20 MeV.cm^2/mg) Under Erase Bias. (b) 1T-1R Circuit Simulation Using CBRAM Compact Model With the Photocurrents For Angular Incident Strike (Case II) from (a). CBRAM Undergoes LRS to HRS Transition Due to Cathode Voltage Being Pulled Up Sufficiently High to Generate a Net Negative Voltage Across the Device ……………………..…… 70

4.15 (a) Simulated CBRAM Resistance After Ion Strike vs. Ion LET Under Write and Read Bias Conditions. Initially the Device is in High Resistance State (HRS) (b) Simulated CBRAM Resistance After Ion Strike vs. Ion LET Under Erase Bias condition For Different Angles of Incidence of Strike (Case I and II). Initially Device is in Low Resistance State (LRS).................. 71

4.16 A CBRAM Crossbar Array Architecture With Word line Driving the Anodes and Bit Line driving the Cathodes. The Voltages V_{WLNS} and V_{BLNS}

Figure Page

Represent the Voltages Applied at Unselected Cells (Dashed Line) to Prevent Read Disturb While Accessing a Cell (Solid Circle)………..…...…… 73

4.17 (a) Example of Single Event Transient Propagating Along the Word Line of a Crossbar Array Due to Drain of 'Off' PMOS of Word Line Driver Being Struck By Ion. Parasitic RC of the Connecting Wires is Considered As Shown. (b)Degradation of a Single Event Generated Voltage Transient As it Propagates Along a Word Line. The 256th Element Experiences a Degraded Voltage Spike Compared to the First element, Where the spike originates …...………...…….74

4.18 (a) No. of Simulated HRS to LRS Upsets For Both 1/2V and 1/3 V Bias Schemes Along a Word Line in a 256x256 Crossbar Array. (b) No. of Simulated LRS to HRS Upsets For Both 1/2V and 1/3 V Bias Schemes Along a Bit Line in a 256x256 Crossbar Array.……………………..………….............75

5.1 (a) Conventional Power Gating Scheme in CMOS Logic Using High-V_{th} Retention Latches. Subthreshold Leakage is Partially Suppressed. (b) Alternative Scheme Employing Flip-flops with Non-volatile Resistive Memory Elements Which Can Store Data in Absence of Power Supply. Thus By Turning Off Power Supply Completely, Leakage Can Be Completely Suppressed. ………………………………………………...…..… 78

5.2 (a) General Flip-flop Topology With Pulse Generator Followed by Slave Latch. (b) Conventional Sense Amplifier Flip-flop With NAND Gate Based Slave SR Latch …………………………………………...……………….. 82

Figure Page

5.3 (a) Non-volatile Flip flop Architecture using CBRAM Devices With Write and Read Circuitry (FF I). (b) Non-volatile Flip-flop Architecture Using CBRAM Devices With Same Write Circuit As FF I But a Different Read Circuit Implementation (FF II). Notations An, Ca Denote the Anode and Cathode Terminals Respectively For the CBRAM Devices A and B. ….... 84

5.4 Comparison of Static Noise Margins During Read Operation For FF I and FF II vs. Supply Voltage (VDD) For HRS/LRS Ratio of 150 and 5. FF II Shows Superior Noise Margin and Hence Greater Tolerance to Signal Noise.…………….…………………………………………….…. 86

5.5 A Timing Sequence of Different Control Signals and Associated Phases of Operation of the Proposed Non-volatile Flip-flop. ………….……........……87

5.6 Non-volatile Flip-flop (FF II) Simulation Results Using CBRAM Compact Model at Nominal VDD = 1.5V. Normal Data Sensing Followed By Data Write and Read Operations Shown For First Data D = ‘1’ and then D = ‘0’. Read Pulse Duration is 2ns. For Storing, CBRAM A is Programmed to LRS to Store Data D= ‘1’ and CBRAM B Stays in HRS. Subsequently CBRAM A is Erased Back to HRS and CBRAM B is Programmed to LRS For Storing D = ‘0’. ……………………………………..……………….…. 88

5.7 Read Failure Occurrence Probability (Log Scale) vs. Supply Voltage Scaling For Different HRS/LRS Ratios. Inset Shows Minimum Read Pulse Durations For Successful Read Operation vs. Power Supply Voltage (VDD)……......….. 90

Figure Page

5.8 Monte Carlo Simulations of Read Operation Assuming Gaussian Distributions of CBRAM LRS and HRS Levels. …………..…………...…….. 91

5.9 Simulated Total Energy Consumption vs. VDD During Normal Data Sensing, Precharge, Read and Write Phases For Write Operations Performed For Different LRS Values. Inset Shows Dependence of Energy Consumption During Write Operation For Different LRS Levels. ….................................… 92

5.10 Estimation of Minimum Sleep Period in Order For Proposed Non-volatile Flip-flop to Be More Energy Efficient Than a Custom Low Leakage CMOS Latch (Black) [88]. Inset Plot Compares the Total Energy Consumed For the Proposed Design (Black) and Non-volatile Flip-flop From [81] (Red) For VDD= 0.8 V and VDD = 0.4 V. ……………………………………….......... 93

5.11 A General Threshold Logic Gate Architecture With Sense Amplifier to Detect Difference in Impedance Between Two Sets of Input Transistors (N_1-N_n and N_{n+1}-N_{2n+1}) Referred to As Input Network. …...………................... 96

5.12 Circuit Implementation of an N-Input Reconfigurable Threshold Logic Gate Using CBRAMs As Programmable Weights. Two Phases of Operation Exist – Precharge and Evaluation, The CBRAMs Are Programmed to Different LRS Values to Implement Input Weight. ……...……………………..…………… 98

5.13 Transient Simulation Showing Input, Output and Control Signals of the Hybrid Threshold Logic Gate of Fig. 5.12 During Evaluation Operation For the Threshold Function $f = ab + bc + ca$. The Circuit Evaluates Every

Figure Page

Time Clock (Clk) is High and Prog is Low. Both f and Its Complement is Implemented Here At *Out* and $\overline{Out}$ Respectively By Programming the Resistance of All the CBRAM Devices to 60 KΩ (Corresponding to an Integer Weight of 1)……………………………………………………..…… 101

5.14 Monte-Carlo Simulation Comparing the Number of Failures (Incorrect Evaluations) Occurring in a CMOS Only Implementation vs. Those Occurring in a Hybrid Implementation of a 5 Input Majority Logic Function As Power Supply (VDD) is Lowered. CMOS Design is Found to Fail At Higher Values of VDD Than the Hybrid Implementation. …..…………………..……….. 102

6.1 (a) Schematic View of a Perceptron, Which Forms the Elementary Unit of an Artificial Neural Network. (b) Schematic of the Biological Counterpart of Perceptron - a Neuron With It's Dendrites (Analogous to Perceptron Inputs), Synaptic Pathways (Analogous to Input Weights) and Axon (Analogous to Perceptron Output). ………………………………….....……....………….. 107

6.2 Two Neurons Connected Via Synaptic Pathway Form the Fundamental Computational Unit of Biological Neural Networks, Reproduced From [97]. (b) General Shape of Neuron Spike (Action Potential) Modeled From [97]. (c) Experimental Data (Points) for Biological STDP Learning Function and Corresponding Analytical Approximation (Solid Line). ……………...…….. 110

6.3 Simulated STDP Behavior with CBRAM Compact Model Plotting Percentage Change in Conductance vs. Relative Spike Timing Difference For Three Different Initial LRS Values. ……………………………...……….……. 114

Figure Page

6.4 Block Diagram of the Simulated Three Neuron Network Block Diagram Using CBRAM Synapses. …………………………………………..………….. 115

6.5 Transient Simulation Results of the Network of Fig. 6.4 Showing How Modification of CBRAM Resistances Based on Input and Output Spike Timing Can Be Used for Associative Learning. …………………………….....115

6.6 (a) Leaky Integrate-and-fire CMOS Neuron Circuit Based on [109] Used in This Work. (b) Sub-blocks Used for The Neuron Circuit - a Transconductance Amplifier For the Comparator (Left) and a Constant-g_m Bias Circuit to Generate the Necessary Voltage Bias Levels (Right). ..……......118

6.7 (a) Micro-Photograph of Wire-bonded Die After Post-processing Containing CMOS Neurons and BEOL Fabricated CBRAM Devices. (b) Schematic of BEOL On-chip CBRAM Cross-section.. …...………………...………..….... 119

6.8 Test Setup For STDP Demonstration Consisting of Two On-chip CMOS Neurons Connected Via a Single CBRAM Synapse (Also Fabricated On-chip in BEOL Processing). ……………... 121

6.9 (a) Sample Spike Output Waveforms From On-chip Pre and Post Neuron Using the Test Setup of Fig.6.. …………………..……………..………. 122

6.10 Experimentally Measured STDP Behavior With On-chip CBRAM Device, Showing Change in CBRAM Conductance vs. Spike Timing Difference, With Pre and Post Neuron Spike Widths of (a) 7 µs and (b) 15 µs. Three Different Initial LRS Values are Used to Exploit The Multi-level Programmability in CBRAM Device………………………………...….... 123

CHAPTER 1

CONDUCTIVE BRIDGE RESISTIVE MEMORY TECHNOLOGY

1.1 Overview of Emerging Non-Volatile Memory Devices

Information storage has become increasingly important in today's microelectronics industry. Demand for greater data storage capabilities while limiting power consumption drives the consumer electronics market. Moreover, leakage power consumption has become the dominant component of power consumption with shrinking device sizes and a critical design consideration for embedded battery-driven electronic applications. Today, the various dominant memory technologies offer different advantages and disadvantages. The Static Random Access Memory (SRAM) offers high speed memory access and is used typically as small on-chip cache memories due to larger area. Dynamic Random Access Memory (DRAM) on the other hand is more suitable for high density main memory due to more compact area and high endurance but is prone to leakage and requires periodic data refresh, thus increasing power consumption. Unlike the volatile SRAM and DRAM technologies, Flash is a non-volatile memory (NVM) technology and has shown promise for scaling down to the 22 nm node. However, disadvantages of flash memory include low endurance, high programming power consumption and relatively slow program speed compared to SRAM and DRAM. Modern computers accommodate the trade-offs associated with different types of memory through hierarchical memory storage. However, the need for a new memory technology that can to bridge the gap between NVM and RAM technologies by providing low power operation with virtually zero leakage, non-volatile data storage and scalability for high density memory that can be easily integrated with standard CMOS processes has become more and more pressing.

This need for alternative non-volatile memory technologies has led to increased research focus on emerging memory devices which can combine the best features of the existing memory types and offer scalability through three dimensional vertical integration beyond the lateral scaling limits of traditional memory technologies.

The various types of emerging memory can be divided into the following major groups-

1) Ferroelectric memories: They are based on the property of reversible electrical polarization under the influence of a sufficiently strong electric field in ferroelectric materials such as $SrBi_2Ta_2O_9$ (SBT) and Pb $(Zr_x, Ti_{1-x})O_3$ (PZT) [1], [2]. The polarization is bi-stable in nature with respect to the electric field. This bi-stability in the polarization forms the basis of Ferroelectric random access memory (FRAM) technology. Although FRAM memory offers low-power and high speed operation and good endurance, it has suffered setbacks due to scalability issues and difficulties in integration with standard processes.

2) Magneto-Resistive memories: Magneto Resistive random access memory (MRAM) technology consists of two layers of ferromagnetic material are separated by a thin insulating layer. While one layer has fixed magnetization (known as the 'fixed' layer), the other layer (called the 'free' layer) is able to change its magnetization in response to external electric field [3]. When current flows through the device, the free layer changes its magnetization, which leads to a measurable change in the electrical resistance. There are two ways to program the cell. In traditional MRAM technology, known as field-writing MRAM, magnetic field is induced in the MRAM cell by passing current through a pair lines that run parallel to the cell. In the more recent version of MRAM cell, known as spin-transfer torque MRAM (STT-MRAM) [4], a spin-polarized current (i.e., current with

electrons of one spin type more than the other) is generated by passing current through the fixed layer which then passes through the free layer, changing its electron spin orientation to be either in the same direction as that of the fixed layer (parallel or low resistance configuration) or in the opposite direction to that of the fixed layer (antiparallel or high resistance configuration). This spin-torque based mechanism helps to reduce the write currents required to program the cell, compared to the traditional field writing MRAM variants. While MRAMs offer high speed, high density and non-volatility, high power consumption remains an issue. The relatively small difference between high resistance and low resistance states is also a concern as voltage level scales down.

3) Phase Change Memory: Phase change memory (PCM) devices typically consist of chalcogenide materials such as the commonly used alloy of Germanium, Antimony and Tellurium ($Ge_2Sb_2Te_5$) known as GST [5]. Reversible resistance change in such devices occurs due to the reversible transition of the chalcogenide material between amorphous (high resistance) and crystalline (low resistance) phases due to Joule heating under electrical bias [6], [7]. PCM memories are unipolar devices i.e. the voltage bias polarity remains the same during write or erase operations, only the bias magnitude and the duration may be different. PCM memory is among the more mature emerging memory technologies at present, having demonstrated scaling potential beyond 16 nm node [8]. High programming energy and write latency as well as resistance and threshold voltage drift over time have been observed in such devices.

4) Resistance Switching Memory: While technically PCM and MRAM are also resistance switching memories, the term resistance switching memory (ReRAM) has come to be used more commonly for devices that undergo bipolar resistance switching due to ion

drift under electric field. Two main categories of ReRAM devices can be considered devices that undergo resistance change due to drift of anions (specifically oxygen vacancies) in certain transition metal oxides (e.g. TaO_x, TiO_2, HfO_x) [9] – [11] and devices that undergo resistance change due to drift of metal cations in chalcogenide glasses (eg. Ge_xSe_y, Ge_xS_y) or other solid electrolytes [12]-[14]. Both these types of ReRAM devices consist of an oxide layer or a solid electrolyte layer sandwiched between two metal layers called the anode and the cathode. The first category is often referred to as RRAM or Ox-RRAM while the second category is known as either Electrochemical Metallization memory (ECM), programmable metallization cell (PMC) or commercially as conductive bridge random access memory (CBRAM) [15]. Both these types of devices demonstrate certain common characteristics such as bipolar resistance switching (unlike PCM), filamentary nature of resistance change and presence of multiple intermediate resistive states between the high resistance state (HRS) and low resistance state (LRS).

In RRAM, under applied electric bias, a soft breakdown of the dielectric oxide material may occur, generating and propagating oxygen vacancies through a hopping mechanism to form a conductive filament (CF), leading to resistance change. In another mechanism, the oxygen vacancies migrate to the cathode, and reduce the metal valency by trapping electrons. This creates a ‘virtual cathode’ which moves towards the anode. Thus a conductive region and non-conductive region are created within the oxide layer and this enables gradual resistance change to occur. The TiO_2 based HP ‘memristor’ is an example of such valence change induced resistance switching devices.

In CBRAM devices, the resistance change occurs due to transport of metal cations instead of anions from anode to cathode due to electrochemical redox reactions under

sufficient electrical bias, ultimately leading to electrodeposition of metal near the cathode which then grows back towards the anode, thus creating a metallic conductive filament. Resistance switches from HRS to LRS once the filament bridges the anode and cathode. Since CBRAM devices are the main focus of this work, a more detailed discussion about their operation principle will be provided in this and subsequent chapters. Throughout this work, the term CBRAM will be used most frequently, but the terms PMC or ECM may also be used as alternative acronyms to refer to the same device technology [16]-[18].

1.2 Conductive Bridge Resistive Memory: Theory

Structurally, CBRAM devices consist of a metal–electrolyte–metal or metal-insulator-metal (MEM/MIM) vertical stack. The two metal layers act as electrodes, one consisting of an electrochemically active metal (referred to as anode) and the other consisting of an inert metal electrode (referred to as cathode). As mentioned earlier, the CBRAM cell achieves resistive switching through the formation and dissolution of nano-sized metallic conductive filaments between the anode and cathode within the solid-state electrolyte material. Filament growth and dissolution occurs due to electrochemical reduction-oxidation (redox) reactions and the electric field assisted transport of metal ions.

Several different combinations of materials have been explored as part of research into improving the desirable performance-related properties such as endurance, retention and switching power consumption. Typically Silver (Ag) or Copper (Cu) have been chosen for the electro-chemically active metal layer while inert metals such as Tungsten (W), Nickel (Ni) or Platinum (Pt) have been used for cathode layer. A large variety of insulating materials can be utilized as the electrolyte layer [19]-[21]. Since the electrolyte layer can be fabricated to be ultra-thin (in the order of tens of nm), ions from the anode can be

transported quickly due to the large electric fields produced across the thin electrolyte layer even with voltages between 1V and 2 V. Thus even insulators such as Silicon Dioxide (SiO_2) have been used for CBRAM devices. Germanium-based chalcogenide glass materials (Ge_xS_y and Ge_xSe_y of different stoichiometric ratios) have been the most widely used as the electrolyte layer. Results [19] comparing the two chalcogenides seem to indicate that CBRAM devices using Ge_xS_y are more thermally stable and less susceptible to electrical degradation during annealing than devices with Ge_xSe_y, when Ag is used as the anode.

Since the underlying resistive switching mechanism and electrical behavior remains more or less the same across the different types of CBRAM technology, the research presented in this work can be expected to be applicable for the CBRAM based memory technology in general. The specific structure of devices examined in this study consist of an Ag-doped chalcogenide glass electrolyte layer (either $Ge_{30}Se_{70}$ or GeS_2) sandwiched between an anode formed with an active metal (Ag) and a cathode formed with an inert metal (W or Ni) [22].

Like other ReRAM devices, the CBRAM is a two terminal passive element capable of resistance change when a voltage, above or below specified thresholds, is applied across its terminals. Ion hopping has been proposed as the primary mechanism of cation transport [12-14]. At the cathode, the cations are reduced and electrodeposited, forming a conducting filament that grows towards the anode. Upon reaching the anode, the filament provides a low resistance path through the film. This constitutes a change from a high resistance state (HRS) to a low resistance state (LRS). Under reverse electrical bias, the filament undergoes dissolution and eventually the device returns to its HRS condition. A typical CBRAM I-V

characteristic is shown in Fig. 1.1(a), which represents a typical resistance switching characteristic for the CBRAM devices. The figure shows a write voltage threshold (HRS to LRS switching) between 200 mV and 300 mV. The erase voltage threshold (LRS to HRS switching) was less than -100 mV. Similar write/erase thresholds are obtained independent of device area. The maximum current allowed to flow through the devices during characterization, i.e., the compliance or programming current, was set to 10 µA for the DC response shown in Fig. 1.1(a). Previous studies have determined that the LRS resistance shows an inverse power law dependence on the magnitude of the programming current compliance [15]. Similar LRS resistance vs. compliance current characteristics was observed for the CBRAMs examined in here as well, as shown in Fig. 1.1(b).

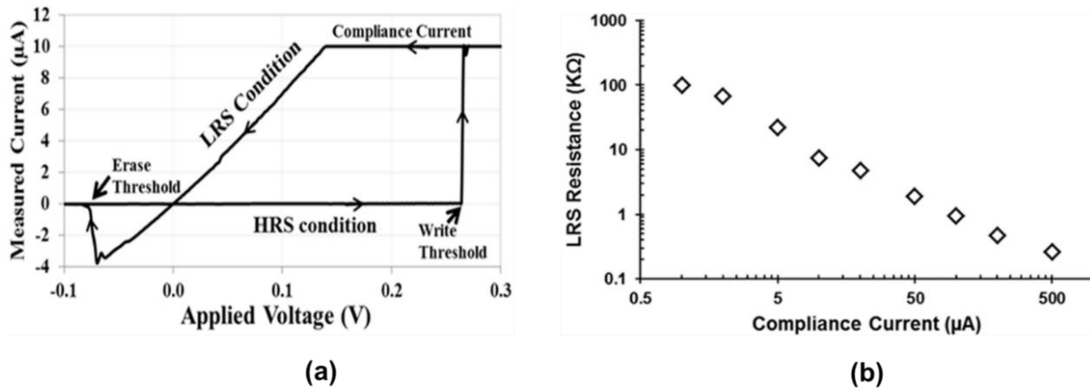


Fig. 1.1. (a) DC I-V characteristics of a CBRAM device with a 10 µA compliance current which shows the switching between high resistance state (HRS) and low resistance state (LRS). (b) Experimentally obtained inverse power law dependence of LRS CBRAM resistance on compliance (programming) current used during DC characterization.

Test chips containing several CBRAM devices arranged in an array were fabricated at Arizona State University (ASU). The test chip fabrication procedure for Ag- $Ge_{30}Se_{70}$ CBRAM devices is as follows [23]. At first, 100 nm of SiO_2 is deposited on a p-doped Si wafer using electron-beam evaporation. Then 100 nm of Ni (device cathode) and 100 nm

of SiO_2 (for isolation) are deposited successively. Device vias are patterned in the isolation layer and then wet etched to expose the Ni layer. The vias are used either as contact holes to the Ni electrode, i.e., cathode vias, or as "device" vias which define the active area of the CBRAM. After etching, 80 nm $Ge_{30}Se_{70}$ and 30 nm Ag films are deposited on the vias using a Cressington thermal evaporator. The wafer is then exposed to UV light (λ = 324 nm, E = 3.82 eV) for 1 hour at a power density of 10 mW/cm^2 in order to allow Ag photo-doping of the chalcogenide film. An additional 35 nm of Ag is then deposited on top of the silver-doped $Ge_{30}Se_{70}$ layer to create the device anode. Aluminum (800 nm) is then deposited and etched to form contacts for both device electrodes. Finally, the wafer is annealed at 120° C for 20 minutes. Fig. 1.2(a) illustrates a cross-section of the CBRAM device. Fig. 1.2(b) shows the top-view photo-micrograph of a CBRAM fabricated with a 500 µm by 500 µm via dimension.

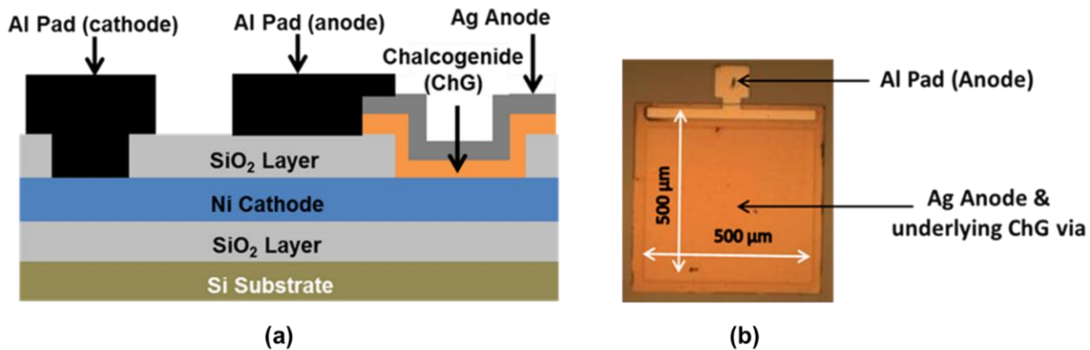


Fig. 1.2. (a) Illustration of CBRAM device cross-section with Ag anode, Ni cathode and chalcogenide ($Ge_{30}Se_{70}$) via layer. (b) Micro-photograph showing top (planar) view of a CBRAM device with via dimension 500 µm by 500 µm.

1.3 CBRAM Programming Dynamics

Previous section described the overall electrical characteristics and switching behavior of CBRAM devices. The programming operation of CBRAMs (and other metallic ion transport based resistive memory devices) consists of a series of electrochemical

oxidation/reduction reactions and subsequent ion migrations occurring under external electrical bias. The oxidation of Silver (Ag) takes place mainly at the anode side and allows for Ag ions to be created for the CF, while reduction of these same Ag ions occurs both at the cathode side and in the electrolyte (Chalcogenide) layer. Thus an electrochemically active metal is required for the anode. The oxidation and reduction reactions are given by (1.1) and (1.2) below respectively-

$$Ag \rightarrow Ag^{+} + e^{-} \quad (1.1)$$

$$Ag^{+} + e^{-} \rightarrow Ag \quad (1.2)$$

In order to overcome the electrochemical energy barrier, an external voltage bias needs to be applied across the device electrodes. A condition where the anode is sufficiently positively biased with respect to the cathode initiates the write process; oxidation of Ag at the anode and its subsequent migration towards cathode followed by reduction to grow the CF. Conversely, a sufficiently negatively biased anode with respect to the cathode starts the erase process; oxidation of Ag in the formed CF and its migration towards the anode where it undergoes reduction to form metallic Ag again. The Silver ions react with the electrolyte to form conductive phases which then combine together to form the conductive filament. Each set of oxidation and reduction reactions at the two electrodes cause a charge transfer and generate an electrode current which has been modeled by the Butler-Volmer equation [24] as,

$$I = I_o \left[e^{\frac{\alpha\eta}{kT/q}} - e^{-\frac{(1-\alpha)\eta}{kT/q}} \right] , \quad (1.3)$$

where I_o is the equilibrium current which refers to the current when the rate of reaction at anode is equal to the rate of reaction at the cathode. The first exponential term refers to

cathode current and the second term refers to anode current, while α is the transfer co-efficient, η is the current-over potential and kT/q is the thermal voltage. However, of relatively greater importance from device modeling perspective is the current flow associated with the Ag^+ ion transport inside the electrolyte, which ultimately leads to formation/growth/dissolution of the conductive filament and thereby determines the change in CBRAM resistance. This current flow is associated with electrical field influenced "hopping" of ions between adjacent sites in the solid electrolyte layer in the device under external bias, and can be modeled by a drift current density following the Mott-Gurney ionic transport model [25] as

$$J_h = 2ZqN_i a v_o e^{\left(-E_a/_{kT}\right)} sinh\left(^{q\alpha E}/_{2kT}\right) \quad , \tag{1.4}$$

where Zq is the total charge of the ions, N_i is the density of ions in the electrolyte, υ_o is the vibration frequency of the ions, α is the ion hopping distance and E is the net electric field at the tip of the growing filament given by

$$E = V/(L + h \cdot (\rho_f/\rho_e - 1)) \quad . \tag{1.5}$$

In Eq. (1.5), V is the potential difference between anode and cathode, ρ_f and ρ_e are the resistivity of the filament and the bulk electrolyte, respectively, and L is the electrolyte layer thickness. As $\rho_f << \rho_e$, the denominator of (1.5) can be simplified to (L-h). When modeling the temporal evolution of the conductive filament inside the electrolyte in CBRAM, the filament is modeled two dimensionally with variable vertical height $h(t)$ and lateral radius $r(t)$. The vertical growth rates of the filament can be modeled as

$$\frac{dh}{dt} = \frac{J_h}{ZqN_m} \quad , \tag{1.6}$$

where N_m is the density of electrodeposited metal (cm^{-3}) in the electrolyte. By combining (1.4) and (1.6), the following expression can be used to model the vertical filament growth rate as

$$\frac{dh}{dt} = v_h e^{\left(-E_a/_{kT}\right)} sinh\left(^{q\alpha E}/_{2kT}\right), \tag{1.7}$$

where v_h is a fitting parameter. Once the filament is formed, the LRS current density can be modeled as J_r, given by

$$J_r = v_r{}' e^{\left(-E_a/_{kT}\right)} sinh\left(^{q\beta V}/_{kT}\right) \ , \tag{1.8}$$

where $v_r{}'$ is fitting parameter and the term βV is a substitute for the vertical electric field E in (1.5) representing the lateral electric field, β being the lateral electric field fitting parameter. The filament can be modeled as either as a conical or cylindrical shape. Microscopic images suggest a conical shape, being wider near the cathode and narrower near the anode [26]. The volume (*Vol*) of a conical filament in the LRS condition can be expressed in terms of the variable top radius r and the bottom radius R as

$$Vol = \frac{\pi L}{3}(R^2 + Rr + r^2)\ . \tag{1.9}$$

The incremental change in the filament volume (*dVol*) given by

$$dVol = \frac{J_r.dt.\pi r(t)^2}{ZqN_m}\ , \tag{1.10}$$

Using (2.8) and (2.9), the lateral filament radius growth rate can be derived as

$$\frac{dr}{dt} = \frac{r.v_r}{LqN_m} e^{\left(-E_a/_{kT}\right)} sinh\left(^{q\beta V}/_{kT}\right) \ , \tag{1.11}$$

where v_r is a fitting parameter (proportional to $v_r{}'$) and r is the current filament top radius. In chapter 3, the above analytical equations will be used to estimate the variable CBRAM resistance and develop a compact model for circuit simulations.

CHAPTER 2

EXPERIMENTAL DEVICE CHARACTERIZATION

2.1 Impedance Spectroscopy Analysis

AC measurements or impedance spectroscopy is a widely used method to characterize electrode-electrolyte-electrode material systems. In order to measure the ac response of a test system [27]-[28], a direct measurement of the impedance in the frequency domain is performed over a wide range of frequencies to obtain the frequency spectra. The impedance of any system can be defined as $Z = V/I$ and contains both resistive (R) and reactive (L and/or C) components. All of these component parameters are determined by applying a voltage across the test system $V(\omega) = V_0 Sin\ (\omega t)$ where ω represents the angular frequency ($2\pi f_0$). Unless the system under test is a perfect resistor, the resulting current will undergo some phase shift (Φ) with respect to the voltage, given by $I(\omega) = I_0 Sin\ (\omega t + \Phi)$. The impedance response can then be calculated as $Z(\omega) = V(\omega)/I(\omega)$. The complex impedance $Z(\omega)$ can be expressed as

$$Z(\omega) = Z'(\omega) - jZ''(\omega), \tag{2.1}$$

where $Z'(\omega)$ represents the real (resistive) component and $Z''(\omega)$ represents imaginary (reactive) component. By varying the input angular frequency (ω), the frequency spectra of the system impedance can be obtained. Another related complex formalism of interest is the electric modulus (M), which can be derived from complex impedance as

$$M(\omega) = j\omega C_0 Z(\omega) = M'(\omega) + jM''(\omega), \tag{2.2}$$

where $C_0 = (\varepsilon_0.A/l)$ is the vacuum capacitance of the same test structure (area A and thickness l). The different regions of a test system are normally characterized by a resistance and capacitance, usually placed in parallel. In the frequency domain, the

different RC elements are separable [29]. From the impedance spectrum, different RC elements associated with different regions of the test system are identified and the values of the individual components are then extracted.

Fig. 2.1 shows a typical complex impedance spectra of a simple parallel RC element in the form of a Cole-Cole plot [30], where the negative of imaginary component ($-Z''(\omega)$) is plotted against the real component ($Z'(\omega)$). It can be seen that the frequency spectra is a semi-circular arc passing through origin and intercepting the real impedance axis at the value of the element resistance R_1.

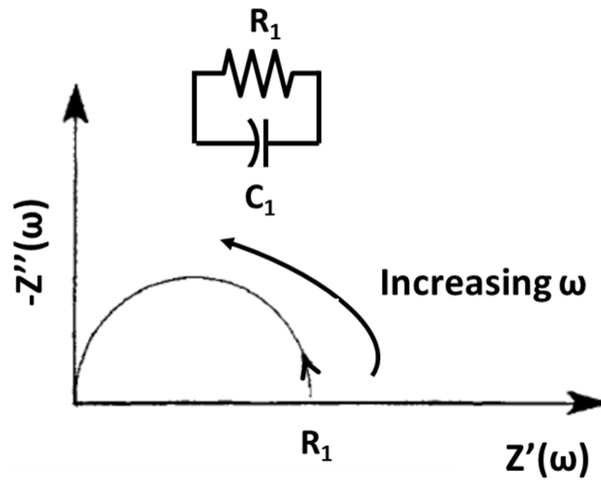


Fig.2.1. Cole-Cole Plot representation of complex impedance spectra of a parallel RC element (R_1-C_1).

This method of extracting the impedance spectra for a system was utilized to characterize the Ag-$Ge_{0.3}Se_{0.7}$-Ni CBRAM devices fabricated at ASU. Since circuit performance is one of the primary technology drivers for applications involving such devices, it is important to examine the electrical impedance of CBRAMs. Electrochemical Impedance Spectroscopy (EIS or IS) has been previously used to characterize different nanodevices and thin films [31]-[34]. The metal-solid electrolyte-metal structure of these devices indicates the presence of capacitive elements along with the programmable device

resistance. In this work [35] impedance spectra of multiple CBRAM devices with different areas are measured experimentally. Measurements are conducted on each device in both high resistance state (HRS) as well as low resistance state (LRS). The associated resistance and capacitance parameters are then extracted to construct scalable equivalent RC circuit models for both states.

As mentioned in Chapter 1, CBRAM devices are exposed to UV light during the fabrication process. This step is known as photo-doping, which causes Ag from the anode to be partially incorporated into the chalcogenide glass. This photo-induced doping [36]-[37] changes the electric properties of the Chalcogenide glass from dielectric to electrolytic (hence ionic conduction is now possible), leading to the formation of a solid-state electrolyte [38]-[40]. Moreover, this step helps to eliminate the need for electro-forming (i.e. the initial high voltage/current programming step), ensuring uniform and repeatable switching behavior of the device. Photo-induced Ag-incorporation leads to creation of two layers within the chalcogenide via layer: a region rich in Ag on the anode side and a region with much lower Ag concentration on the cathode side. This effect is observed from TEM images of the device as well as from results obtained with atomic profile measurements. Impedance spectra analysis corroborates this two-layer structure.

Experimental Details

Both DC and AC frequency response measurements were performed on the CBRAMs with different via areas: 100 µm by 100 µm, 250 µm by 250 µm, 400 µm by 400 µm and 500 µm by 500 µm. The resistance switching characteristics were obtained by sweeping the voltage from -0.5V to 0.5V with an Agilent 4156C semiconductor parameter analyzer. Impedance measurements were performed with an Agilent 4284A LCR meter on CBRAMs

in both HRS and LRS. Impedance spectra were obtained between 20 Hz and 1 MHz with a small signal AC voltage (10 mV RMS) applied across the device. Since this applied voltage is well below the device switching threshold voltage, the state of the device is not disturbed during the small signal measurement.

Impedance Spectra Characterization of CBRAM in HRS

Transmission Electron Microscopy (TEM) was performed on a HRS CBRAM to investigate the device cross-section. TEM was conducted with a JEOL ARM200 microscope and images were acquired for a magnification of $2x10^6$ with 200keV electron and a beam current of 14 pA. Fig. 2.2(a) shows the cross-section of a device in the HRS using high angular dark field. All the major layers are identified, including the Ag anode, Ni cathode, chalcogenide active via and the SiO_2 isolation layer. The Ag and Ni layers show crystalline granularity, while the chalcogenide layer is amorphous and formed by two different layers within via: a photo-doped Ag-rich layer (lighter in color) and an Ag-poor region (darker). This is seen in Fig. 2.2(b) which is a magnified view of the ChG-Ni interface.

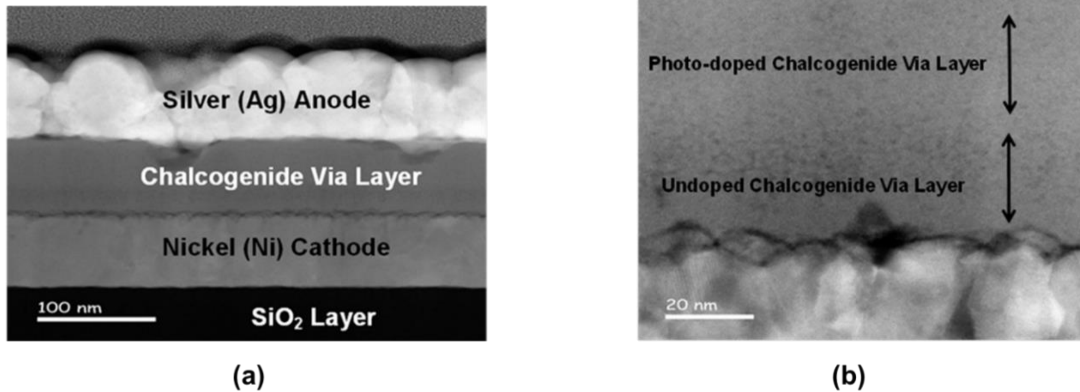


Fig.2.2. (a) TEM image of a CBRAM device (HRS condition) cross-section using high angular dark field. (b) Magnified view of the same cross-section near the chalcogenide via-Ni cathode interface, indicating presence of two layers (one lighter and another darker) within the chalcogenide (ChG) more clearly.

To estimate the composition of the layers within the CBRAM via, energy dispersive x-ray scattering (EDS) was used to obtain profiles of species present in the device via. The electron beam was scanned along a line in the plane of device via and a characteristic emission x-ray spectrum was retrieved at each point of the scanned line. The profiles obtained for the CBRAM are shown in Fig. 2.3. From the obtained atomic profiles along the device, it can be seen that the Ge and Se profiles are almost identical and indicate the ChG film thickness. The Ag profile is high and constant in the beginning, indicating the Ag anode layer, and then drops to a lower value which stays at more or less the same level, indicating the presence of a photo-doped Ag rich layer within the via. The sharp fall in Ag profile near the Ni cathode indicates the boundary with a second layer, which has considerably less Ag concentration.

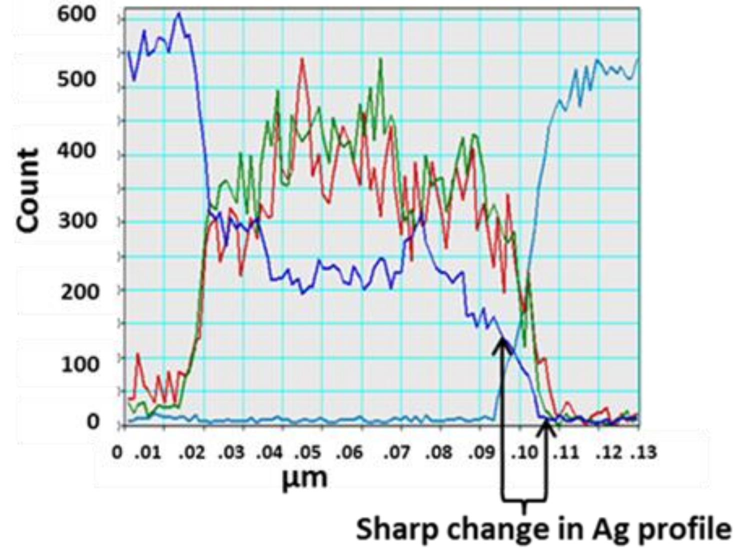


Fig.2.3. Atomic profiles of Ge (red), Se(green), Ag(dark blue) and Ni(light blue) along the CBRAM device, showing the sharp change in Ag count near the cathode side, indicating existence of a Ag gradient as modeled in the impedance spectra equivalent circuit. The thickness of the Ag- poor chalcogenide layer can be estimated to be approximately 5 nm based on the obtained profile.

It is known that device illumination with UV light causes Ag to diffuse into the chalcogenide (ChG) via layer [36]-[40], forming a solid-state electrolyte. During this process, Silver diffuses from the top Ag electrode towards the bottom Ni electrode, reacting with the ChG to form different phases within the film. After annealing, which is the final processing step used on the devices studied in this work, binary phase like Ag_2Se may appear [41]. After processing, the CBRAM via may be described as a non-uniformly doped electrolyte with a significantly reduced Ag doping concentration close to the bottom Ni electrode. Measured spectra of all the HRS CBRAM devices across via areas demonstrated similar characteristics that indicate presence of two regions within the ChG via with different electrical characteristics. Based on the experimental impedance spectra and the physical characterization results discussed above, an equivalent circuit consisting of a contact resistance (R_{cont}) in series with two separate parallel RC configurations is extracted.

Fig. 2.4 plots the imaginary components of complex impedance (*-Z''(ω)*) (Fig. 2.4(a)) and of its corresponding electric modulus (*M''(ω)*) (Fig. 2.4(b)) vs. their respective real components *Z'(ω)* and *M'(ω)* for a 100 µm x 100 µm CBRAM in HRS. The plots indicate two separate regions, each with distinct RC time constants.

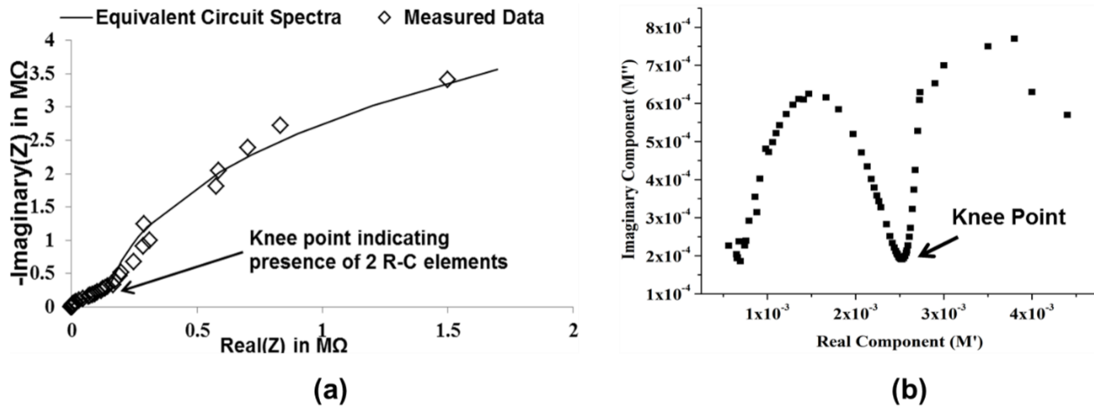


Fig.2.4. (a) HRS impedance spectra of a CBRAM device (100 µm by 100 µm diameter) showing both measured data and fitted characteristics. (b) Imaginary component vs. real component of Electric Modulus (M (ω)) derived from complex impedance data of (a).

The top RC configuration, composed of R_1 and C_1, determines the impedance of the photo-doped portion of the ChG film and the second RC configuration, composed of R_2 and C_2, is associated with the un-doped (Ag-poor) chalcogenide (ChG). Each of the parallel RC elements produces a semicircle in the complex impedance plane.

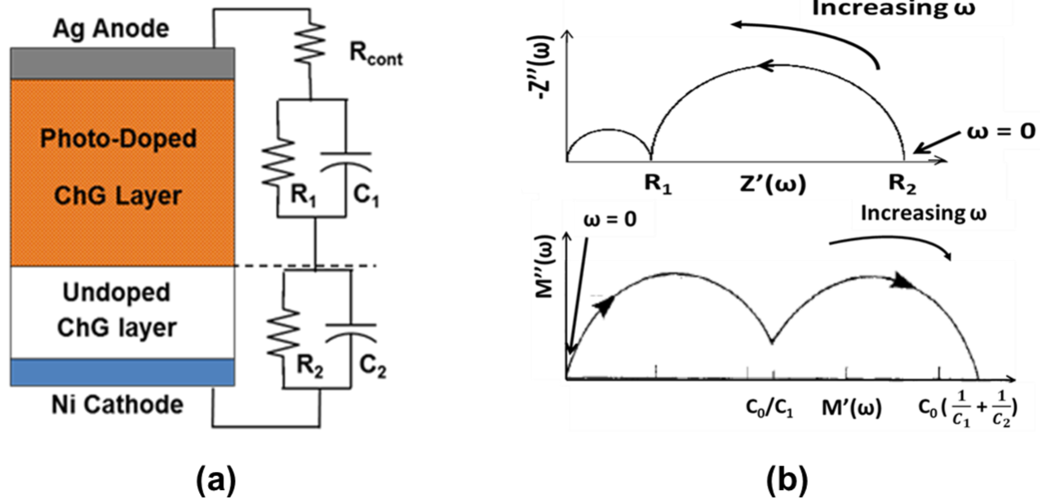


Fig.2.5. (a) Photo-doped HRS CBRAM RC equivalent circuit with two regions with distinct RC characteristics. (b) Idealized complex plane representation of real and imaginary components of impedance (Z(ω)) (top) and electric modulus (M(ω)) (bottom) associated with 2 parallel RC element circuit.

Mathematically, HRS CBRAM complex impedance is given by

$$Z(\omega) = (R_1 // \frac{1}{j\omega C_1}) + (R_2 // \frac{1}{j\omega C_2}), \tag{2.3}$$

Using (2.1), the real and imaginary components of impedance can be derived (ignoring the small R_{cont}) as

$$Z'(\omega) = \frac{R_1}{1+(\omega R_1 C_1)^2} + \frac{R_2}{1+(\omega R_2 C_2)^2}, \tag{2.4}$$

$$Z''(\omega) = \frac{R_1(\omega R_1 C_1)}{1+(\omega R_1 C_1)^2} + \frac{R_2 \omega R_2 C_2}{1+(\omega R_2 C_2)^2}, \tag{2.5}$$

Corresponding real and imaginary components of *M(ω)* (Eq. (2.2)) can be derived from (2.3) as

$$M'(\omega) = \left(\frac{(\omega R_1 C_1)^2}{1+(\omega R_1 C_1)^2}\right) . \frac{C_0}{C_1} + \left(\frac{(\omega R_2 C_2)^2}{1+(\omega R_2 C_2)^2}\right) . \frac{C_0}{C_2}, \tag{2.6}$$

$$M''(\omega) = \left(\frac{\omega R_1 C_1}{1+(\omega R_1 C_1)^2}\right) . \frac{C_0}{C_1} + \left(\frac{\omega R_2 C_2}{1+(\omega R_2 C_2)^2}\right) . \frac{C_0}{C_2}, \tag{2.7}$$

Fig. 2.5(a) shows the HRS equivalent circuit with the two parallel RC elements and Fig. 2.5(b) shows the complex plane plots using the analytical equations derived above.

The extracted value for R_{cont}, which represents the resistance of the probe contact, was less than 100 Ω and is independent of via size. Fig. 2.6 plots the average values for R_1 and R_2 as a function of via area, respectively. It can be seen that both R_1 and R_2 vary inversely with via area as expected.

It can also be observed that R_2 (the resistance of the un-doped layer) is nearly 100 times greater than the resistance of the photo-doped layer, R_1. Since the resistances are in series, it is the un-doped layer resistance that determines the DC HRS impedance. The HRS resistance ranges from approximately 10 MΩ for the smallest via area to approximately 200 kΩ for the largest.

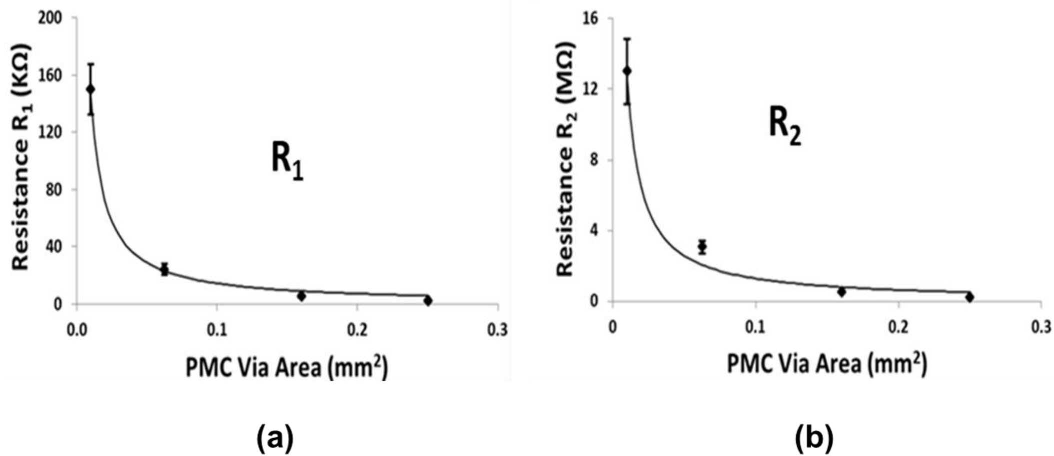


Fig. 2.6. Plots showing scaling of HRS CBRAM resistance parameters R_1 and R_2 with via area. Both R_1 and R_2 vary inversely with via area.

Fig. 2.7 plots the dependence of C_1 and C_2 on CBRAM via area. The capacitance of the photo-doped region (C_1) exhibits a higher-order, non-linear dependence on via area. This may be due to variations in the dielectric properties of Ag-doped GeSe caused by non-

uniform incorporation of Silver into the material. As the Fig. 2.7 (b) shows, unlike the capacitance of the doped chalcogenide material, the capacitance in the undoped layer (C_2) scales linearly with via area.

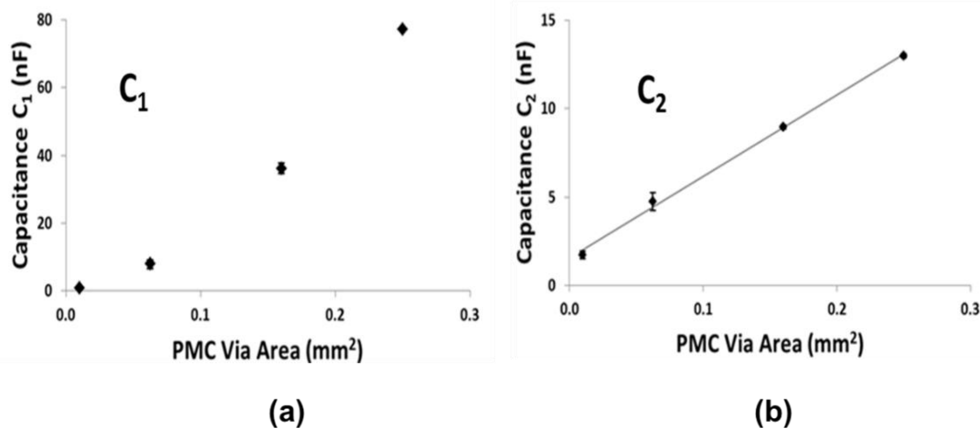


Fig.2.7. Scaling of HRS CBRAM capacitance parameters showing non-linear scaling of HRS CBRAM capacitance parameter C_1 (a) with area and linear scaling of HRS CBRAM capacitance parameter C_2 (b) with via area.

Impedance Spectra Characterization of CBRAM in LRS

During the write process (HRS to LRS switching), Ag^+ ions from the Ag-rich photo-doped layer drift towards the cathode, eventually creating a conductive Ag filament across the entire via. In this condition, the device is in LRS. The resistance across the chalcogenide film can now be assumed to be dominated by the resistance of the conductive filament and thus the LRS CBRAM can be modeled as a contact resistance in series with one parallel RC element. The impedance of this configuration can be represented on the complex plane with a single semicircle, as shown in Fig. 2.1 earlier.

Prior to small signal impedance measurement, a positive dc voltage sweep at a compliance (limiting) current of 10 μA was used to set each CBRAM device into a fixed low resistance state (LRS). For the LRS device equivalent circuit shown in Fig. 2.8 (a) the

equivalent circuit can can be described with a contact resistance R_{cont}, on-state resistance, R_{on} and the LRS device capacitance, C_{on}. A typical complex impedance Cole-Cole plot of a 100 μm x 100 μm device obtained experimentally is shown in Fig. 2.8 (b).

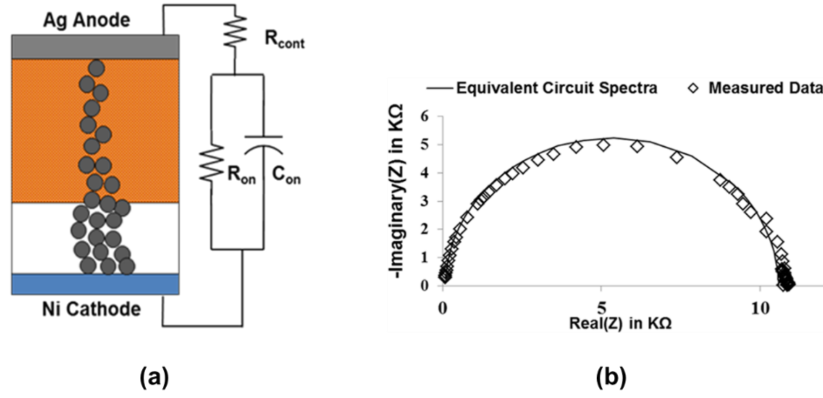


Fig.2.8. (a) LRS CBRAM equivalent circuit. (b) LRS impedance spectra of a CBRAM device (100 μm via diameter) showing both measured data and fitted characteristics, indicating single RC element.

Fig. 2.9 (a) plots average values of R_{on} vs. via area. As the figure shows, the on-state resistance is less than 10 kΩ and does not seem to depend on via size. This non-scalability can be attributed to the fact that R_{on} is determined by the resistance of the conductive Ag-filament formed across the film, which is likely independent of the device area. Finally, we observe the ratio of HRS to LRS resistance (R_2/R_{on}) ranges from approximately 2,000 for the smallest area device to 40 for the largest. Fig. 2.9 (b) plots the dependence of the LRS capacitance C_{on} on CBRAM via area. As with the HRS capacitance in the un-doped layer, the capacitance C_{on} scales linearly with via area. When the device switches into the LRS condition, the filament creates a low resistance pathway between the anode and cathode terminals while the LRS capacitance C_{on} is then in effect, the series combination of HRS capacitances C_1 and C_2, i.e.,

$$C_{on} = \frac{C_1 C_2}{C_1 + C_2}\,. \tag{2.8}$$

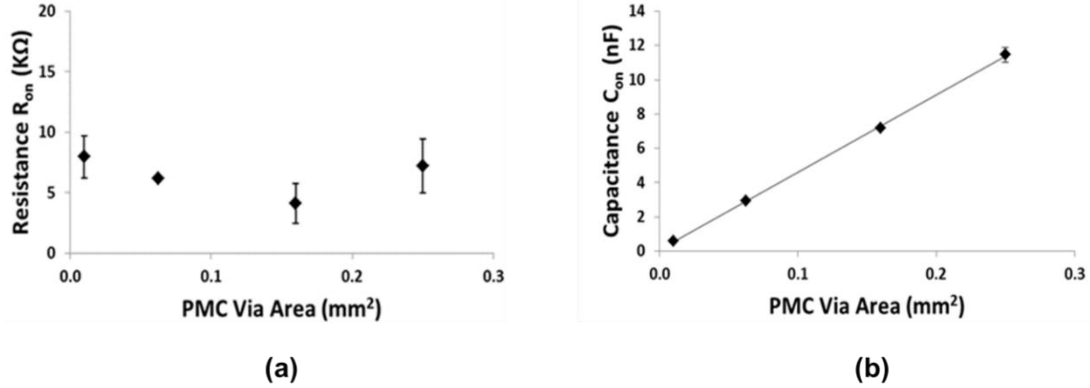


Fig. 2.9. (a) LRS Resistance parameter R_{on} vs. area. (b) LRS Capacitance parameter C_{on} vs. area.

Fig. 2.10 plots the extracted C_{on} as well its value obtained from Eq. 1, using the values for C_1 and C_2 obtained from the HRS fit. The good agreement between the two sets of values for C_{on} suggests that the application of a HRS to LRS switching bias has not appreciably altered the Ag concentrations across most of the doped and un-doped regions layers. Since the conductive filament formed in LRS is only a few tens of nm in diameter, a significant change in Ag concentration throughout the rest of the film is unlikely, especially considering the large dimensions of the devices (≥100 μm) characterized. Therefore, in LRS the Ag concentration, aside from the filament, has not changed, and thus capacitances in the doped region, C_1, and un-doped region, C_2, have not been measurably altered.

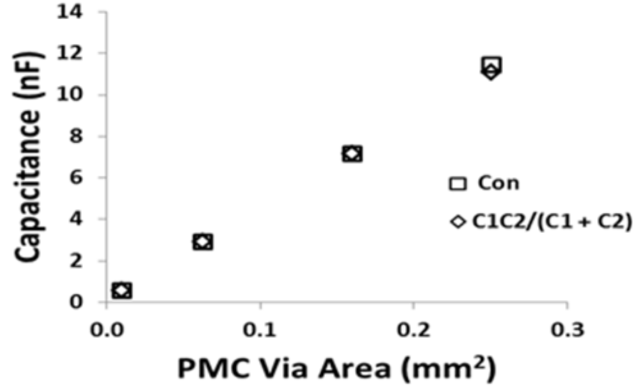


Fig.2.10. Capacitance vs. Area showing close match between C_{on} and series combination of C_1 and C_2.

Based on the linear scaling of extracted capacitance parameters C_2 and C_{on}, the capacitance per unit area of the CBRAM devices is calculated to be approximately 4.6 μF/cm^2. The dielectric constant (ε_r) of the undoped chalcogenide layer can be calculated from the HRS capacitance (C_2) per unit area (A) and the estimated thickness of the Ag layer (d) (which can be obtained from Ag profile measurement via EDS as shown in Fig. 2.3) as,

$$\varepsilon_r = \left(\frac{C_2}{A}\right).d/\varepsilon_0 \tag{2.9}$$

Using an estimated Ag layer thickness of 5 nm as observed in the EDS measurements, the value of ε_r is calculated to be approximately 26, which is somewhat larger than values previously reported for un-doped GeSe chalcogenide glass [42]. This discrepancy may be attributed to a condition whereby the "undoped" layer still contains residual quantities of Ag from the photo-doping process which can increase its relative permittivity over a pure un-doped film.

2.2 Quasi-Static Resistance Switching Characterization

The characteristic switching behavior of CBRAM devices was experimentally measured using the standard voltage double sweep using an Agilent 4156C semiconductor parameter

analyzer. Fig. 2.11(a) shows typical measured I-V characteristics, while Fig. 2.11(b) identifies the key electrical parameters that can be extracted from such a plot.

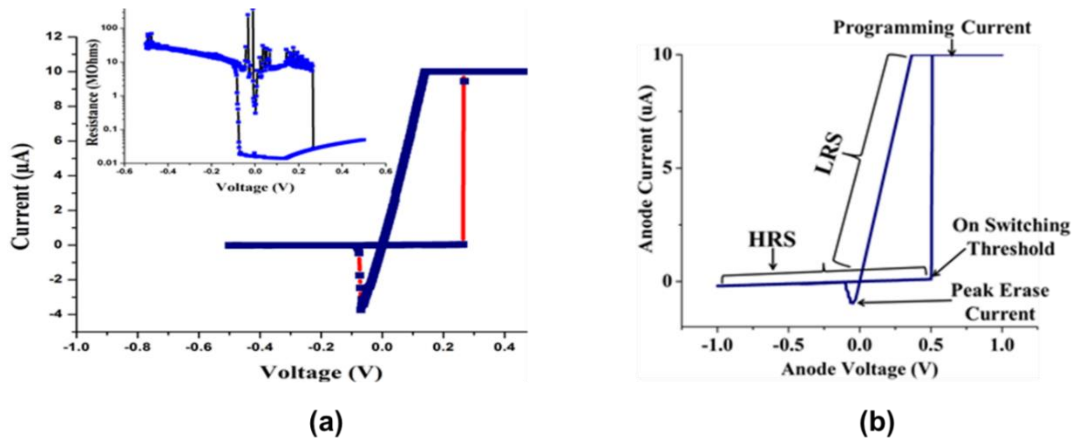


Fig. 2.11 (a) A typical measured DC I-V characteristics of a CBRAM device with a 10 µA compliance current, with the corresponding resistance vs. voltage plot (semi-log scale) shown as inset. (b) Another typical I-V characteristics with some key parameters identified such as on switching threshold, set programming current (compliance current), peak reset current, high resistance state (HRS) region and low resistance state (LRS) region.

CBRAM devices, like many other resistive memory technologies, demonstrate multi-level resistance programmability in its LRS state [43]. The most common demonstration of this property is performed by controlling the current through the device during programming, either using the in-built compliance current function of the 4156C instrument, or at circuit level by controlling the saturation drain current of the access transistor in a 1T-1R CBRAM element. The higher the programming current used, the lower is the final programmed LRS level. This is believed to be due to greater charge flow allowed at higher compliance leading to a thicker conductive filament (CF) in the LRS.

The switching behavior of Ag-GeSe CBRAM devices (5 µm x 5µm cross-section) was recorded over several cycles (≥ 25) for a wide range of compliance current levels. Experimental I-V characteristics at 100 µA compliance current over 20 write-erase cycles

are shown in Fig. 2.12 (a). Fig. 2.12(b) shows the decrease in LRS with increasing compliance current. Fig. 2.12 (c) and 2.12(d) show the dependence of LRS and peak erase current respectively on compliance current. As shown in Fig. 2.12(c) that LRS (R_{on}) is dependent on the compliance current ($I_{compliance}$) as,

$$R_{on} = \frac{0.363}{I_{compliance}{}^{1.14}} \; . \tag{2.10}$$

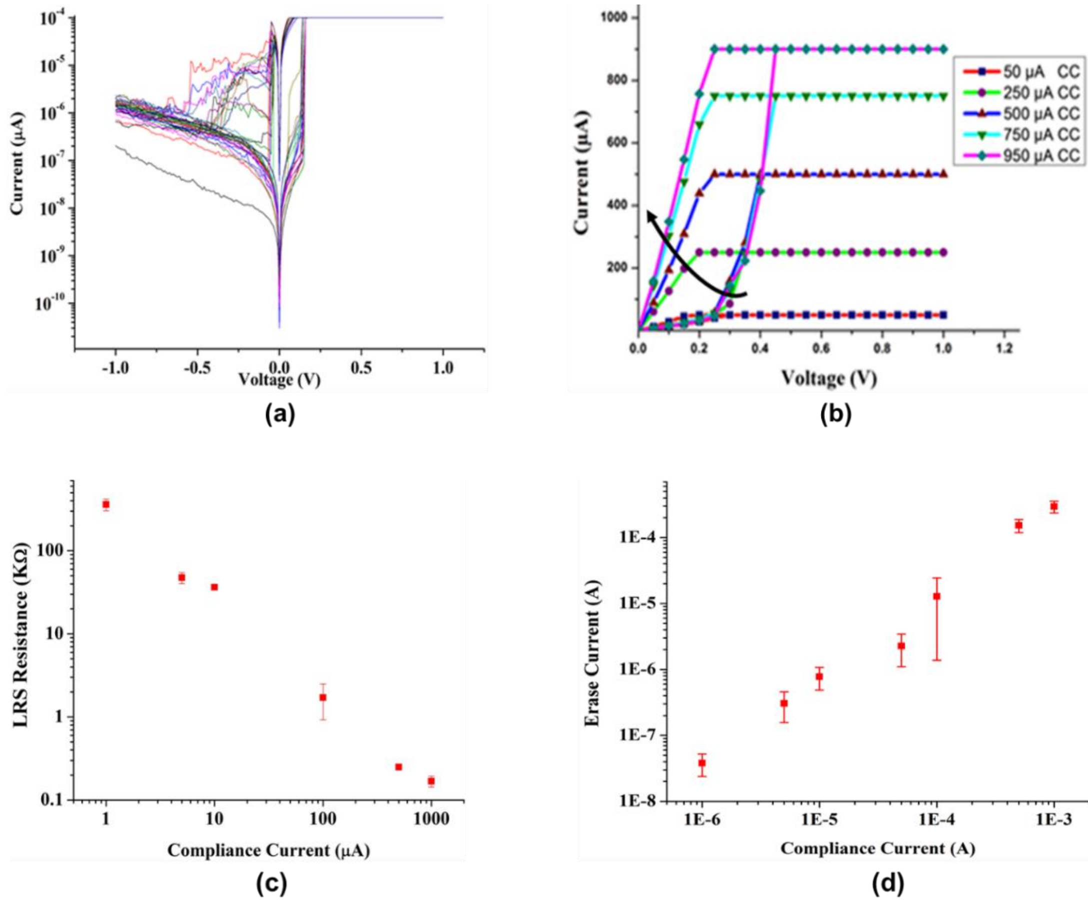


Fig. 2.12 (a) Experimental quasi-static I-V switching characteristics at 100μA compliance current for 20 cycles, (b) I-V characteristics for positive voltage double sweep between 0 V and 1 V at different compliance currents (CC). Arrow indicates direction in which LRS decreases. (c).Experimental dependence of LRS resistance on compliance current (log-log plot) during DC voltage sweep. (d) Experimental dependence of peak erase current on compliance current (log-log plot) during DC voltage sweep.

Another characterization test involved repeated DC sweeps performed at successively higher stop voltages so as to induce a gradual change in resistance. The compliance current was not reached during these sweeps. A series of positive voltage double sweeps starting from 0 V with successively increasing stop voltages (from 65mV to 120mV) on an already 'On' device made it possible to observe gradual decrease in resistance, likely due to increase in CF diameter . Such a set of positive sweeps is shown in Fig. 2.13 (a). For each successive sweep, the device starts from the resistive state defined by the previous sweep. An identical test was then performed for performing gradual erase. A gradual increase in resistance was obtained by applying a series of negative voltage double sweeps with successively increasing stop voltages (from -0.5V to -1.5V for each successive sweep, as shown in Fig. 2.13(b). This is most likely due to a gradual decrease in the CF diameter.

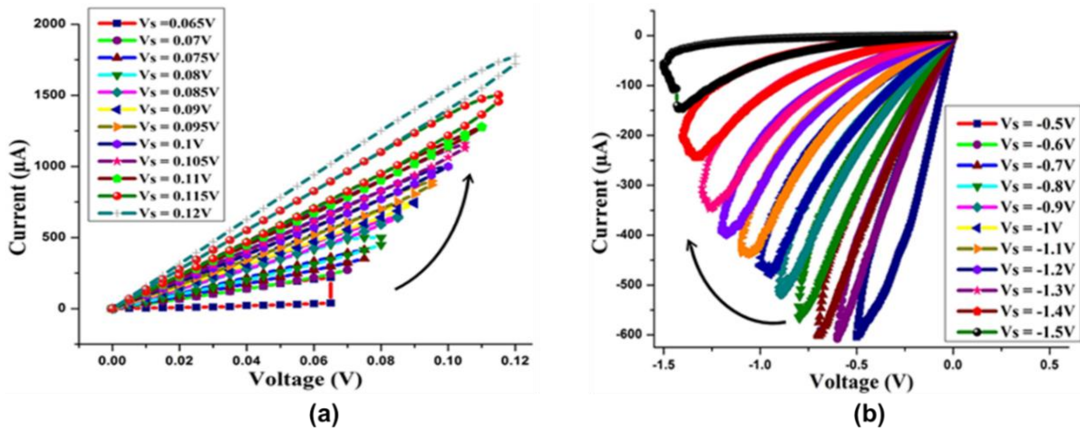


Fig.2.13 Incremental resistance change in CBRAM under DC bias: (a) resistance decrease under positive dc bias with repeated sweeps while increasing stop voltage (V_S) from 65 mV to 120 mV. Arrow indicates direction in which resistance gradually decreases. (b) Resistance increase under negative bias with repeated sweeps while increasing stop voltage (Vs) from -0.5 V to -1.5 V. Arrow indicates direction in which resistance gradually increases.

2.3 Pulsed Mode (Transient) Characterization

Multi-level Resistance Programmability

Transient characterization of the CBRAM involved applying positive and negative "write/erase" voltage pulses to decrease or increase the resistance of the device, respectively. The transient test setup is shown in Fig. 2.14 (a). An arbitrary waveform generator (Tektronix AWG520) is used to generate voltage pulses of either polarity and of different magnitudes and pulse widths. The waveform generator output is applied to a CBRAM device in series with a current limiting resistor R_L. In order to "read" the CBRAM resistance, a low voltage pulse is applied after each write/erase pulse. During the read operation, the voltage measured across the limiting resistor is used to sample the CBRAM resistance. The value of the limiting resistance is important. As mentioned previously earlier, the change in CBRAM resistance is a function of current flowing through the cell. A lower value of resistance will allow a larger voltage drop across the CBRAM as well as greater current flow through the device. Thus the limiting resistor is one of the factors controlling the amount of change in CBRAM resistance in this setup. The "read" pulse magnitude is set low so as not to disturb the resistive state. For this purpose, a read pulse not greater than 200 mV magnitude was used. Fig. 2.14(b) and 2.14(c) demonstrate a 'hard' write-erase operation whereby the PMC device is programmed to a LRS by a single positive pulse and then into a HRS by a single negative pulse. It is worth noting that the minimum voltage required to change the resistive state of the device is dependent on the voltage sweep rate [44]. Hence minimum voltage required to switch the device is greater for pulsed voltage (with much faster sweep rates) compared to DC or quasi-static voltage.

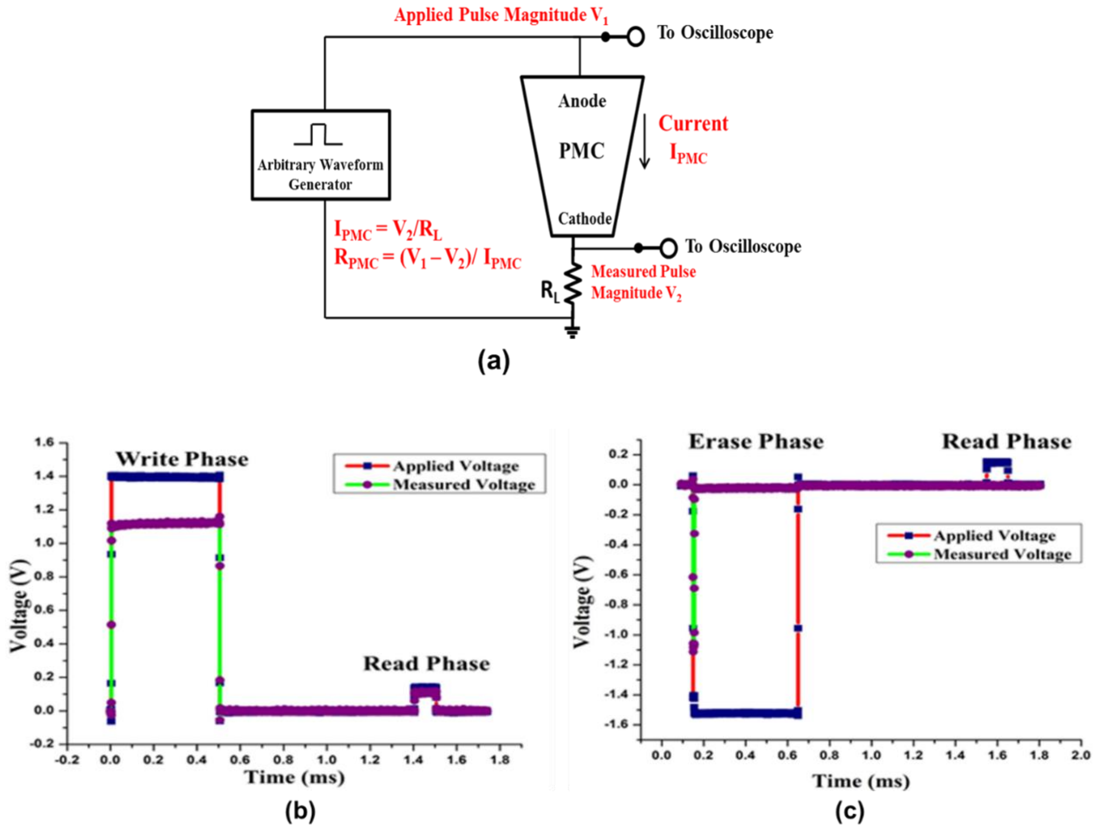


Fig.2.14 Transient voltage pulse based PMC programming:(a) experimental test setup (measured PMC resistance during read operation = $(V_1 - V_2)/I$); (b) Typical PMC write and read voltage transients showing applied voltage and measured output voltage; (c) Typical PMC erase and read voltage transients showing applied voltage and measured output voltage.

By varying the applied voltage magnitude and pulse width, the conductive filament lateral dimension and consequently the amount of resistance change in the device can be controlled. So it is possible to set the device into different LRS by applying voltage pulses of same pulse duration but of different amplitudes. Fig. 2.15(a) shows a typical write-erase operation used to demonstrate this behavior [45]. The simulation results (solid lines) are obtained using a CBRAM compact model based on equations presented in chapter 1 (the model is discussed in detail in chapter 3). Fig. 2.15(b) shows distribution of experimentally measured LRS for voltage pulses of amplitudes between 1.4V and 2V and 25 μs width.

Multiple write operations (≥ 20) were performed starting from an initial HRS (≥ 0.5 MΩ). Fig. 2.15(c) shows the mean and standard deviation of the LRS data presented in (b).

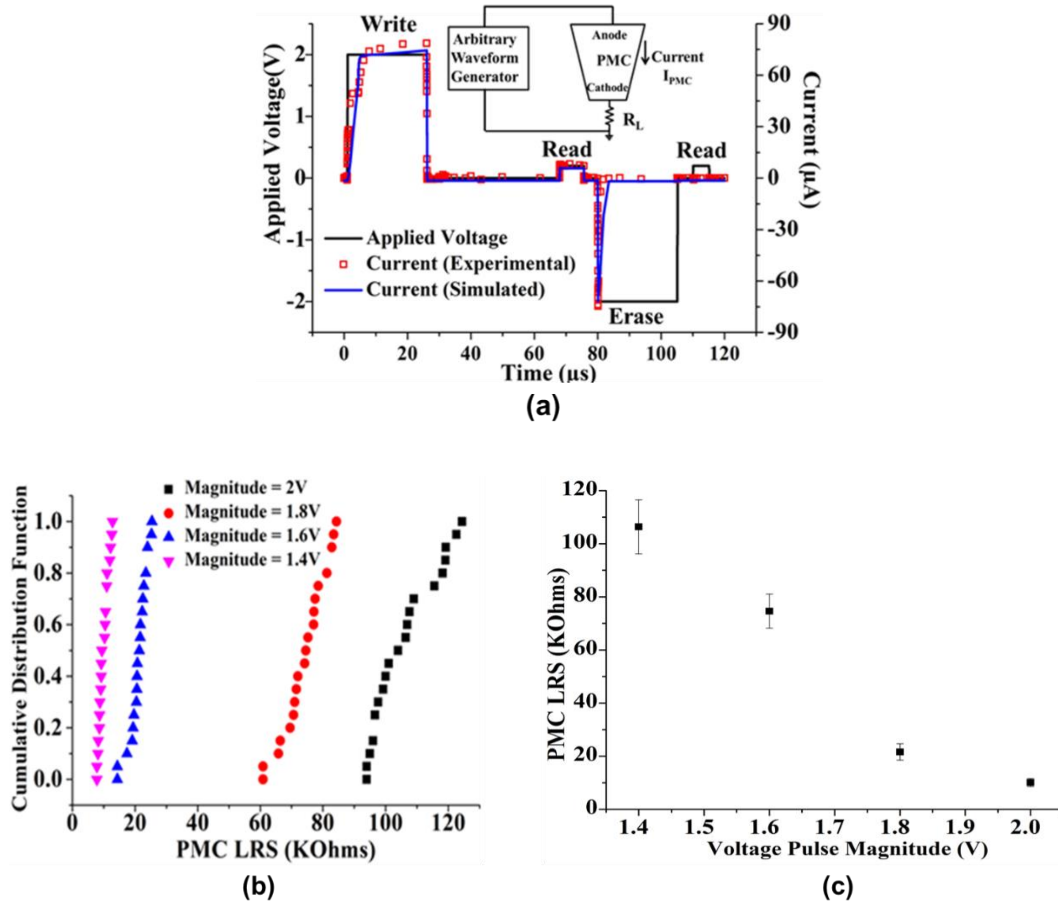


Fig.2.15 (a) Transient write-erase operation performed on CBRAM in series with a current limiting resistor (experimental setup shown in inset). Each write and erase pulse is followed by a low magnitude read pulse. (b) Experimental data showing multiple LRS for 4 different applied voltage pulse magnitudes. (c) LRS obtained from data of (b) vs. applied pulsed voltage amplitude.

A potential application of the multi-level programmability of CBRAM devices and other resistive memory devices is in neuromorphic circuit applications. In these circuits, the CBRAM provides a compact single device implementation of the synaptic conductance modulation observed in biological neurons (analogous to numerical weight change in artificial neural networks). For an electronic synapse, resistance should be capable of

changing gradually with the application of a voltage pulse (each pulse being analogous to the voltage spikes generated by biological neurons for conductance modulation of synaptic pathways) [46]-[47]. Controlling the magnitude and pulse width will determine the amount of resistance change after each step. Fig 2.16 shows a typical experimental result obtained for incremental resistance programming using the current limiting resistor setup described earlier (Fig. 2.14 (a)). The resistance of a CBRAM device was first gradually decreased to a low value and then gradually increased to restore the initial high resistance value. PMC resistance values were calculated based on measured output voltage during read pulse applied after each step. These results are obtained for constant magnitude positive and negative voltage pulses for program and erase, respectively. Specifically, ±1.5V magnitude pulses were used of width 100 µs width for gradual programming and 200 µs width for gradual erase. From the measured resistance values, each program/erase operation can be estimated to cause an average change of approximately 3% of the initial and final CBRAM resistance before and after the program and erase sequence.

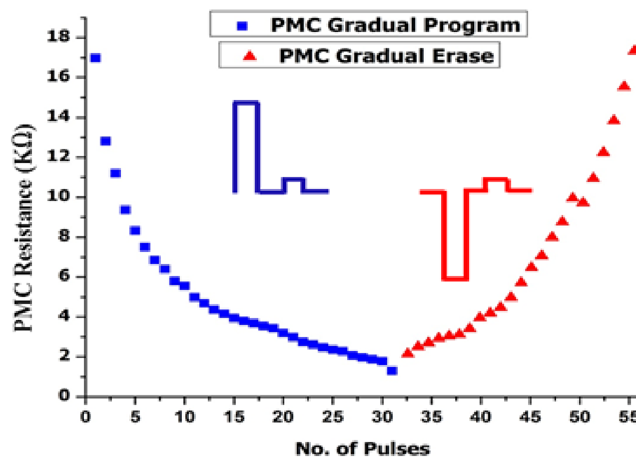


Fig. 2.16 Incremental resistance programming of a CBRAM device with 100 µs, 1.5 V write pulses and 200 µs, -1.5 V erase pulses, each pulse followed by 100 mV read pulses.

To demonstrate the effect of applied voltage pulse width on the change in resistance, a sequence of programming voltage pulses with increasing magnitudes and fixed pulse width were applied to the CBRAM device as shown in Fig.2.17 (a). Three different pulse widths of 100 µs, 500 µs and 1 ms were used. The results are shown in Fig. 2.17 (b). It can be seen that the change in CBRAM resistance is accelerated by using longer pulse width, as the largest change in CBRAM resistance occurs for 1ms pulse width sequence followed by the sequence with 500 µs and lastly 100 µs widths. This behavior is expected, as longer pulse widths allow more radial CF growth.

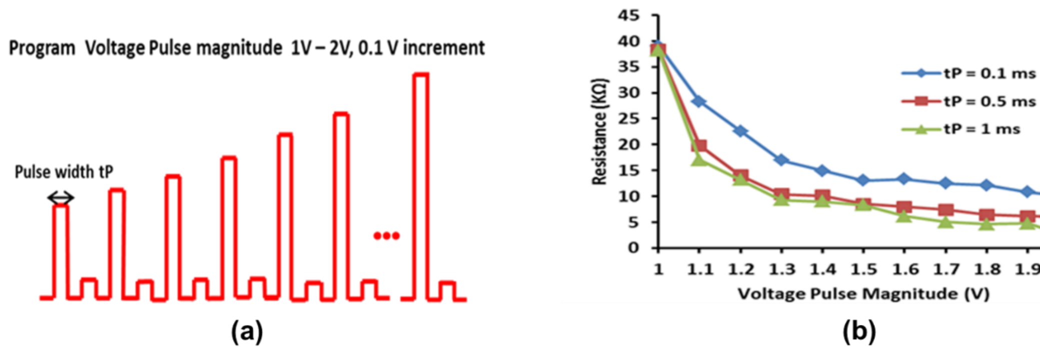


Fig 2.17 Incremental change in resistance depends on the width of applied voltage pulses. (a) Applied sequence of programming voltage pulses of pulse width tP. (b) Resistance vs. Applied voltage pulse magnitude for tP = 0.1 ms, 0.5 ms and 1 ms.

CHAPTER 3

CBRAM COMPACT MODEL FOR CIRCUIT SIMULATIONS

3.1 Resistance Estimation from Filament Dimensions

Based on the derivation of the conductive filament growth rates in section B of chapter 2, a time dependent expression for the change in resistance of a CBRAM device under electrical bias can be derived. For the conical filament assumed previously, the variable filament resistance $R_f(t)$ can be derived as [48],

$$R_f = \frac{\rho_f h}{\pi r R}, \tag{3.1}$$

where ρ_f is the filament resistivity, R is the bottom surface radius of the filament and h and r are the variable height and top radius, calculated using (2.6) and (2.10) respectively. The resistance of the filament in the initial state (in the HRS) can be expressed as,

$$R_{f,i} = \frac{\rho_f h_0}{\pi r_0 R}, \tag{3.2}$$

where h_0 and r_0 are the initial height and radius respectively of the filament ($h_0 < L$). These initial conditions in the 'Off' state may vary from device to device based on several factors. e.g., the filament height h_0 depends on the amount of photo-doping performed during fabrication or the magnitude and duration of the erase pulse. For the purposes of modeling, the following values were assumed: $h_0 = 10$ nm and $r_0 = 0.1$ nm. The resistance of the bulk portion of the electrolyte (excluding the filament) can be expressed as,

$$R_e = \frac{\rho_e . L}{\pi . (r_c^2 - R_{CF,bottom}{}^2)}, \tag{3.3}$$

where ρ_e is the bulk electrolyte resistivity. The overall resistance of the device can therefore be expressed as a parallel combination of the filament resistance and bulk resistance as,

$$R_{PMC} = \frac{1}{\frac{1}{R_e}+\frac{1}{R_f}}. \qquad (3.4)$$

Hence the overall CBRAM resistance is dependent on the applied voltage across the device (*V*) and the current flowing through the CBRAM can be expressed as,

$$I_{PMC} = V/R_{PMC} \qquad (3.5)$$

3.2 Compact Model Transient Simulation

In order to enforce the compliance current functionality implemented by the 4156 parameter analyzer instrument during I-V characterization using by the quasi-static voltage sweep, the current computed using (3.5) can be compared with a compliance current parameter (I_{comp}) during each simulation step and when $I_{CBRAM} \geq I_{comp}$, the voltage across the device becomes different from the applied voltage (*V*) and is given by,

$$V_{comp} = I_{comp}.R_{PMC}.sgn(V), \qquad (3.6)$$

where *sgn(V)* represents the signum function to assign the appropriate sign to V_{comp} based on the polarity of the applied voltage *V*. Fig. 3.1 shows simulation results for a voltage sweep induced write-erase operation and the corresponding change in the simulated filament height and radius.

To obtain the CBRAM resistance switching behavior, the voltage applied across the CBRAM/PMC element is ramped from 0 V to sufficient positive voltage and then to a sufficient negative voltage, before returning to 0 V again. When applied voltage is positive, the current through CBRAM model increases due to the vertical and lateral growth of conductive filament (based on equations described in chapter 1), until it reaches a compliance current (set using measuring instrument or by external control circuitry). At this point, the actual voltage across the element reduces to maintain constant current

through it. Due to the current being constant and the reduced voltage across the element, the filament radius and on state resistance of the CBRAM (which is defined by the filament dimension as per equation 3.1-3.4) also remain constant for the remaining time the applied voltage is positive. Then as the voltage polarity becomes negative, the filament begins to shrink laterally first and then vertically, until it returns to the original HRS state. Fig. 3.1 shows simulation results for the quasi-static voltage sweep induced write-erase operation of the CBRAM compact model described above for a voltage sweep between ±1 V and for a compliance current of 1 μA. As seen from Fig. 3.1, the experimentally observed effects of compliance current setting the CBRAM LRS and the reduction of voltage across the CBRAM can be captured by the compact model. The filament height remains unchanged once it becomes equal to the chalcogenide layer thickness (set at 60 nm in the model to match fabricated devices) i.e. when the vertical filament bridges the two electrodes.

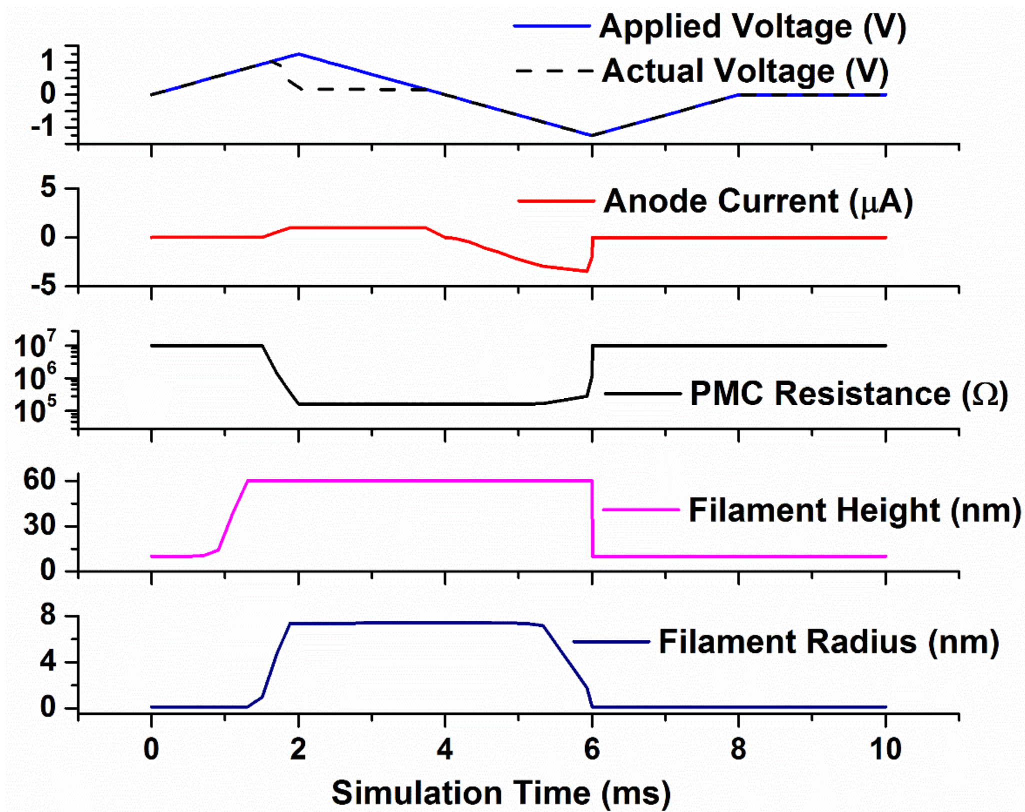


Fig. 3.1 Simulated voltage double sweep (from -1 V to +1V and back to -1V) to generate I-V characteristics of CBRAM model showing applied voltage and corresponding change in the filament height and radius.

Transient simulations can also be performed with the model to demonstrate write, read and erase operations. Fig. 3.2 shows simulation results for a 1T-1R CBRAM element for a sequence of write operation followed by a read operation and an erase operation followed by a read operation. In the case of a 1T-1R cell, the saturation current of the access NMOS helps to limit the current through the device (without need for explicitly enforcing a compliance current parameter I_{comp} within the code). It can be seen that during write operation, first the filament grows vertically until a height $h = L$ is reached. Only then is lateral filament growth is initiated and radius r increases as long as write pulse is applied. The increase in r is rapid at first and then flattens off gradually. Simultaneously with the change in h and r, the CBRAM resistance R_{CBRAM} decreases from its initial HRS given by

(3.2) to a LRS. The final value of the LRS is dependent on the current allowed to flow through the CBRAM model and the duration of the write pulse. During read, a low voltage pulse of 0.25 V is applied. If the CBRAM is in LRS, a read current is registered as shown in the figure. A subsequent erase operation with a negative anode voltage and similar NMOS gate voltage as shown causes the filament to first shrink laterally, causing r to reduce to r_0. This is followed by the vertical filament height decrease until h is restored to its initial h_0 and the resistance returns to HRS once again. A read operation performed after this erase operation detects negligible current flow. Thus such a model can be used to simulate multi-level resistance switching in CBRAM devices for a variety of circuit applications, as will be explored in subsequent chapters.

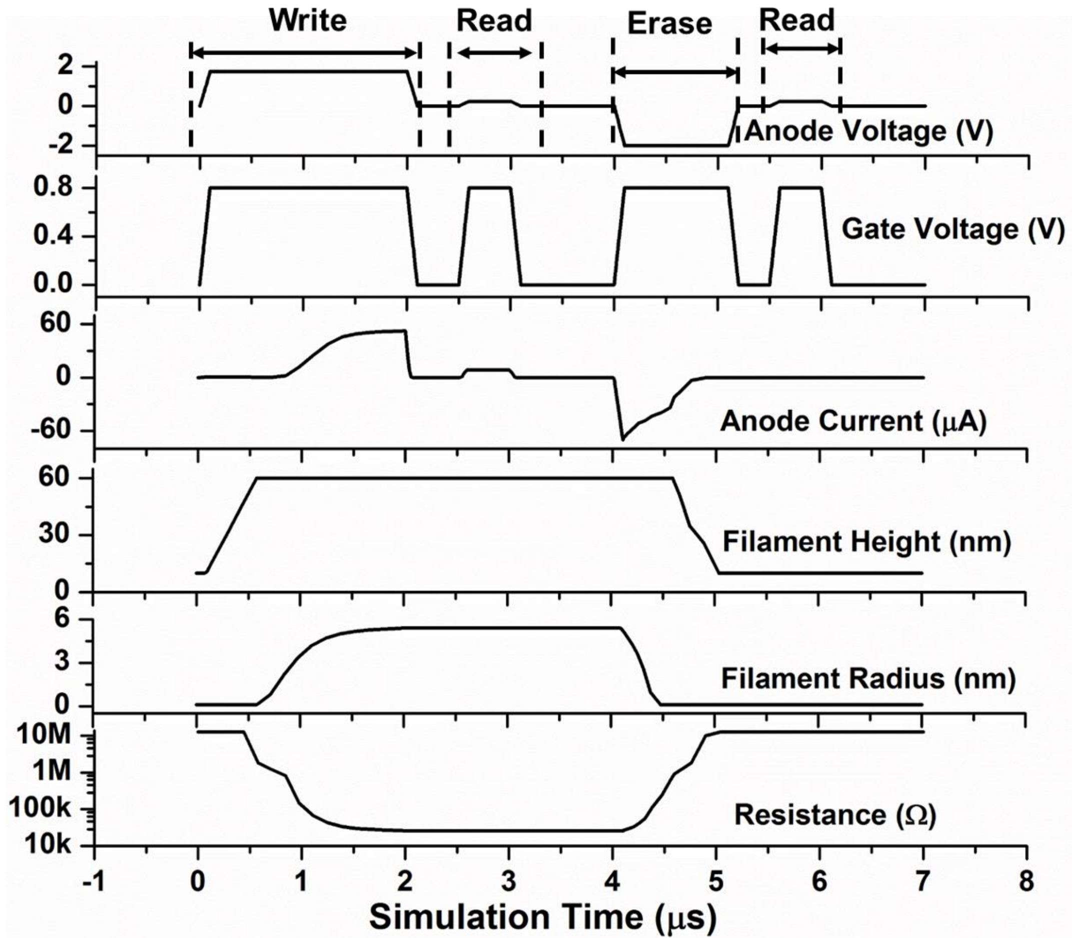


Fig. 3.2 Simulated pulsed voltage write-read-erase-read sequence for a 1T-1R simulation using the CBRAM compact model. During write, vertical filament growth occurs, followed by increase in filament radius, LRS value being determined based on the radius. No change in filament dimensions occur during read. During erase, filament radius decreases first followed by decrease in filament height, thus causing LRS to HRS transition.

3.3 Model Fitting to Experimental Data

Simulation results using this model were fit to experimental device characteristics. The experimental data was collected for Ag-$Ge_{0.3}Se_{0.7}$ based two-terminal CBRAM devices and also for Ag-GeS_2 based 1T-1R CBRAM memory elements. Fig. 3.3(a) show the fit to the experimental I-V characteristics for a two-terminal CBRAM. Fig. 3.3 (b) and 3.3(c) compare experimental and simulated CBRAM LRS and peak reset current values

respectively for different compliance currents. Finally, Fig. 3.3(d) compares the experimental and simulated LRS values and CF radius vs. applied voltage pulse magnitude.

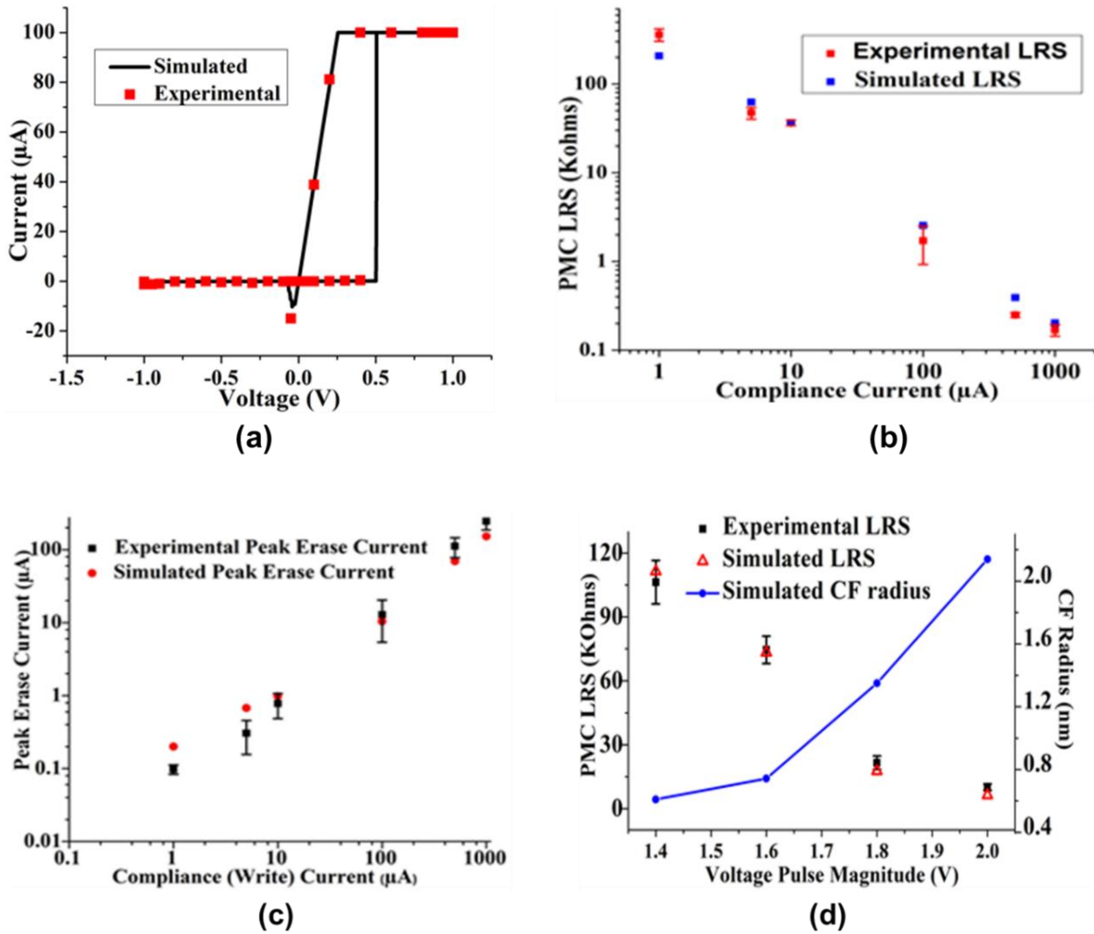


Fig. 3.3 Comparison of simulation and experimental results- (a) I-V characteristics at I_{comp} = 100μA, (b) LRS dependence on compliance current, (c) Peak reset current dependence on compliance current and (d) LRS dependence on applied voltage pulse during transient characterization and the simulated filament radius.

Commercially manufactured CBRAM memory arrays usually consist of 1T-1R elements, i.e., a CBRAM device in series with a NMOS selector switch. The Ag-GeS_2 based CBRAM devices used for characterization in this work were fabricated in back end of line (BEOL) in a 130 nm standard CMOS process [49]. Fig. 3.4(a) shows a cross-section of a 1T-1R element. Fig.3.4 (b) shows simulated I-V characteristics of a 1T-1R CBRAM

(during a single sided anode voltage sweep) for different applied gate voltages. In the case of a 1T-1R cell, the applied gate voltage magnitude determines the maximum programming current through the NMOS and consequently the LRS level during a write operation. The source is grounded during these operations.

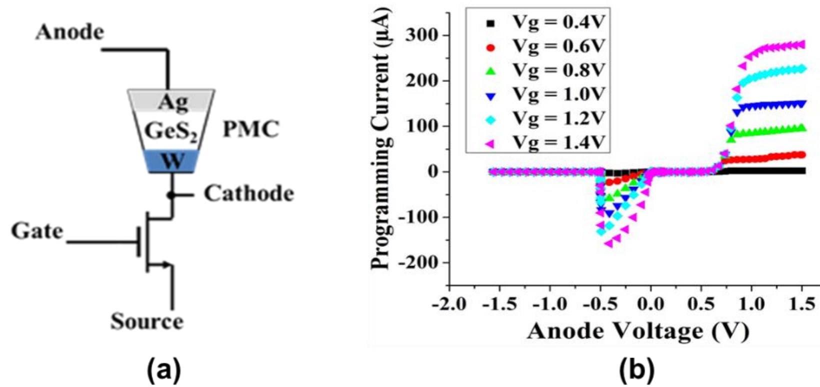


Fig. 3.4 (a) A 1T-1R CBRAM device with Ag anode, GeS_2 via layer and W cathode, (b) Simulated I-V characteristics during anode voltage sweep for the 1T-1R CBRAM with source grounded for different fixed gate voltages.

Fig. 3.5 shows experimentally obtained HRS and LRS levels during 100 write-erase operations performed by sweeping the anode voltage. Stable HRS (LRS) conditions were obtained by sweeping the CBRAM anode from 0 to +2V (-2V) while applying NMOS gate bias of 0.7V and keeping NMOS source at 0V. HRS is approximately 5MΩ and LRS is approximately 10KΩ.

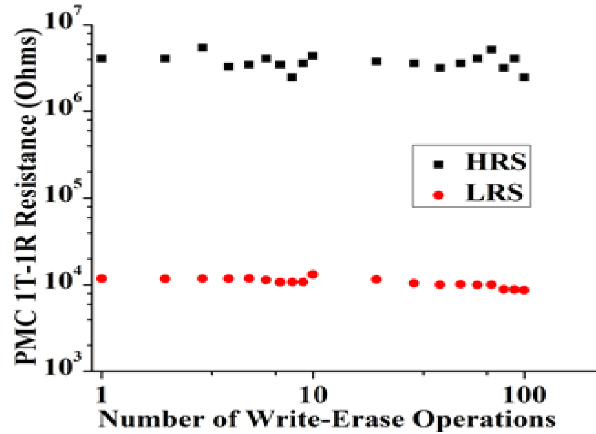


Fig.3.5 Experimental resistance cycling of 1T-1R element during 100 write-erase operations performed by sweeping the anode voltage.

The on switching threshold voltage (during HRS to LRS transition) of CBRAM devices exhibits a dependence on the applied voltage ramp rate [44], i.e., greater the rate of voltage rise the lower the switching voltage threshold. This behavior is demonstrated in Fig. 3.6(a) and the model was fit to experimental data to estimate the value for the parameter α which determines the vertical growth rate of filament from equation (1.4). The parameter β that determines the lateral or radial change in filament (from equation (1.8)) for the write operation can be extracted by fitting the model simulation results to the multiple LRS levels, as shown in Fig. 3.6(b). Fig. 3.6(c) shows programming current (NMOS selector saturation current) vs. gate voltage data for a 1T-1R memory cell (corresponding to the different LRS levels) and the corresponding fit obtained from simulation. Lastly, Fig. 3.6(d) shows comparison between simulation and experimental data for peak currents during erase for anode voltage sweep at different gate voltages, which enables extraction of the value of β during erase operation..

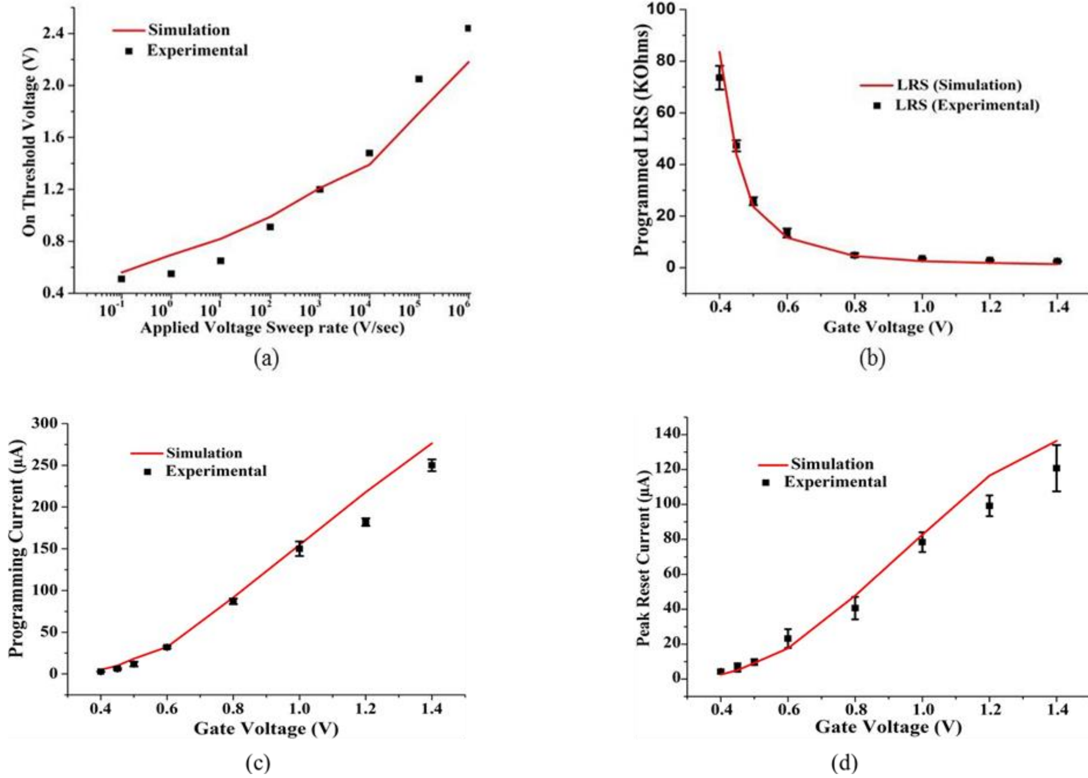


Fig.3.6 Comparison of CBRAM compact model simulation results and corresponding experimental data: (a) dependence of on switching threshold voltage (during HRS to LRS transition) on the applied voltage ramp rate, (b) multilevel LRS in 1T-1R CBRAM during anode voltage sweep for different NMOS gate voltages, (c) programming currents corresponding to the different applied gate voltages for each LRS level shown in (b), and (d) peak erase currents in 1T-1R CBRAM during erase operation voltage sweep after programming at different gate voltages.

In addition to DC response characteristics, the model was also verified for transient pulsed voltage inputs, as it is useful in circuit simulations. A comparison between a typical transient pulsed write-erase operation performed experimentally and via simulation is shown in Fig. 3.7 with the associated voltage and current transients. No current limiting resistor was used for this operation and it can be seen that write-erase times of the order of 100 ns can be achieved with such devices.

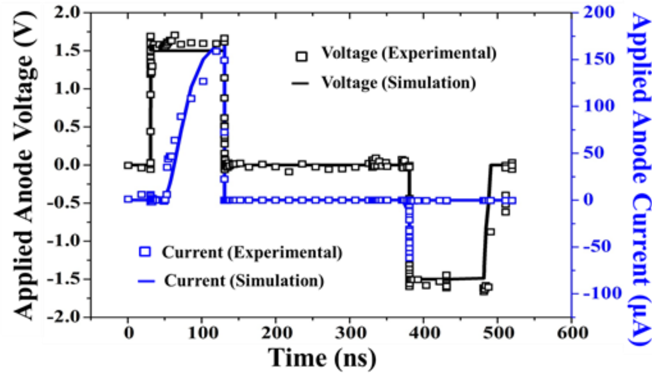


Fig.3.7 Simulated and experimentally obtained anode voltage and current transients of a two terminal CBRAM device during a typical pulsed write erase operation with cathode grounded.

Table 3.1 defines and lists the values for all the relevant optimized simulation parameters for both $Ge_{0.3}Se_{0.7}$ based and GeS_2 based CBRAM devices based on the fitting of the simulation results to experimental data described throughout this chapter. Appendix A contains the complete Verilog-a code.

TABLE 3.1 CBRAM Compact Model Simulation Parameters

Symbol	Quantity	Fitted Values	
		$Ag\text{-}Ge_{0.3}Se_{0.7}$ CBRAM	Ag- GeS_2 CBRAM
v_h	Vertical growth fitting parameter	0.5 m/s	0.5 m/s
v_r	Lateral growth fitting parameter	0.1 m/s	0.1 m/s
E_a	Activation Energy	0.3 eV	0.46 eV
L	CBRAM via thickness	60 nm	60 nm
α	Ion hopping distance	24 nm	15nm
β	Electric field parameter	0.25, V > 0, 0.22 , V < 0	0.52, V > 0, 1.4, V < 0
ρ_e	bulk electrolyte resistivity	8000 Ω.m	8000 Ω.m
ρ_f	filament resistivity	$5x10^{-4}$ Ω.m	$5.3x10^{-4}$ Ω.m
r_c	CBRAM cell via radius	2.5µm	500nm
T	Cell Temperature	300 K	300K

CHAPTER 4

SINGLE EVENT EFFECTS IN CBRAM MEMORY ARRAYS

4.1 Single Event Effects Overview

Ionizing radiation strikes can generate charged carriers in semiconductor devices (e.g. MOSFETs) [50] and lead to generation of voltage/current transients that can potentially cause effects such as loss of stored data in memory storage elements, reversal of logic states, transient malfunctioning of circuits or even permanent damage to devices or circuits [51]-[54]. All such effects can be referred to as single event effects (SEEs). The term single event upset (SEU) is used to denote cases where a SEE leads to loss of data in memory or propagation of erroneous value in logic. Studies based on scaling trends predict an increase in SEE susceptibility by approximately 40% per process node [55]. Today, single-event susceptibility has become a commonly used reliability metric for even mainstream integrated circuit products [56]. In recent decades, the study of single event effects has seen renewed interest due to a number of factors. Firstly, experimental evidence of upsets due to indirect ionization caused by protons and neutrons as well as due to direct ionization from heavy ions brought the realization that not only electronics for space missions, but also commercial non-hardened parts in near earth orbits can be susceptible to single event upsets due to higher abundance of protons over heavy ions in the environment. Secondly, with scaling down of device sizes resulting in higher clock speeds and increasingly complex functionalities in ICs, susceptibility to single event effects will only increase, possibly giving rise to new failure mechanisms and affecting even terrestrial electronic parts. Lastly, as the number of dedicated radiation hardened foundries continue to shrink drastically, electronic systems for space applications today use more and more commercial

non-hardened parts, which requires modifications to design methodologies to incorporate various radiation hardening by design (RHBD) strategies to tackle increased SEE sensitivity.

In most cases, SEUs are ‘soft errors’, i.e., errors from which the circuit can recover as opposed to permanent damage and loss of functionality in components and devices. The shape, polarity and duration of single event transients (SEEs) depend on the energy transferred from the ionizing particles as it travels through the struck material. As a charged particle travels through a material, the particle slows down due to loss of kinetic energy. This loss of energy can be divided into two components based on the manner in which it loses energy, i.e,. either through radiation emission or through Coulombic interactions. The total energy loss per unit length traversed (also called stopping power) can thus be written as,

$$\left(\frac{dE}{dx}\right) = \left(\frac{dE}{dx}\right)_E + \left(\frac{dE}{dx}\right)_N, \tag{4.1}$$

where $\left(\frac{dE}{dx}\right)_E$ is the electronic energy loss due to Coulomb interactions (i.e., the ionization), and $\left(\frac{dE}{dx}\right)_N$ is the nuclear energy loss (due to radiation emission and nuclear interactions). The term $\left(\frac{dE}{dx}\right)_E$ is better known as linear energy transfer (LET), usually expressed in the units of MeV.cm^2/mg.

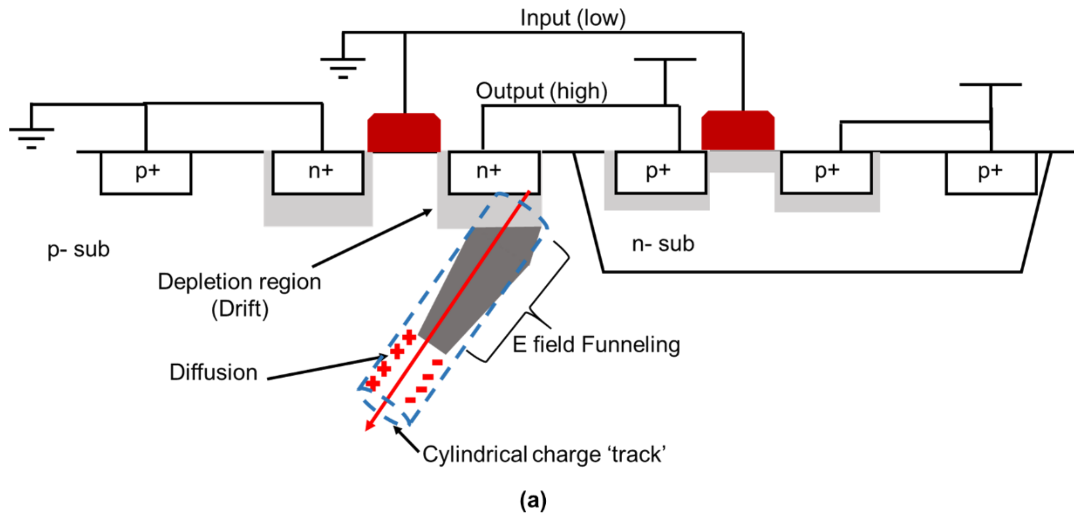
SEEs consist of two processes: charge generation and deposition within semiconductor devices due to ionizing radiation during a strike, followed by charge collection by electric field present within the device (e.g., reverse biased p-n junctions) after the strike [57]. An overview of both these processes will now be given.

a) Charge Generation: charge generation can be either due to direct ionization by the incident particle itself or indirect ionization by secondary particles created by nuclear reactions between the incident particle and the struck device. Direct ionization occurs when an energized charged particle (ion) releases electron hole pairs while traversing within semiconductor device. Such a particle comes to stop within the material after having lost all its energy and the distance travelled (referred to as range) can be estimated from its LET value. Direct ionizations are typically caused by ions with atomic number higher than 2 (also referred to as heavy ions).Indirect ionization on the other hand are usually caused by lighter particles such as protons and neutrons which may undergo inelastic collisions within the semiconductor material, leading to nuclear reactions with the target nucleus (Although in recent highly scaled technologies, direct ionization from protons have also been observed). The products of such reactions can be heavier ions which may produce enough charge carriers along their path via direct ionization to cause SEUs. Such secondary ions can be scattered in the direction of original particle, which can introduce angle of incidence dependencies.

b) Charge Collection: The generated charge carriers can be most efficiently collected by a strong electric field via the drift transport mechanism. This is why reverse-biased p-n junctions with high electric field in their depletion region are typically the most sensitive to SEEs [58]. Apart from drift within the junction depletion, diffusion can cause carriers generated outside the depletion region to travel to the junction and get collected. Such charge collection leads to the generation of a transient current 'spike' at the junction contacts (eg. drain/source in a MOSFET). Another significant aspect of the charge collection phenomenon is the collapse of junction electric field due to the highly conductive

particle ‘track’ of generated charged carriers; an effect known as ‘field funneling’. This effect can cause the electric field to collapse within the junction and extend it beyond the junction, thus enabling efficient charge collection via drift even in the substrate.

Fig. 4.1(a) shows the different processes (drift, funneling and diffusion) associated with SEEs in a CMOS Inverter cross-section. Here it is the ‘off’ NMOS drain which is most likely to be affected by a SEE. Typically, the impact of SEEs can be modeled for device and circuit simulations by a transient current source across the struck node (e.g. between drain and bulk in MOSFETs). Fig. 4.1(b) shows parameters such as rise and fall time constants used to define such a current transient model, along with the various effects that define the shape of the current pulse. The rapid rise of the current transient is due to the fast drift transport, whereas the slower decay is due to the gradual carrier diffusion transport.



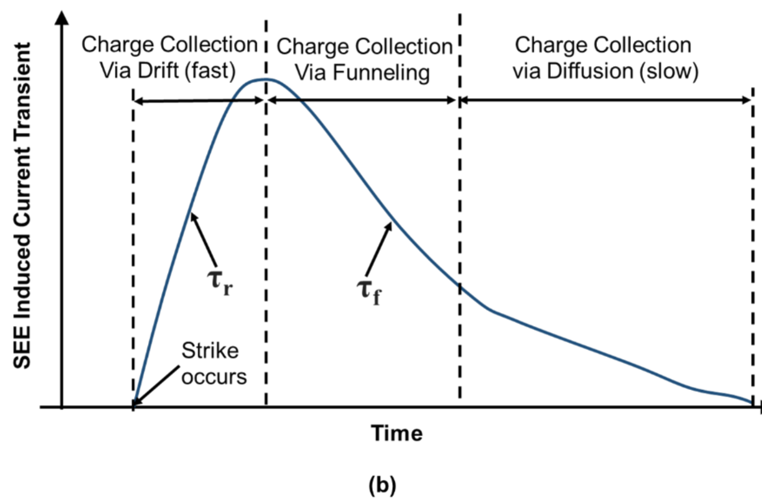


Fig. 4.1. (a) Schematic illustration of processes typically associated with single event effects using example of a CMOS inverter. (b) SEE induced current transient showing contributions from the various charge collection processes.

4.2 Radiation Effects in CBRAM Memory

The 1T-1R memory element (CBRAM device in series with an access NMOS transistor) described in previous chapters can potentially provide an alternative to CMOS based memory elements such as SRAM that is less sensitive to SEUs. CBRAM devices themselves have demonstrated high tolerance (i.e., an ability to retain programmed state) to total ionizing doses as large as 10 Mrad (SiO_2) [59], [60]. However, radiation effects in

peripheral CMOS circuits can increase their vulnerability to data loss. Experimental data have shown CBRAM-based serial EEPROM memory to exhibit data retention capability at higher total ionizing dose (TID) levels compared to flash and other types of emerging non-volatile memory devices [61].

To further investigate the effect of TID on CBRAM performance, parameters such as switching endurance and retention of CBRAM devices on a test chip were measured before and after exposure to ionizing radiation over several days [62]. In endurance testing, the ability of the CBRAM devices to undergo resistive switching between LRS and HRS repeatedly is investigated. Each cycle consists of the following sequence: write pulse, read pulse and erase pulse followed by another read pulse. Same setup as described earlier in chapter 2 for pulsed testing with a current limiting resistor is used for these cycling test resistance sampling. The goal was to test the ability to cycle these devices even after exposure to high doses of total ionizing radiation (TID). For this purpose, several two terminal CBRAM devices were initially placed in a gamma (γ) ray radiation chamber (Gamma-cell 220 in ASU) with anode and cathode contacts left floating. At periodic intervals (after a certain amount of TID exposure) some of the devices were removed, leaving the remaining devices to be exposed to higher TID. The removed devices were cycled as described above, until the HRS began to decrease. The devices exposed in this manner to gamma-ray radiation were compared to the behavior of control devices not exposed to radiation. The HRS and LRS levels recorded were between 100 KΩ and 100 Ω. In order to ensure reliability of write-erase operations, longer pulses than those normally used were applied to program the device; the write/erase pulses used for these experiments were thus ± 1.5 V, 10 ms width, while the read pulses were 75 mV, 1 ms. A typical result

for the HRS and LRS levels sampled during the cycling tests for a control device as well as devices exposed to different TID levels (maximum TID being 4.62 MRad) is shown in Fig. 4.2. For the control devices tested in this work, the maximum number of cycles achieved was generally between 1.5×10^4 to 2×10^4 cycles. As shown in Fig. 4.2, the control begins to exhibit a decrease in the HRS resistance after 10^4 cycles. After 1.5×10^4 cycles, the control shows little difference between the HRS and the LRS (i.e., failure).

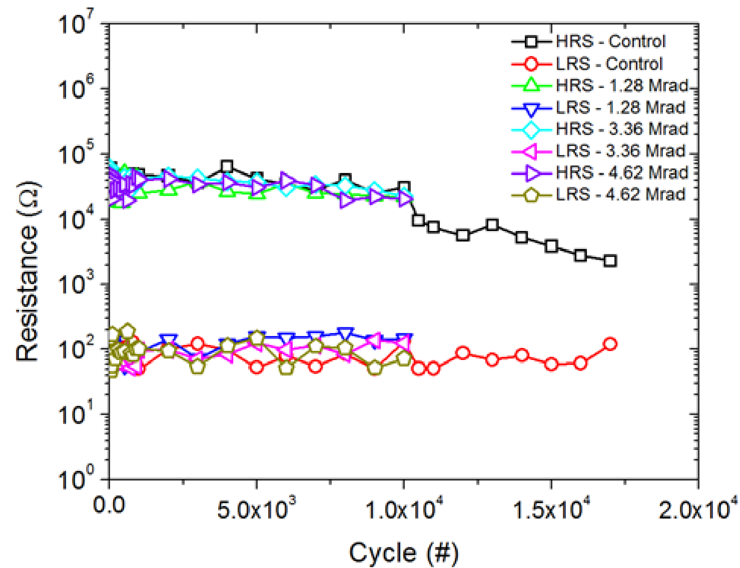
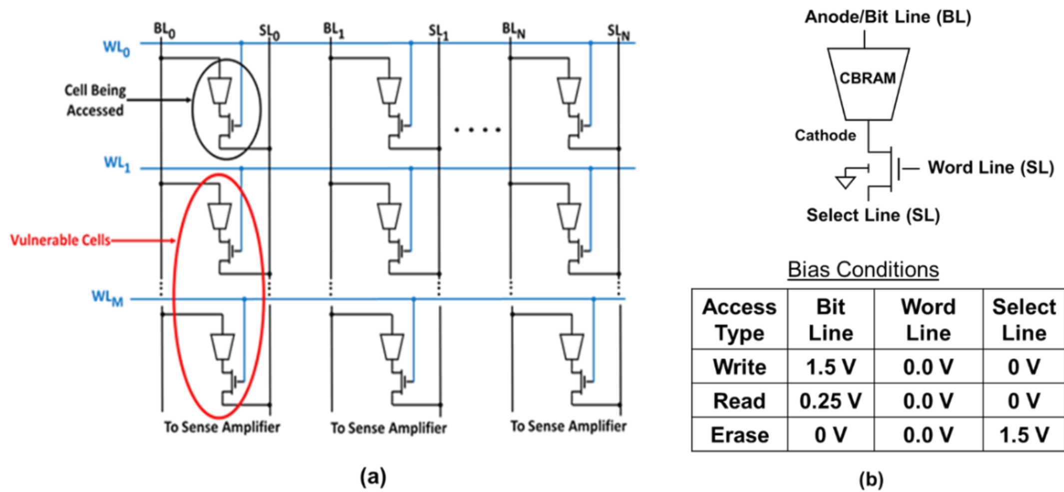


Fig. 4.2 Endurance of CBRAM devices exposed to Co-60 gamma-rays. HRS and LRS resistance as a function of the number of write-read-erase-read cycles for a control CBRAM and a CBRAM exposed to 4.62 MRad.

Although TID response on CBRAM devices have been studied in some detail, the response to single event transients is still under investigation. Recent work involving heavy ion and pulsed laser testing on an embedded resistive memory array indicated susceptibility of in the resistive memory peripheral control circuits to single-event functional interrupts (SEFIs) due to single-event upsets [63]. Fig. 4.3 (a) shows a 1T-1R array architecture. It can be seen that 1T-1R elements in the same column as an element being accessed experience the same bias on the bit line (BL) and select line (SL). Thus, all the 'inactive'

1T-1R elements that share the same bit and select lines with another element being accessed are potentially vulnerable to single event transients. Fig. 4.3(b) shows the different possible bias conditions experienced by such an inactive element during write, read and erase operations. Based on practical device operation, the programming voltage (Vdd) in this work is taken to be 1.5 V. Since the selector (access) NMOS of such an element experiences different bias conditions at the drain and source terminals during write, erase and read operations, all these scenarios need to be investigated separately to assess single event susceptibility.



Access Type	Bit Line	Word Line	Select Line
Write	1.5 V	0.0 V	0 V
Read	0.25 V	0.0 V	0 V
Erase	0 V	0.0 V	1.5 V

Fig.4.3. (a) 1T-1R array architecture with word lines (WL) driving access NMOS gates in the same row and elements in same column sharing the same bit and select lines. (b) Bias conditions in an inactive 1T-1R element (circled in (a)) that is susceptible to single event induced upsets during write, read and erase operation being performed on another cell in the same column.

In the subsequent sections in this dissertation, we focus on the sensitivity of CBRAM memory arrays to single event upsets (SEUs), specifically stored data inversions. Experimental results from heavy ion testing are presented that demonstrate data inversions in pre-programmed CBRAM devices. The results show evidence of not only HRS to LRS transitions but also of LRS to HRS transitions that have not been reported for such devices

in previous studies. In order to analyze the experimentally observed data inversions, two types of circuit simulation approaches for SEEs in CBRAM arrays are used: one based on a purely analytical model of single event transients and the other incorporating device level TCAD simulations. Based on the simulation results, a threshold incident ion LET (linear energy transfer) value for upsets are estimated under different bias conditions. Finally, SEE simulation results are also presented for another type of CBRAM memory, the crossbar array configuration. The crossbar array consisting of only CBRAM devices with CMOS drivers at the edges of the array can be susceptible to multiple event upsets unlike the 1T-1R array and analysis of such upsets will be also presented.

4.3 CBRAM Heavy Ion Testing: Experimental Results

Heavy ion strike experiments on 1T-1R test chip were performed at the Ion Beam Analysis of Materials (IBeAM) laboratory in the LeRoy Eyring Center for Solid State Science at Arizona State University, using a Cockroft-Walton gas-insulated high frequency 1.7 MeV tandem accelerator with a beam line and analysis chamber at room temperature (27 °C) [64]. The test chip consisted of a 1T-1R array of 11 1T-1R elements fabricated using methods described in Section II. After being programmed to HRS or LRS the array was subjected to a beam of a specific type of ions (H, He or O) in the accelerator. Appropriate voltage bias (discussed in more detail in section V) was applied to the array during ion beam exposure so as to investigate both HRS to LRS and LRS to HRS transitions. For this purpose a constant bias voltage of 2V was used for anode/select line. Four separate tests were performed for 1.5 MeV Hydrogen, 2 MeV Helium and Oxygen ions (2 MeV and 4 MeV) at a fluence of approximately 10^{15} ions/cm^2. After each test,

resistances of the array elements were measured by applying a dc voltage sweep from 0 V to 25mV at the anode, a constant 1.2V at the gate and 0V at the source.

Fig. 4.4 shows the measured mean resistance of all the pre -programmed HRS and LRS elements vs. incident ion LET measured after each exposure. It can be observed that when pre- programmed to HRS (≈ 5 MΩ), the mean resistance level of the cells decreased with increasing LET. After exposure to the highest measured LET (4 MeV Oxygen ions), one HRS programmed device resistance decreased to 59 KΩ. This significant change in HRS reduces the ratio of HRS to LRS to below a factor of 10 and can hence be interpreted as an upset, as the sense amplifier may not distinguish between the two states for such a low sense margin. Another HRS programmed device resistance was observed to be ≈ 161 KΩ post exposure. While the HRS to LRS ratio for this device was still > 10, such a change can leave it vulnerable to multiple event upsets (MEUs). The abrupt lowering of HRS of these devices at the highest LET causes the standard deviation from the mean HRS to increase drastically as can be seen in Fig. 4.4. Based on these results, a SEU error cross - section was also plotted, as shown in Fig. 4.4. When programmed into LRS state (10 KΩ), none of devices showed any significant deviation from pre-programmed values, as also shown in Fig. 4.4. Hence based on the experimental results, a threshold LET for an upset (HRS to LRS transition) can be predicted to be ≈ 8 MeV.cm^2/mg.

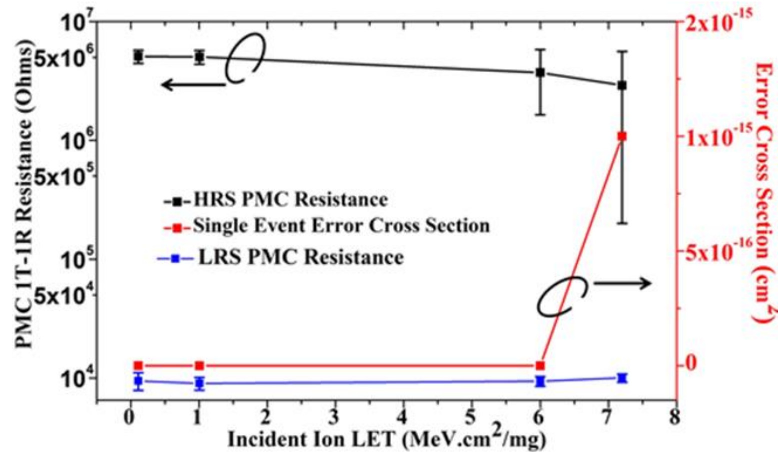
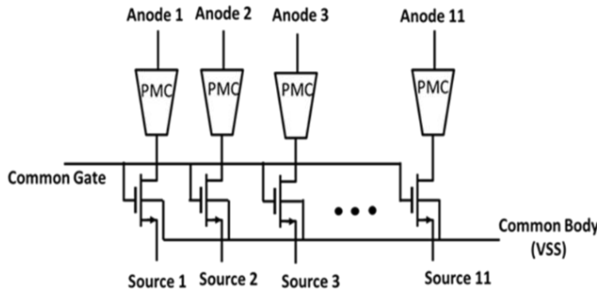
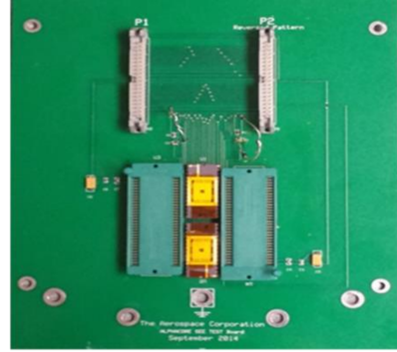


Fig. 4.4. Experimental results showing mean resistance of 1T-1R test array elements after ion strike vs. incident ion LET. HRS CBRAMs undergo decrease in resistance with increasing LET. Also shown is the corresponding SEU error cross-section vs. LET. It can be seen that LRS CBRAMs subjected to ion strike do not show any indication to revert to HRS.

Further testing was performed at a much higher LET at the 88 inch cyclotron facility in Lawrence Berkeley National laboratories [65]. An incident ion LET of 58.78 MeV/(mg/cm^2) (Xe ions) was used at a fluence of 10^9 particles/cm^2 (10^6 particles/cm^2/s flux for 1000 seconds). Of the two arrays, one array was pre-programmed to have devices in HRS and in the other array all devices were pre-programmed to be in LRS. Two sets of control devices (devices not exposed to heavy ion testing) were also programmed separately before the experiment to HRS and LRS and then measured after the test duration to verify data retention capability. Fig. 4.5(a) shows a schematic of the 1T-1R array used for collecting data, while Fig. 4.5 (b) shows the printed circuit board that was designed to hold the packaged parts in order to apply bias during programming and also during exposure.



(a)



(b)

Fig. 4.5. (a) 1T-1R array used for heavy ion testing. (b) Smartboard designed at Aerospace for electrical characterization and biased heavy ion testing.

After exposure, all devices were measured with a DC voltage sweep up to 50 mV at the anode, with gate biased at 1.4V, to measure any change in pre-radiation resistive states of each 1T-1R CBRAM. During exposure the HRS programmed array was subjected to write bias conditions (bit line voltage = 1.5 V, word line voltage = 0V, select line voltage = 0 V, from Fig. 4.3(b)), while LRS programmed array was subjected to erase bias conditions (bit line voltage = 0 V, word line voltage = 0V, select line voltage = 1.5 V. The 1T-1R CBRAM is most likely to experience a bit upset under these bias conditions, as will be discussed in detail subsequently. We can define two resistance ratio parameters as shown in (4.2) and (4.3) for write and erase bias cases. The parameter in (4.2) is a ratio of pre-strike resistance to post-strike resistance for the HRS programmed devices under write bias. The parameter in (4.3) is a ratio of post-strike resistance to pre-strike resistance for the LRS programmed devices under erase bias. For both the parameters, a value of 10 or more indicates a change in resistance by an order of magnitude or more during heavy ion testing and constitutes an upset.

$$Write\ Bias\ Resistance\ Ratio = \frac{Pre\text{-}Strike\ Resistance}{Post\text{-}Strike\ Resistance} \quad (4.2)$$

$$Erase\ Bias\ Resistance\ Ratio = \frac{Post-Strike\ Resistance}{Pre-Strike\ Resistance} \quad (4.3)$$

Fig. 4.6(a) shows the ratio values (given by (4.2)) for each of the HRS programmed devices before and after exposure, while Fig. 4.6(b) shows the same for the HRS programmed control devices. Four out of eight HRS devices experienced an upset due to a resistance decrease by a factor of 10 or more. There was no significant change in the HRS control array resistance levels as can be seen in Fig. 4.6(b).Fig. 4.7(a) shows the ratio values (given by (4.3)) for each of the LRS programmed devices before and after exposure, while Fig. 4.7 (b) shows the same for the LRS programmed control devices. Two out of eight LRS devices experienced an upset due to a resistance increase by a factor of 10 or more. As shown in Fig. 4.7(b), the control devices successfully retained their programmed LRS levels. In the following section, we analyze these two types of upsets in more detail.

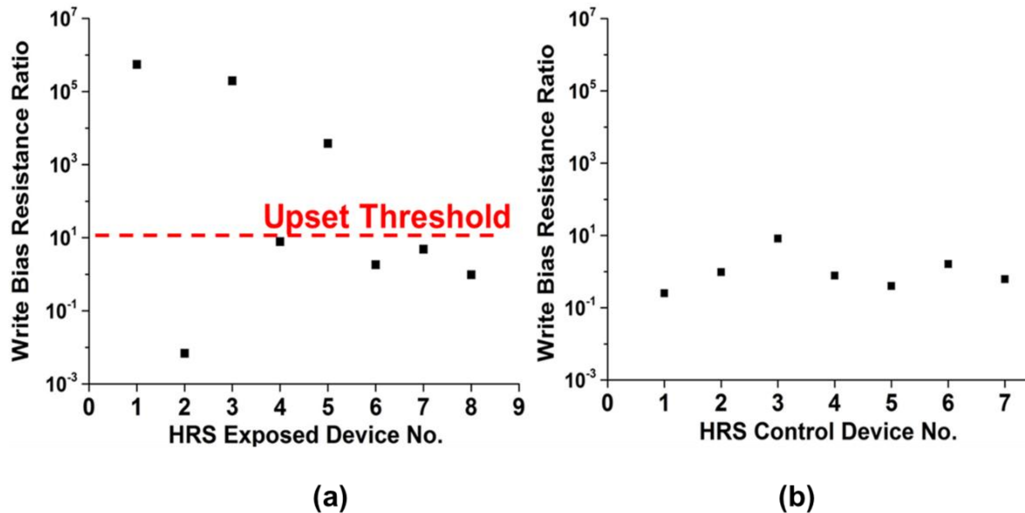


Fig.4.6. Plot showing resistance ratio parameter defined by (4.2) for (a) each HRS programmed exposed device, and (b) each HRS programmed control (unexposed) device. Devices with a ratio of 10 or more are considered to have experienced single event upset in the form of HRS to LRS transition.

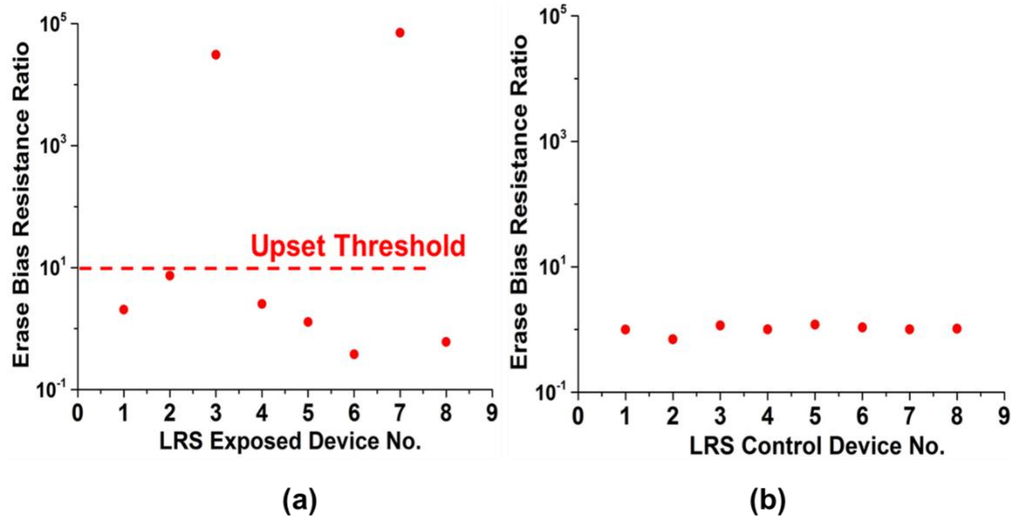


Fig.4.7. Plots showing resistance ratio parameter defined by (4.3) for (a) each LRS programmed exposed device, and (b) each LRS programmed control (unexposed) device. Devices with a ratio of 10 or more are considered to have experienced single event upset in the form of LRS to HRS transition.

4.4 Voltage Dependent CBRAM Resistance Change

As seen earlier from experimental data on transient characterization and the discussion regarding CBRAM compact model in chapter 3, the change in resistance in CBRAMs has a non-linear dependence on the magnitude and duration of voltage applied across it. This dependence can be estimated analytically as well based on the equations for dynamic filament growth rates used for developing the compact model. Assuming a cylindrical filament, the device resistance is dominated by the filament resistance (R_{CF}), which is given by,

$$R_{CF} = \frac{\rho_f . h + \rho_{e.}(L-h)}{\pi . r^2}, \qquad (4.4)$$

where ρ_f is the filament resistivity, ρ_e is the bulk electrolyte resistivity and h and r are the variable filament dimensions whose time rate of change is defined by equations (1.7) and (1.11) respectively. Hence an incremental change in filament resistance (dR_{CF}) can be

expressed in terms of incremental change in h and r by equations (4.5) and (4.6) respectively as,

$$dR_{CF} = \frac{(\rho_f - \rho_e)\, dh}{\pi . r^2}, \tag{4.5}$$

$$dR_{CF} = -2R_{CF} . \frac{dr}{r} \; . \tag{4.6}$$

Substituting *dh* and *dr* from equations (2.7) and (2.11) respectively, dR_{CF} can be expressed in terms of the applied voltage differential from anode to cathode (*V*) as,

$$dR_{CF} = (\frac{(\rho_f - \rho_e)}{\pi . r^2}). v_h . e^{\left(-E_a/_{kT}\right)} sinh\left(^{q\alpha E}/_{2kT}\right). dt, \tag{4.7}$$

$$dR_{CF} = -(2R_{CF}). \frac{v_r}{LqN_m}. e^{\left(-E_a/_{kT}\right)} sinh\left(^{q\beta V}/_{kT}\right). dt \tag{4.8}$$

Fig. 4.8 demonstrates the relationship between change in CBRAM resistance (ΔR_{CF}) and applied voltage pulse magnitude and duration via simulation using the CBRAM compact model.

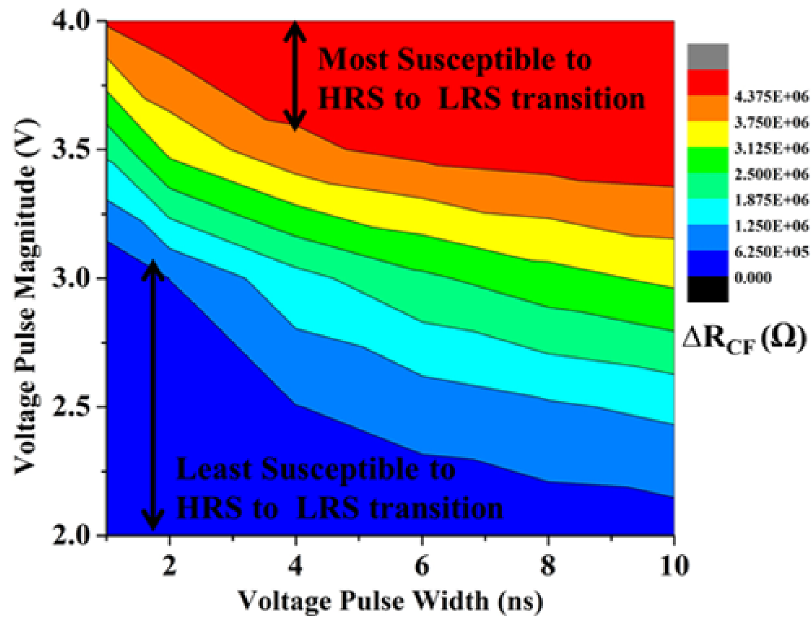


Fig.4.8 Contour plot showing simulated change in CBRAM resistance (ΔR_{CF}) vs. magnitude and duration of a voltage pulse applied across the CBRAM. A higher value of ΔR_{CF} increases likelihood of a transition to LRS.

4.5 Single Event Effect Simulation: 1T-1R Array

As discussed earlier, the MOSFET drain/source to body junction is a particularly sensitive strike location in bulk CMOS technologies [66]. Strikes in these locations can cause forward biasing of the respective p-n junction, leading to generation of drain voltage transients which can even go negative [67]. In this work, drain node of 1T-1R element is used to denote the common terminal connecting the CBRAM cathode with the access NMOS. Since strikes at this common terminal should have the maximum impact on the CBRAM state, in this work only ion strikes at the drain are considered. Two possible types of SEUs can happen: a HRS 1T-1R element changing to a LRS and a previously LRS cell changing to a HRS. As mentioned earlier, two simulation approaches will be discussed below.

i. Analytical Simulation Approach

In this approach, an analytical model for estimating the carrier generation rate during SEE and using the estimated value to compute the transient current magnitude, the current being assumed to be a pulse of constant magnitude [64]. The linear energy transfer (LET) of the charged particles striking the CMOS layer of 1T-1R memory can be expressed in terms of the total energy lost per unit path length traversed by the particle (E) as $E = LET.\rho$, where ρ is density of the irradiated material. This in turn can be used to calculate specific ionization (SI), i.e., number of ion pairs formed per unit volume traversed by the charged particle as $SI = E/(w.\pi.r^2)$, r being the path track radius and w being a material property that indicates the energy required to generate a single ion pair. The ion pair generation rate (G) can hence be calculated as $G = SI/\tau$ (in ions/cm^3.s), τ being the ion transit time. In this way the ion pair generation rate (G) can be estimated for a given incident ion LET.

Magnitude and duration of current and voltage transients generated due to a heavy ion strike on a MOSFET depend heavily on the existing bias conditions at the time of ion strike [68]. The most likely bias condition for a SEU in a 1T-1R memory array is depicted in Fig. 4.9 (a), i.e., when a write operation is being performed on a 1T-1R cell in a row. During this operation, all cells in the same row receive a voltage VDD on the bit line and 0 V on the select line. If a cell in this row is not being written (word line is at 0 V), but its access NMOS drain is struck by an ion during this time, the resulting current transient can cause the voltage at the NMOS drain (CBRAM cathode) to drop below the initial condition (VDD) and can even go negative. Hence a net transient positive voltage difference develops across the CBRAM and makes it susceptible to a SEU. If this cell is in HRS and undergoes transition to a resistance comparable to the LRS, then data stored in the cell switches to a '1' from a '0'. Under this operating condition, an ion strike-induced transient voltage pulse generated at the drain (also the CBRAM cathode) can be modeled by the equivalent circuit shown in Fig. 4.9 (b), which consists of the CBRAM in series with the parasitic drain-bulk junction diode and an ion-induced current (modeled by an ideal current source). This current (I_{ph}) resulting from electron hole pair generation can be defined as [69]

$$I_{ph} = qAG(x_j + L_e)\,, \tag{4.9}$$

where A is the drain area (in cm^2), x_j is junction depth of the drain (in cm), L_e is electron diffusion length (in cm) and G is the carrier generation rate (in ions/cm^3.s) which can be estimated from the incident ion LET as described earlier. For a given LET, the generation rate is obtained by assuming an ion strike profile diameter of 200 nm and 5 ps duration [58], [70]. The ion strike generated current is then estimated using (4.9). The circuit

simulation is performed using the CBRAM model discussed previously (variable resistance R_{CF}) in parallel with an intrinsic CBRAM capacitance (C_P) whose value is estimated by extrapolating experimentally measured device capacitance. The diode is modeled as the drain-body junction in a 130 nm CMOS process. For a sufficiently high photocurrent, the drain (CBRAM cathode) voltage, which is initially at VDD, drops to a negative value. This generates a corresponding net positive voltage transient across the CBRAM. Fig. 4.10(a) shows a voltage transient waveform generated during simulation and corresponding transition of HRS 1T-1R element to a LRS for an LET of 8MeV.cm^2/mg. Fig. 4.10(b) shows the simulated CBRAM resistance after transient generation in the circuit of Fig. 4.9 (b) for different ion LET values. A threshold LET can be estimated when the CBRAM resistance becomes comparable to the designated LRS for the binary memory array (taken to be ~10 KΩ). Finally, it may be noted here that an increase in strike profile diameter will reduce generation rate (G), which in turn will reduce the photocurrent generated. As an example, for the simulation shown in Fig. 4.10 (a), the corresponding photocurrent magnitude is ≈ 540 µA. A 10% increase (decrease) in the nominal profile diameter can cause this photocurrent magnitude to decrease (increase) to approximately 440 µA (656 µA). On the other hand, an increase (decrease) in ion strike duration can increase (decrease) the duration of the transient voltage pulse generated across the CBRAM, which in turn can increase (decrease) the total change in CBRAM resistance.

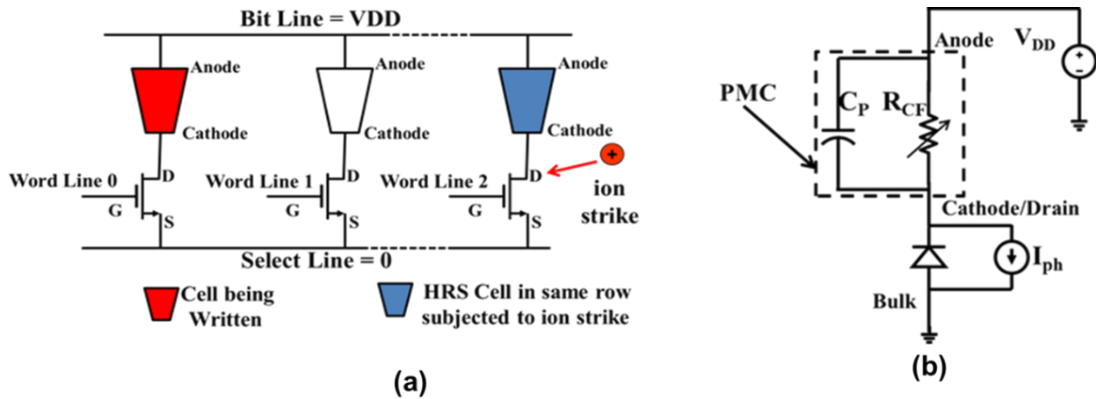


Fig.4.9 (a) Ion strike induced single event on a HRS element (Word Line 2) in 1T-1R array while another element in same row (Word Line 0) is being written, which constitutes the most likely bias condition for a possible SEU due to HRS to LRS transition and (b) equivalent circuit for the HRS cell (word line 2) to model the single event induced resistance change.

The simulation results thus show that spurious HRS to LRS transitions may occur in 1T-1R CBRAMs during write operation on another element in the same row leading to errors in stored data. It is also worth noting that the likelihood of an upset increases substantially in presence of higher positive bias (VDD) on the bit line during the ion strike, as this increases the voltage transient magnitude across the CBRAM, which in turn increases change in CBRAM resistance. For the case when there is lower bias being applied on the bit line, such as the read operation (≈ 200 mV), the threshold LET for upset is expected to be much larger due to lower magnitude of the voltage transient generated. Thus, during erase or storage mode, when the bias on bit line is 0 V, the possibility of an upset would be highly unlikely for the same reason.

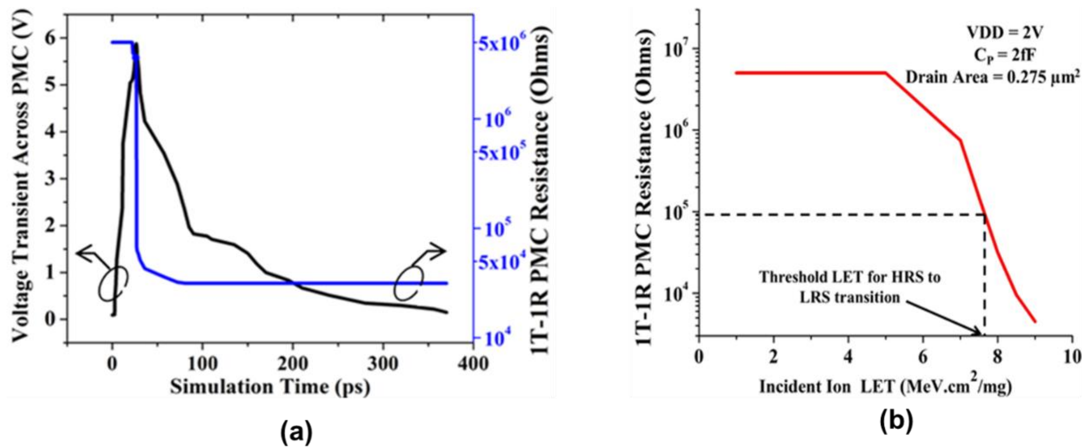


Fig.4.10 (a) A simulated voltage transient generated across 1T-1R CBRAM and its corresponding change in resistance for LET = 8MeV.cm^2/mg and (b) simulated results showing 1T-1R CBRAM resistance after ion strike event vs. incident ion LET. A threshold LET ≈ 7.5 MeV.cm^2/mg for undesired HRS to LRS transition following ion strike is indicated, assuming the minimum HRS to LRS ratio to be 10 for the devices under consideration.

There is also the possibility of a SEU characterized by the transition of an LRS 1T-1R cell to HRS. In this case, the SEU-induced transient voltage generated at the cathode will need to be more positive than the anode voltage, thus creating a net negative voltage drop across the CBRAM. Fig. 4.11(a) shows the condition where an ion strike occurs on the drain of a LRS cell, while another cell in the same row is being erased. The bit line (anode) in this case is grounded, while the select line is at VDD (2V). From the equivalent circuit of the cell under consideration shown in Fig. 4.11(b), it can be seen that only a negative voltage can be generated at the drain due to forward biasing of the drain-body junction, while the source-body junction remains reverse biased. Hence LRS to HRS transition is unlikely for such a single event.

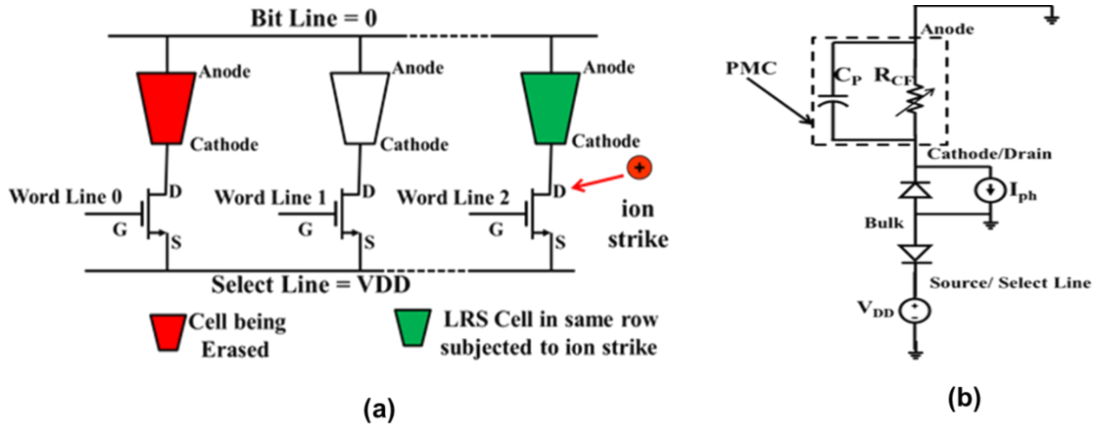


Fig.4.11 (a) Ion strike induced single event on a LRS element (Word Line 2) in 1T-1R array while another element in same row (Word Line 0) is being erased and (b) equivalent circuit for the encircled LRS cell to model the single event induced resistance change.

ii. TCAD based Simulation Approach

In the previous section, a simplified analytical model for ion-strike induced photocurrents was used to assess single event susceptibility. However, such a model predicts a LRS to HRS upset is not likely to occur, similar to conclusions in [71] for Oxide-based RRAM devices. Such a model then does not explain the LRS to HRS upsets observed experimentally for the same bias conditions as for simulation, shown earlier in the experimental results section. In order to investigate possible mechanisms that may cause LRS to HRS upsets and also to obtain accurate estimates of generated photocurrents under different bias conditions (i.e. write, read and erase), three dimensional single event simulations using a device level TCAD tool can be useful. In this section, simulations were performed on a representative NMOS structure with Silvaco 3D ATLAS with a fixed resistor connected to the drain representing the CBRAM resistance state (i.e., either HRS or LRS) prior to strike [55]. Two different cases of a single event strike at the NMOS drain (i.e., CBRAM cathode) are considered: a strike with normal incidence (case I) and an angular incidence strike (case II), as shown in Fig. 4.12. The photocurrent transients

(denoted by I_{ph}) generated at the drain terminal were then used as inputs to Cadence Spectre circuit simulator tool to perform single event simulations on a biased 1T-1R cell utilizing the calibrated CBRAM device compact model from chapter 3, as also shown in Fig. 4.12.

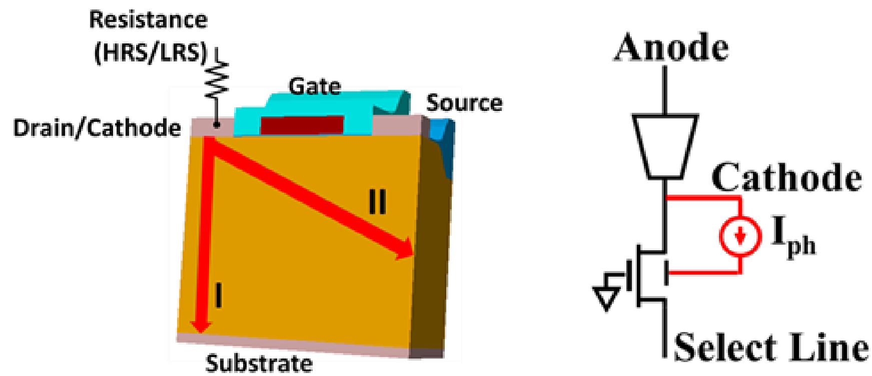


Fig. 4.12. Simulated TCAD structure with fixed resistance representing CBRAM, and corresponding circuit simulation using photocurrent (I_{ph}) from TCAD.

Upset Simulation under Write/Read Bias

As shown in Fig. 4.3 (b), write and read bias conditions are identical, differing only in the magnitude of bit line (CBRAM anode) bias (Vdd = 1.5 V >> $V_{Read} \leq 0.25$ V). Fig. 4.13 (a) shows example of photocurrent transients obtained from ATLAS device simulations for the two angles of incidence at the same LET (25 MeV.cm^2/mg) under write bias conditions. There is no significant change in the nature of the generated drain current transients for the two different angles of incidence. This is true for read bias conditions as well. Under write/read bias, the source-body junction collects very little charge, as source terminal is grounded under these conditions. Hence majority of charge is collected across the drain-body junction. Thus similar photo-current responses can be expected for the two cases when the device is under write or read bias. Fig. 4.13(b) compares the generated photocurrents for normal incidence (case I) under write and read bias conditions with the

same incident ion LET (25 MeV.cm^2/mg). Due to the lower bias during read operation, the electric field across the drain-body depletion region is weaker, which can explain the lower magnitude of photocurrent generated under read bias compared to that during write bias for the same LET. Fig. 4.13 (c) shows circuit simulations using the CBRAM compact model with the photocurrent transients from Fig. 4.13(b) as input. It can be seen that under write bias, the CBRAM cathode (i.e. NMOS drain) voltage drops from VDD to almost 0 V, thus creating sufficiently large voltage transient across the device to initiate a change in resistance from the initial HRS to a LRS of 88 KΩ. However, in the read bias case, the voltage across the CBRAM device is not sufficient to initiate a transition and the HRS remains unaffected. Hence, an upset is much more likely to occur under write bias conditions than under read bias. Finally, it can be noted that during a single event under write or read bias, the CBRAM can only experience a net positive voltage drop from anode to cathode. This suggests that only HRS to LRS transitions, not LRS to HRS switching, can occur under write or read bias conditions.

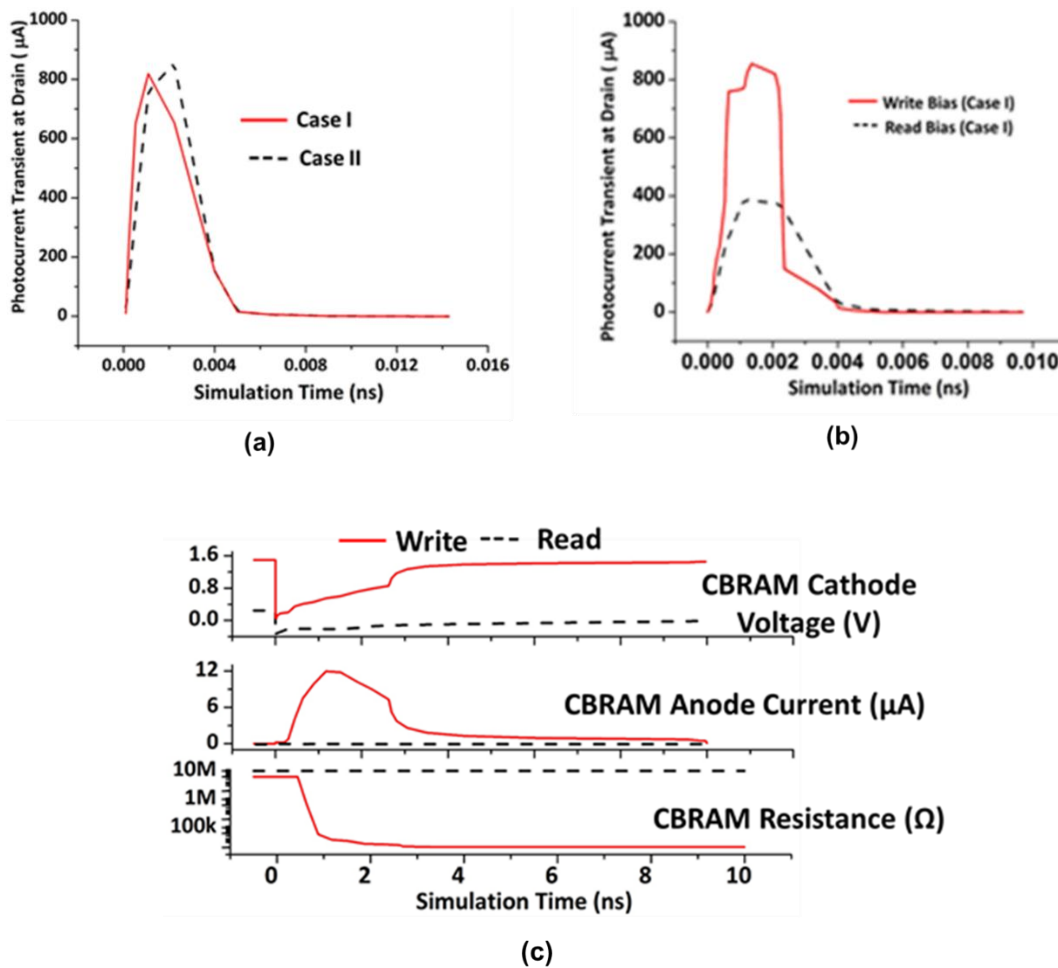


Fig. 4.13. (a) Generated photocurrent transients for cases I and II for the same incident ion LET (25 MeV.cm^2/mg). (b) Comparison of photocurrents generated under write and read bias (case I) for the same LET (25 MeV.cm^2/mg). (c) 1T-1R circuit simulation using CBRAM compact model with the photocurrents from (b). CBRAM undergoes HRS to LRS transition under write bias but not under read bias.

Upset Simulation under Erase Bias

Under erase bias conditions, the drain node of an 'idle' 1T-1R element is grounded (as the anode/bit line is grounded), while source is biased at VDD. Fig. 4.14(a) compares the photocurrent transients obtained for the two angles of incidence (I and II) for the same LET (20 MeV.cm^2/mg) under erase bias. It can be seen that normal incidence strike (case I) which are furthest away from source-body junction generates only a positive current

transient, while the angular strike (case II) generates a largely negative photocurrent at the drain, likely due to a parasitic bipolar effect caused by the forward biasing of the body-drain junction [57]. This reversal of current direction, for an angular strike, increases the probability of a sufficiently positive drain voltage transient being generated to cause an upset in the form of a LRS to HRS transition. For a normal incident strike, the generated voltage transient across CBRAM cathode is still a small positive value and hence is not conducive for a LRS to HRS transition. Fig. 4.14(b) shows the 1T-1R circuit simulation results for case II with the photocurrent transient from Fig. 4.14(a). Initially a write operation is performed to change the initial state from HRS to LRS. Due to the subsequent application of the photocurrent from Fig. 4.14 (a), a large positive voltage is generated at the cathode for case II that is sufficient to cause a reverse transition from LRS to HRS transition as shown. Thus, a LRS to HRS transition is the most likely form of upset for an inactive 1T-1R while under erase bias condition. It is worth noting here that during experimental heavy ion testing (as described in Section III), LRS to HRS upsets were observed under erase bias even though the test board was oriented normal to the ion beam direction. However it is possible that nuclear reactions occurring within the test structure during exposure can still produce recoiling ions that enter at an angle other than normal. Such angular strikes (i.e. case II) are more likely to occur at the high fluence level used during this testing (10^9 particles/cm^2), and can explain the experimental observance of LRS to HRS upsets.

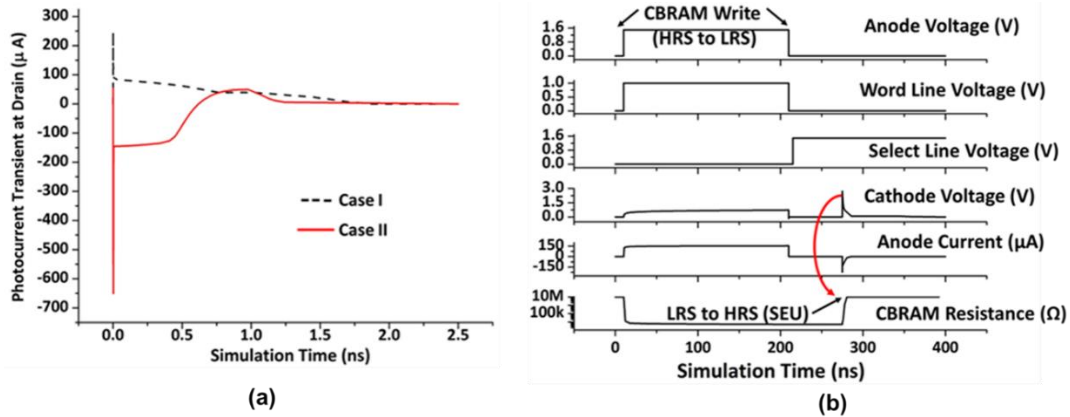


Fig. 4.14. (a) Generated photocurrent transients for cases I and II for the same incident ion LET (20 MeV.cm^2/mg) under erase bias. (b) 1T-1R circuit simulation using CBRAM compact model with the photocurrents for angular incident strike (case II) from (a). CBRAM undergoes LRS to HRS transition due to cathode voltage being pulled up sufficiently high to generate a net negative voltage across the device.

LET Dependency of 1T-1R SEUs

Using the TCAD based simulation methodology, the change in 1T-1R CBRAM resistance due to an ion strike with different LET values was estimated for the various bias conditions and incidence angles. Fig. 4.15(a) shows the simulation results for write and read bias conditions for cases I and II. There was no significant difference observed between case I and II. It can be seen that little resistance change occurs under read bias even at the highest LET, while under write bias, LETs greater than 15 MeV.cm^2/mg can be critical as the HRS decreases to below 100kΩ above the level. Fig. 4.15(b) plots the resistance change from initial LRS under erase condition. Only Case II is likely to cause LRS to HRS transitions as discussed above. A LET greater than 20 MeV.cm^2/mg can be estimated from the plot to be critical for causing such transitions. It should be noted here that due to the multilevel LRS capability of a CBRAM device, multiple-event upsets can be possible during HRS to LRS transitions, as the device may switch to an intermediate

state. However, the erase process is much more abrupt, marked by rapid filament dissolution. Hence LRS to HRS transitions are likely to occur as single event upsets.

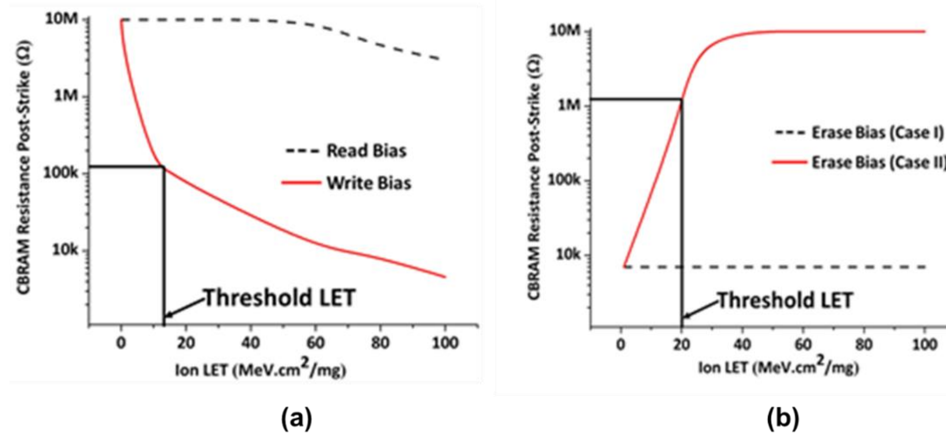


Fig. 4.15. (a) Simulated CBRAM resistance after ion strike vs. ion LET under write and read bias conditions. Initially the device is in high resistance state (HRS). (b) Simulated CBRAM resistance after ion strike vs. ion LET under erase bias condition for different angles of incidence of strike (case I and II). Initially device is in low resistance state (LRS).

4.6 Single Event Effect Simulation: Crossbar Array

In this section we focus on a crossbar CBRAM array architecture and its susceptibility to single event upsets. The crossbar CBRAM architecture consists of only CBRAM devices without selector transistors. Instead CBRAM devices on the same word line (WL) or bit line (BL) share a common driver. This configuration is attractive for the next generation of low power non-volatile memories due to greater integration density ($4F^2$) compared to 1T-1R architectures and hence is currently the focus of extensive research [72], [73]. In this section, we analyze such an array for single event upsets. Fig. 4.16 shows an MxN crossbar array architecture, with the word lines connected to the CBRAM anodes and the bit lines connected to the CBRAM cathodes. The figure shows a cell being accessed (circled) on word line 0 (WL_0) and bit line 0 (BL_0). The voltages applied on the selected lines, V_{WLS} and V_{BLS} can be set appropriately for write, erase or read operations. The voltages V_{WLNS}

and V_{BLNS} represent the applied voltages on the other word lines and bit lines to prevent disturbing the states of the unselected cells (encircled with a dashed box in Fig. 4.16) while accessing this cell. Also, all the other cells along the same word line and bit line as that of the accessed cell also experience a net nonzero voltage drop across them while accessing the selected cell. These cells are referred to as 'half-selected' cells (encircled with solid box in Fig. 4.16). In the crossbar array, single event transients generated at the output of the shared word line or bit line drivers (assumed to be inverters in this work) may affect several of these 'half-selected' cells simultaneously. The generated net voltage spike across these cells can be either positive or negative depending on whether the PMOS or the NMOS of the driver respectively, is struck by an ion. Hence both HRS to LRS and LRS to HRS transitions can occur. As the generated voltage transient propagates across all the cells along a given word line or bit line, the possibility of multi-bit upsets (MBUs) exists in such a configuration. However, due to parasitic resistances and capacitances of the wires and the loading effect of other cells, the generated spike continues to degrade as it propagates. Thus depending on the incident ion LET, different numbers of cells may experience transitions [74]. This is in contrast to a 1T-1R array where a transient generated at one element does not affect other cells due to the isolation provided by the individual access transistors.

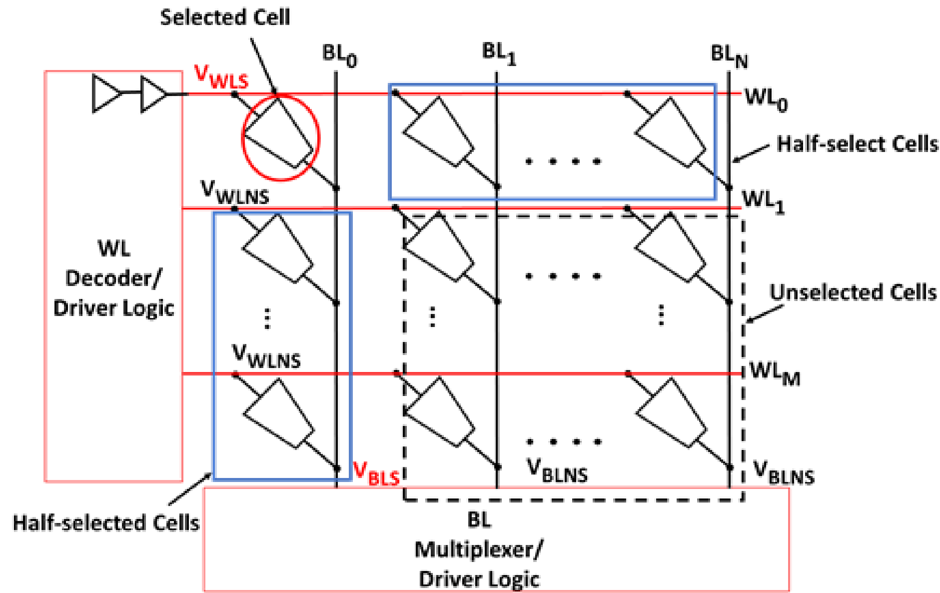


Fig. 4.16. A CBRAM crossbar array architecture with word line driving the anodes and the bit line driving the cathodes. The voltages V_{WLNS} and V_{BLNS} represent the voltages applied at unselected cells (dashed line) to prevent read disturb while accessing a cell (solid circle).

There are two popular biasing schemes used for a crossbar array: 1/2 V and 1/3 V schemes. The two schemes differ in the voltages applied on word and bit lines of unselected cells to prevent accidental disturb of other cells during read write operation on a cell in the crossbar. In the 1/2 V scheme, both V_{WLNS} and V_{BLNS} are set to VDD/2, where VDD is the programming voltage. In 1/3 V scheme, V_{WLNS} is set to 1/3 of VDD and V_{BLNS} is set to 2/3 of VDD. The 1/3 V scheme has been shown to enable higher read margin than 1/2 V scheme [75]. Fig. 4.17(a) shows a situation where a single event generated voltage spike (V_{Spike}) in the 'off' PMOS can cause a net positive voltage drop across the cells in the same word line. Depending on the biasing scheme employed, the voltage generated across each cell in the same word line can be (V_{Spike} – VDD/2) (for 1/2 V scheme) or (V_{Spike} – 2/3 VDD) (for 1/3 V scheme). Fig. 4.17(b) shows the degradation of the generated voltage

spike as it propagates along the word line. The parasitic RC values for the word lines and bit lines are chosen appropriately to correspond to a 130 nm process.

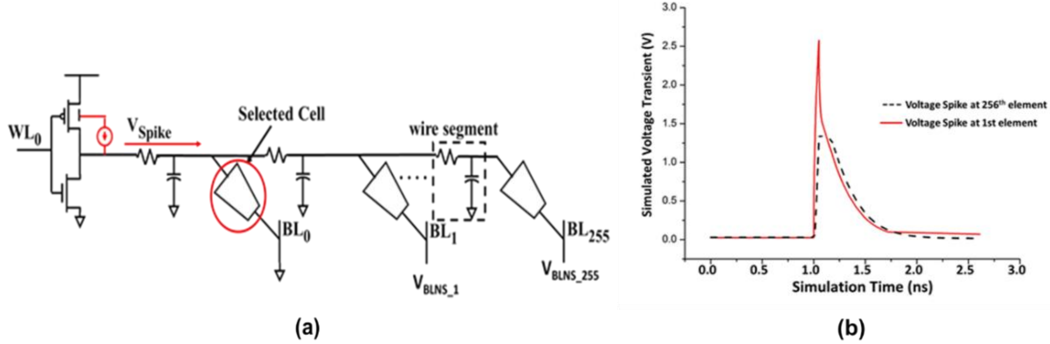


Fig. 4.17. (a) Example of single event transient propagating along the word line of a crossbar array due to drain of ‘off’ PMOS of word line driver being struck by ion. Parasitic RC of the connecting wires is considered as shown. (b) Degradation of a single event generated voltage transient as it propagates along a word line. The 256th element on the word line experiences a degraded voltage spike compared to the first element, where the spike originates.

Circuit simulations using the same CBRAM compact model used earlier are performed for both the biasing schemes to demonstrate multiple bit upsets in a 256x256 array and the results are shown in Fig. 4.18. Two representative cases are considered for comparison of the two bias schemes: 1) HRS to LRS upsets along a word line (Fig. 4.18(a)) and 2) LRS to HRS upsets along a bit line (Fig. 4.18(b)). From Fig. 4.18(a), it can be seen that the number of upsets is higher for the 1/2 V scheme compared to that for the 1/3 V scheme. This can be attributed to the larger voltage drop across unselected cells along the word line in 1/2 V scheme (V_{Spike} – VDD/2) compared to that for the 1/3 V scheme (V_{Spike} – 2/3 VDD). From Fig. 4.18(b), it can be seen that the number of upsets is higher, albeit slightly, for the 1/3 V scheme since a higher voltage drop (V_{Spike} – 1/3 VDD) exists in the 1/3 V scheme compared to that for the 1/2 V scheme (V_{Spike} – 1/2 VDD). The smaller difference shown in Fig. 4.18(b) is likely due to the more abrupt transition from LRS to HRS observed

in CBRAM devices which causes transitions to be less sensitive to differences in the magnitudes of voltage drops across the cells. For HRS to LRS transition as discussed earlier, the device can occupy several intermediate LRS values.

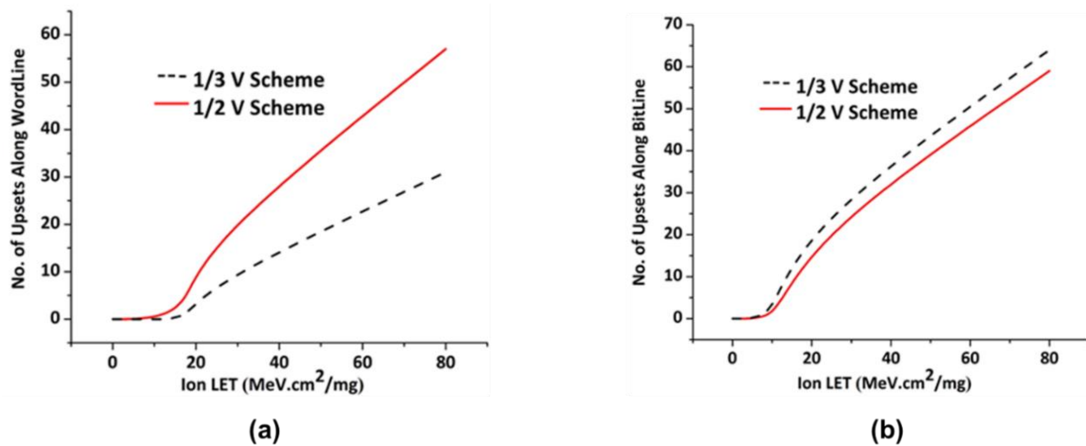


Fig. 4.18. (a) No. of simulated HRS to LRS upsets for both 1/2V and 1/3 V bias schemes along a word line in a 256x256 crossbar array. (b) No. of simulated LRS to HRS upsets for both 1/2V and 1/3 V bias schemes along a bit line in a 256x256 crossbar array.

4.7 Conclusions & Future Work

With continuous shrinking of CMOS devices, susceptibility to single event induced upsets have increased significantly to the point of being a concern for modern day integrated circuits operating at near earth orbits or even terrestrial environments. Emerging non-volatile resistive memory technologies such as CBRAM have demonstrated high tolerance towards total ionizing dose of radiation, but study of the susceptibility of such technologies to ion-strike induced transient radiation effects also needs to be investigated. In this work, experimental data from heavy ion testing performed on CBRAM devices were presented which indicate possibility of single event induced upsets in the form of stored data inversions. Evidence of low resistance state to high resistance state inversions were observed for the first time in such devices. To investigate the mechanisms behind such

upsets, two different circuit simulation and modeling approaches were presented: one based on analytical SEE model and the other based on device level TCAD simulations. Simulation results using both these approaches were presented. Two different types of architectures -1T-1R and crossbar arrays were considered and the difference in their response to single event transients was analyzed. It was found that the analytical model may not predict LRS to HRS transitions observed experimentally, while the TCAD based model showed that angle of incidence of ion strike can be an important factor on occurrence of upsets. In contrast to 1T-1R architecture, the crossbar array can be susceptible to multiple bit upsets due to absence of isolation and such upsets were simulated for two different crossbar biasing schemes (1/2 V and 1/3 V). In future, further heavy ion testing can be carried out to estimate the dependency of SEE susceptibility on angle of incidence of strike.

CHAPTER 5

NON-VOLATILE FLIP-FLOP ARCHITECTURE USING CBRAMS

5.1 Background & Motivation

Power consumption has become the main performance metric of concern for modern day CMOS circuit applications. As CMOS technology scales down, leakage power has come to dominate the overall power consumption and leakage power reduction has become increasingly important for low power applications with limited battery lifetimes such as biomedical implants or wireless sensor networks among others. Many such applications can experience prolonged idle state between active periods, during which retention of data as well as low leakage is required [76]. Embedded memory and data storage registers consume a large portion of the total power budget in low power VLSI systems [77]. For leakage reduction, a high-V_{th} NMOS transistor is used to disconnect the power supply during standby mode; a technique known as power gating. Fig. 5.1(a) shows the power gating scheme in CMOS logic circuits. When the circuit enters standby mode, the $\overline{Sleep}$ input goes low to turn the high-V_{th} NMOS off, thereby reducing circuit leakage. However, in order to retain data for eventual restoring operation after standby period, latches employing high-V_{th} transistors are required as shown. Such high-V_{th} latches still consume power at a rate proportional to standby duration. Hence this technique can only reduce leakage, but not eliminate it completely. Moreover, while high-V_{th} transistors can suppress subthreshold leakage, they do not reduce other types of leakage due to effects such as gate-tunneling and band-to-band tunneling that become more and more prominent in processes below 45 nm.

In order to overcome these issues, one possible solution is to store the data in non-volatile memory devices (instead of high-V_{th} CMOS latches) prior to standby. Resistive memory devices such as those belonging to the CBRAM/PMC family can be well suited for suppressing leakage virtually completely for low power applications, since they store data as a resistance level rather than as charge. Using such devices can remove the load on the power supply (supply voltage VDD) completely during inactive periods. Such a scheme is outlined in Fig. 5.1(b). Before the circuit enters its inactive mode, the data in all I/O flip-flops are first stored within their respective resistive memory elements. At the onset of inactive period, sleep input goes high, thus turning of power supply and ensuring virtually zero leakage.

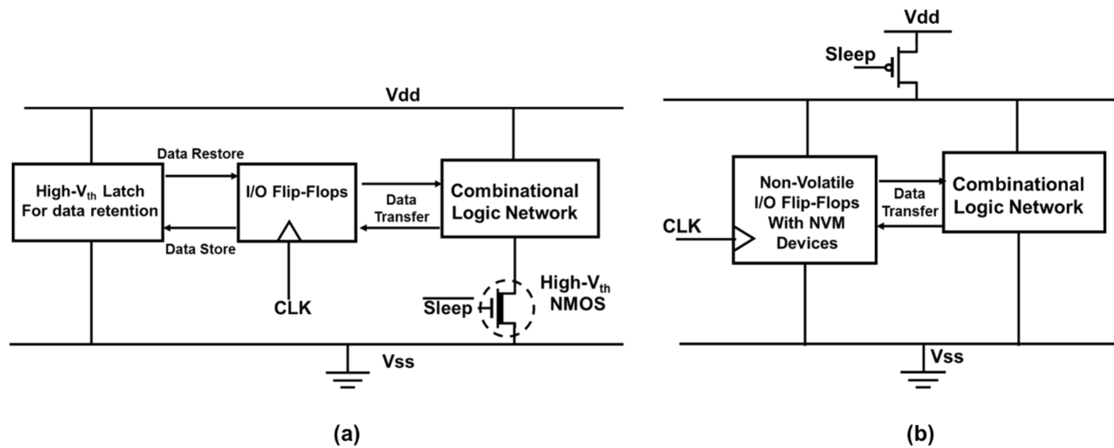


Fig. 5.1. (a) Conventional power gating scheme in CMOS logic using high-V_{th} retention latches. Subthreshold leakage is partially suppressed. (b) Alternative scheme employing flip-flops with non-volatile resistive memory elements which can store data in absence of power supply. Thus by turning off power supply completely, leakage can be completely suppressed.

Previously proposed non-volatile flip flop circuits employing emerging various resistive memory devices belonging to different non-volatile technologies have been largely derived from the conventional master-slave flip flop topology. In [78], a master-slave latch based

non-volatile flip-flop was proposed which uses complimentarily programmed magnetic tunnel junction (MTJ) memory devices to retain data. This design required the use of high power supply voltage IO circuits operating at 1.8V. Another concern for this design is the high programming currents (~200 μA) required to write and erase the MTJ devices. Also, since the difference between HRS (5.5 KΩ) and LRS (2.9 KΩ) of the MTJ devices is small, read operations performed at sub-nominal VDD levels to restore data after sleep period may not be reliable, thus increasing power consumption. In [79], a non-volatile tristate inverter-based master-slave flip-flop is proposed which uses a single ended memristor based latch for data retention in the absence of power. In this work, a fixed LRS of 1 KΩ was assumed for the memristor, leading to high current consumption during write/erase operation which required the use of an additional conditional write circuit. Moreover, since this design uses the simplified generic HP memristor model [80] which does not correspond to a specific practical device technology, accurate estimates for power consumption cannot be obtained. In [81], a type of oxide based resistive memory technology was considered, where two complimentarily programmed Al/TiO_2/Al based devices were used to store data for a master-slave flip-flop topology. This work explores the reliability of read operations to restore stored data to flip flop at sub-optimal Vdd levels unlike previous works. Such low Vdd read operations may require the incorporation of additional 'always on' dummy transistors to balance the inherent mismatch in master-slave circuit due to tristate inverters. Compact model based simulations for the ReRAM devices considered in this work were not presented. The TiO_2 based ReRAM devices used in this work demonstrate a HRS/LRS ratio less than 2, which is significantly lower than the ratios

of 100 or more demonstrated by CBRAM technology. Multi-level programming capability of the TiO_2 devices and its possible impact on the flip-flop performance was not explored.

In this work [82], a non-volatile flip flop is proposed which is based on the sense amplifier flip flop architecture and uses the conductive bridge based resistive memory devices discussed so far. The following features about the proposed design are listed below:

1) From the device perspective, we can take advantage of the prolonged state retention capability, very high HRS to LRS ratio and also commercially demonstrated BEOL integration feasibility [49] of CBRAM devices compared to many other NVM technologies. These device properties can also improve reliability of proposed circuit to operate at sub-threshold voltage levels.

2) The sense amplifier architecture enables a compact CMOS flip-flop design which has demonstrated high performance at sub threshold voltage levels [83] along with near-zero setup time, true single phase clocking and also reduced clock load in comparison to master slave flip-flop topology. An additional advantage of the sense amplifier architecture is its inherent symmetry in comparison to the master slave flip-flop which eliminates need for additional transistors [81].

3) The detailed physics based CBRAM compact model fitted to experimental data for CMOS BEOL integrated CBRAM devices (as discussed in chapter 4) is used to simulate non-volatile flip-flop operation and make reasonable estimate of flip-flop performance metrics such as power consumption and reliability. The effect of the multi-level LRS programming capability demonstrated by CBRAM devices on flip-flop performance is also explored in this work.

4) Circuit techniques to improve reliability of read operation under sub-nominal or even sub-threshold VDD levels are explored and extension of such a circuit topology to computing applications beyond data storage such as threshold logic function computation is discussed.

Various aspects of the proposed design will now be discussed in detail in the following sections.

5.2 Non-Volatile Sense Amplifier Flip-flop Circuit Architecture

As shown in Fig. 5.2(a), the sense amplifier flip-flop generally consists of two stages: the pulse generating stage followed by a slave latch for storing the acquired data. The conventional sense amplifier flip flop combines a clocked sense amplifier first stage with a NAND-based slave S-R latch [84] as shown in Fig. 5.2(b).

The flip-flop operation is as follows: sense amplifier outputs nodes $\bar{S}$ and $\bar{R}$ are precharged to power supply (VDD) when clock signal (CLK) is low. As the clock signal goes high, depending on the input data and its complement (nodes D and $\bar{D}$), only one of nodes $\bar{S}$ and $\bar{R}$ discharges to ground, while preventing the other signal from discharging. This transition in turn triggers the slave latch and the resulting outputs (Q and $\bar{Q}$) are stored until CLK remains high. The additional NMOS with VDD as input shown in Fig 5.2(b) was incorporated in [85] to provide a discharge path to ground when input data changes after the rising edge of CLK, thereby avoiding any potential change in the flip-flop state. Although variations of this volatile sense amplifier flip-flop exist [86], [87], in this work this basic design is modified to incorporate non-volatile data storage and data restore functionality.

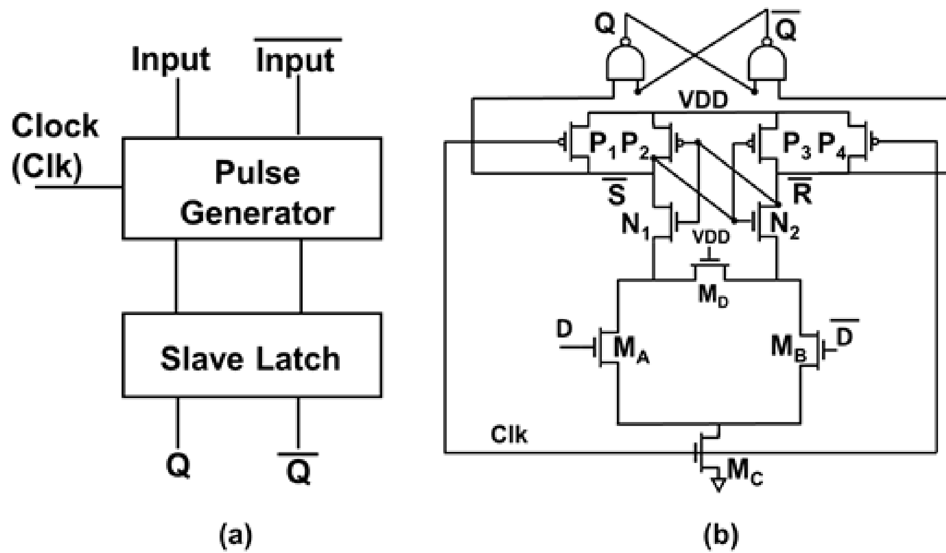
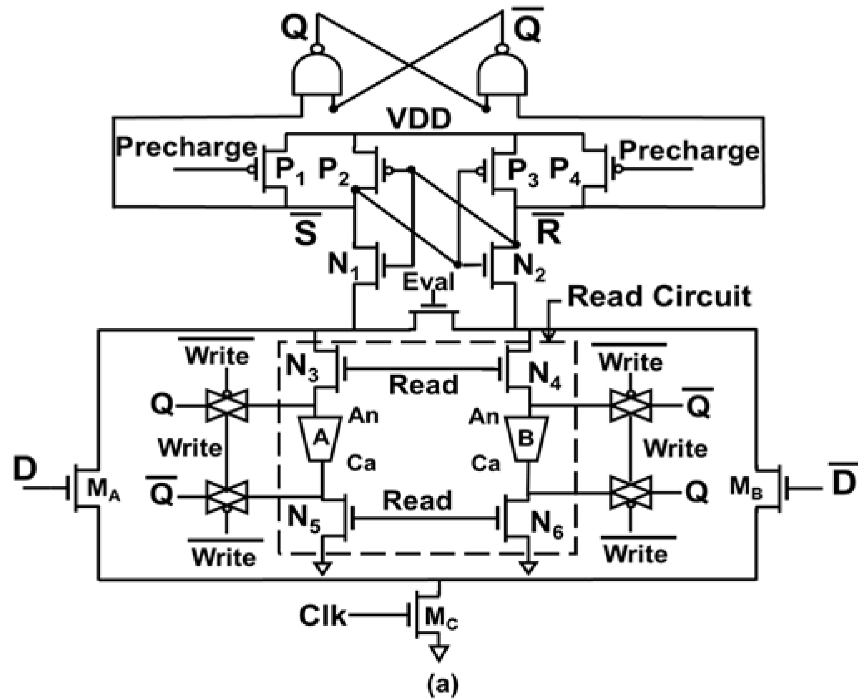


Fig.5.2 (a) General flip-flop topology with pulse generator followed by slave latch. (b) Conventional sense amplifier flip flop with NAND gate based slave SR latch.

Two potential non-volatile flip-flop topologies (called FF I and FF II for convenience) are shown in Fig. 5.3(a) and 5.3(b) respectively. Both utilize complementary resistance programming of two CBRAM devices to store the data during inactive period and then restore the saved data to the outputs by performing a CBRAM read operation upon wake-up. The mode of operation for both the flip flops shown in Fig. 5.3 are the same and can be described as follows. The flip-flop enters the normal mode of operation as the clock (CLK) goes from low to high, and acts like the conventional sense amplifier flip flop of Fig. 5.2(b). The Evaluate (Eval) signal is set to VDD during normal operation for reasons explained above. Instead of the clock (CLK) signal, a separate signal (Precharge) is used to set the nodes $\bar{S}$ and $\bar{R}$ to VDD at the rising edge of clock. When Precharge goes low, PMOS transistors P_1 and P_4 are on and thus $\bar{S}$ and $\bar{R}$ are at VDD. In preparation for storing the latch outputs (Q and $\bar{Q}$) in the CBRAM devices, the circuit enters write phase of operation (Write = VDD, Precharge = VDD, Eval = VDD). During this phase, the outputs

Q and $\bar{Q}$ are used to program the two CBRAMs in a complimentary manner (one to HRS, other to LRS) such that the correct latch outputs can be restored. At the end of the sleep period, the flip-flop goes through two phases in order to restore the stored data: a Precharge phase followed by Read/Restore phase. During the Precharge phase (Write = 0, Precharge = 0, Eval = VDD), both the nodes $\bar{S}$ and $\bar{R}$ are charged to VDD. This is followed by the Read phase (Write = 0, Precharge = VDD, Eval = 0, Read = VDD), during which as Precharge is high, nodes $\bar{S}$ and $\bar{R}$ are free to discharge to ground. The node connected to the LRS programmed CBRAM can discharge faster than the node connected to the HRS programmed CBRAM. The cross-coupled inverter latch in the sense amplifier senses the difference in voltage in the two nodes and restores the correct complimentary output states. It may be noted here that the CLK signal goes high only during normal operation and stays low otherwise, thus allowing the main pull down network (PDN) involving transistors activated by the complimentary data inputs (M_A, M_B) to operate separately from the PDN involving CBRAM read/write circuitry.



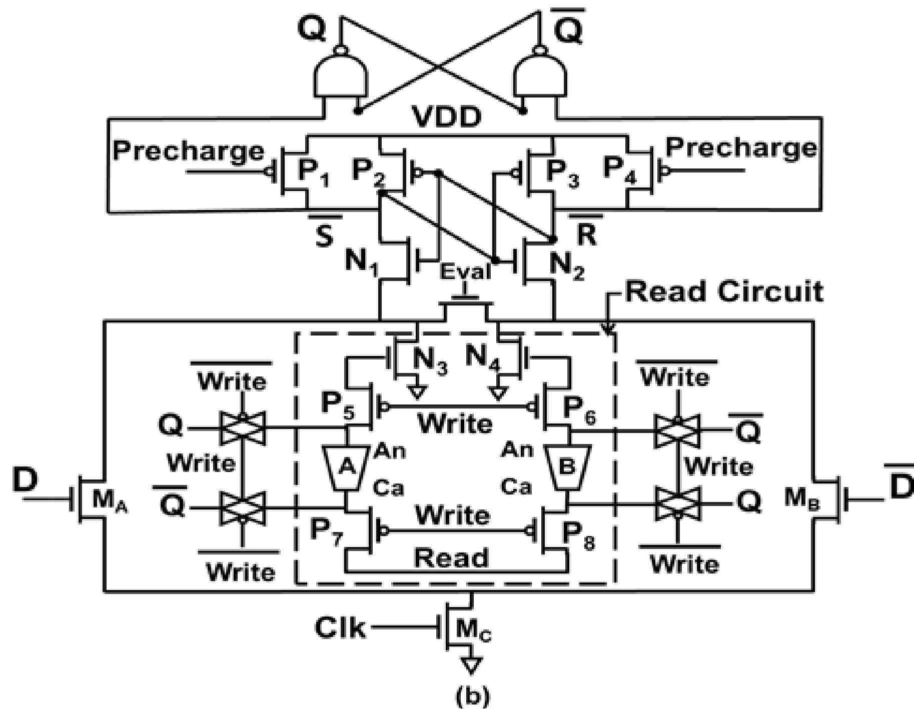


Fig.5.3. (a) Non-volatile flip flop architecture using CBRAM devices with write and read circuitry (FF I). (b) Non-volatile flip flop architecture using CBRAM devices with same write circuit as FF I but a different read circuit implementation (FF II). Notations An, Ca denote the anode and cathode terminals respectively for the CBRAM devices A and B.

Read Circuit Implementation

While differential sense amplifier based designs offer high speed and low power consumption, their operation may suffer from errors due to device mismatch and noise. Hence reading and restoring the data to the flip flop outputs is crucial to ensure reliable operation even at ultra-low voltage supply levels. CMOS sense amplifier flip-flops have already demonstrated reliable operation at sub-threshold voltage supply levels [86]. In the read circuit implementation shown in the design of Fig. 5.3(a) (FF I), the Read signal triggers both sides of the sense amplifier simultaneously through NMOS transistors N_3-N_6. Hence both the nodes $\bar{S}$ and $\bar{R}$ start discharging at the same time but at different rates determined by the difference in impedance of the discharge paths through the complimentarily programmed CBRAM devices A and B. This simultaneous initiation of the read operation is in contrast to the read circuit of Fig. 5.3(b) (FF II) where the Read signal is used to trigger the two sides of the differential sense amplifier (i.e. N_3 and N_4) separately through the complimentarily programmed CBRAM devices.

This ensures that one of the two nodes ($\bar{S}$ and $\bar{R}$) receives the Read signal earlier and begins discharging before the other, while also slowing down the initial discharge rate of the other node via the feedback action of the cross-coupled inverters. This modification can improve the robustness of the non-volatile flip-flop (FF II) over previous implementations. It should be noted here that the delay parameters associated with normal operation of flip-flop are only dependent on the original sense amplifier flip flop discharge path(M_A, M_B, M_C) and are not affected by the additional read/write circuit which remain inactive during normal data sensing. The simulated CLK-to-Q delay in the proposed design

was approximately 93 ps, which is similar to previously reported sense amplifier delay values for 130 nm technology node [87].

The static noise margin of a differential sense amplifier is a measure of its reliability and can be defined as the maximum DC voltage offset that can be applied to one side before the amplifier senses incorrectly. The static noise margin during read operation of FF II is compared with that of FF I in Fig. 5.4 for two different HRS/LRS ratios of 5 and 150 across a range of voltage supply (VDD) levels. It can be seen that FF II demonstrates much higher read noise margin than FF I in all cases.

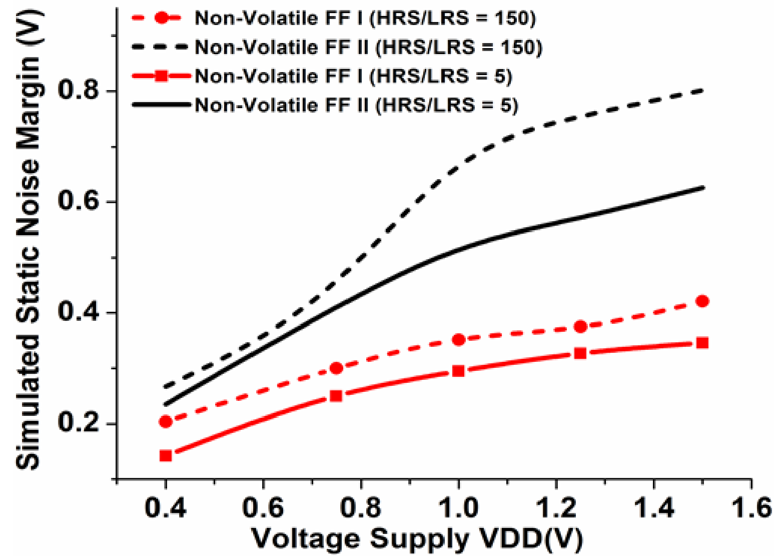


Fig.5.4. Comparison of static noise margins during read operation for FF I and FF II vs. supply voltage (VDD) for HRS/LRS ratio of 150 and 5. FF II shows superior noise margin and hence greater tolerance to signal noise.

Write Circuit Implementation

The active high Write signal controls the buffers (transmission gates) through which the outputs Q and $\bar{Q}$ are applied to write the CBRAM devices A and B to a HRS or LRS appropriately during the write phase. A minimum programming voltage pulse amplitude of 1.2 V was experimentally estimated for the devices considered in this work [49]. Hence

for operating VDD lower than 1.2 V, voltage level shifters can be added to the existing write circuit. Finally, it can be noted that all the control signals (Precharge, Evaluate, Write and Read) can be globally generated for all the non-volatile flip flops in the entire system. Fig. 5.5 shows the timing diagram for the different control signals in the design, while transient simulation results using the CBRAM compact model for a 130 nm process design kit demonstrating all the different phases of operation for storing and restoring data for D= '1' and D= '0'are shown in Fig. 5.6.

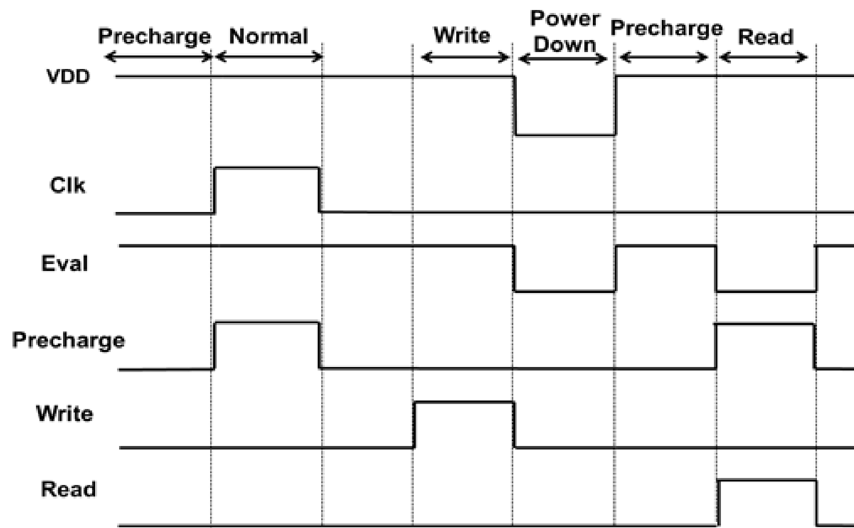


Fig.5.5 A timing sequence of different control signals and associated phases of operation of the proposed non-volatile flip-flop (both Fig. 5.3(a) and 5.3(b)).

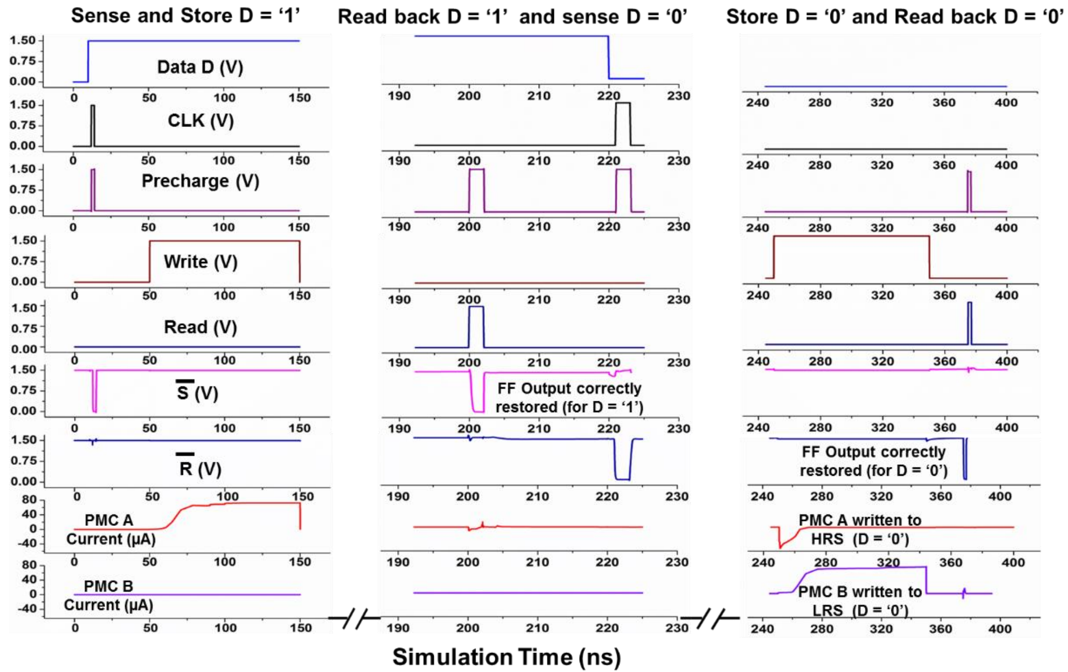


Fig. 5.6.Non-volatile flip-flop (FF II) simulation results using CBRAM compact model at nominal VDD = 1.5V. Normal data sensing followed by data write and read back operations shown for first Data D = '1' and then D = '0'. Read pulse duration is 2ns. For storing Data D= '1', CBRAM A gets programmed to LRS and CBRAM B stays in HRS. Subsequently for storing D = '0', CBRAM A is erased back to HRS and CBRAM B is now programmed to LRS.

5.3 Flip-Flop Energy & Reliability Estimation

In this section, the energy consumption and read operation reliability of the non-volatile flip-flop (FF II) of Fig. 5.3(b) are analyzed across a range of voltage supply levels. As CBRAM memory technology is still maturing, significant variations in resistance levels can be expected. While HRS levels in CBRAM devices have been found to fluctuate, a lower bound of 1MΩ seems easily achievable. On the other hand, the LRS can be accurately programmed depending on the applied bias conditions as discussed earlier. Degradation of both HRS and LRS levels can also occur in a device over time, thus lowering the sense margin during read operation. In order to estimate probability of correct read operation, Monte-Carlo simulations were performed across a range of supply voltage

levels (1.5V-0.2V) based on the MOS device process variation and mismatch parameters provided for the 130 nm design kit. A set of sub-nominal HRS to LRS ratios (5, 10 and 100) were considered. Fig. 5.7 shows the results obtained after 1000 Monte-Carlo simulations performed for each supply voltage level and HRS to LRS ratio. All transistors in Fig. 5.3 (b) were sized appropriately to improve tolerance to process variations and mismatch at low voltage supply levels. It can be seen that less than 1% failure probability can still be obtained for a HRS to LRS ratios of 10 and 100 even at VDD level of 250 mV and 200 mV respectively. Hence it can be concluded that CBRAM devices will be ideal for ensuring reliable non-volatile operation for ultra-low power operation. Another parameter of interest is the read time required for successful data restore without failure. The inset of Fig. 5.7 shows the simulated minimum read times vs. supply voltage (VDD) level. It can be seen that read time has to be increased for sub threshold operation to ensure correct read operation. The area overhead due to the read/write circuitry was approximately 1.26 μm^2 for nominal supply voltage operation and increased to 2.8 μm^2 for subthreshold operation due to upsizing of read circuit transistor sizes for improved reliability.

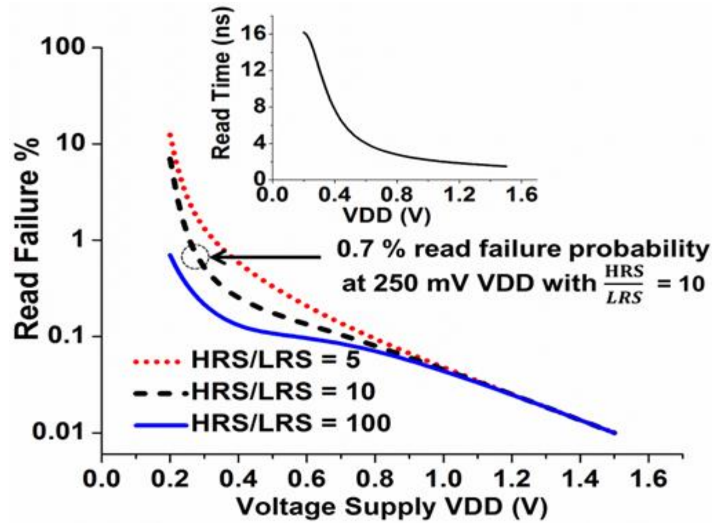


Fig.5.7 Read Failure occurrence probability (log scale) vs. supply voltage scaling for different HRS/LRS ratios. Inset shows minimum read pulse durations for successful read operation vs. power supply voltage (VDD).

To further explore the impact of CBRAM resistance level variations on the reliability of the read operation, additional Monte-Carlo simulations were performed at a subthreshold VDD of 250 mV while incorporating a Gaussian distribution of CBRAM HRS and LRS levels. As shown in Fig. 5.8, three different standard deviation (σ) values (10%, 15 % and 20 % of mean resistance value (μ)) for different mean LRS levels. For the HRS, a fixed mean and standard deviation of 2 MΩ and 25% of mean respectively are maintained in all the cases. It can be seen that even for a LRS σ/μ of 15 % at 150 KΩ, less than 1% failures occur during read operation.

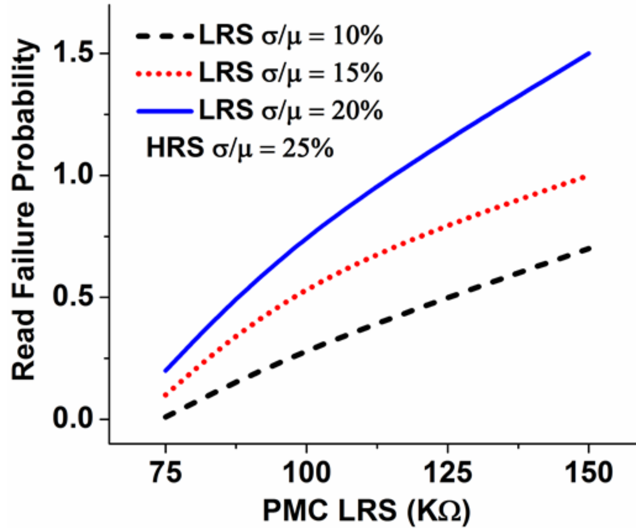


Fig. 5.8 Monte Carlo simulations of read operation assuming Gaussian distributions of CBRAM LRS and HRS levels.

The choice of the LRS level is a tradeoff between greater reliability of read/restore operation vs. higher write energy consumption. A higher LRS level can reduce the sense margin and increase read failure probability while lowering the write energy consumption and vice versa. Fig. 5.9 plots the simulated total energy consumed during normal sensing, precharge, read and write phases of operation for different programmed LRS levels against the supply voltage. The inset shows only the energy consumption during write operation on a CBRAM vs. the LRS. Programming the CBRAM to a lower LRS level increases the programming current and hence the power consumed increases. All simulations are performed using the non- volatile flip-flop of Fig. 5.3(b) (FF II). Based on above results, it can be seen that even for low HRS to LRS ratios of around 10 while taking into account CBRAM resistance variability, reliable read operation at sub-threshold supply voltages (≈ 250 mV) can still be possible, while also consuming only 372 fJ during write operation.

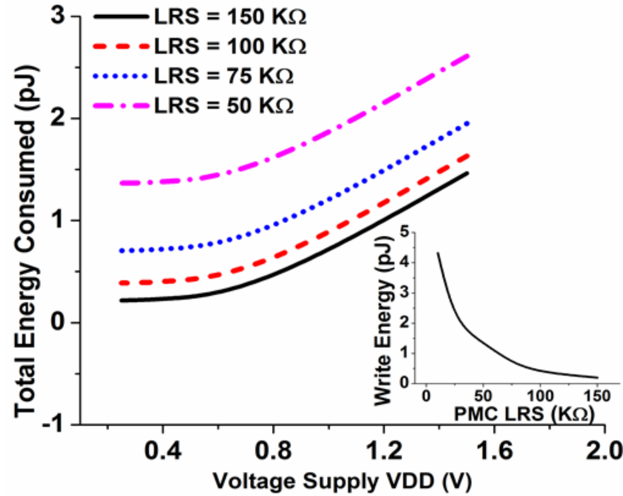


Fig. 5.9 Simulated total energy consumption vs. VDD during normal data sensing, precharge, read and write phases for write operations performed for different LRS values. Inset shows dependence of energy consumption during write operation for different LRS levels.

While the proposed non-volatile flip-flop has no leakage power consumption in contrast to conventional low-leakage CMOS-based designs, it consumes additional energy during the read-write cycle before and after the sleep period. Fig. 5.10 compares the leakage energy dissipated over time in a custom designed low leakage CMOS latch [88] with the total energy dissipated during each read-write cycle of the proposed non- volatile flip-flop (FF II) at VDD =0.4 V for different LRS levels. A minimum sleep period duration for the proposed design to be more energy efficient can then be obtained for each LRS. Thus, for example the non-volatile flip-flop becomes more energy efficient for a sleep period greater than 500 ms for a LRS of 150 KΩ during each write. Inset of Fig. 5.10 compares the total energy consumption of earlier reported design [81] with the design proposed here for two separate VDD values, 0.8 V and 0.4 V.

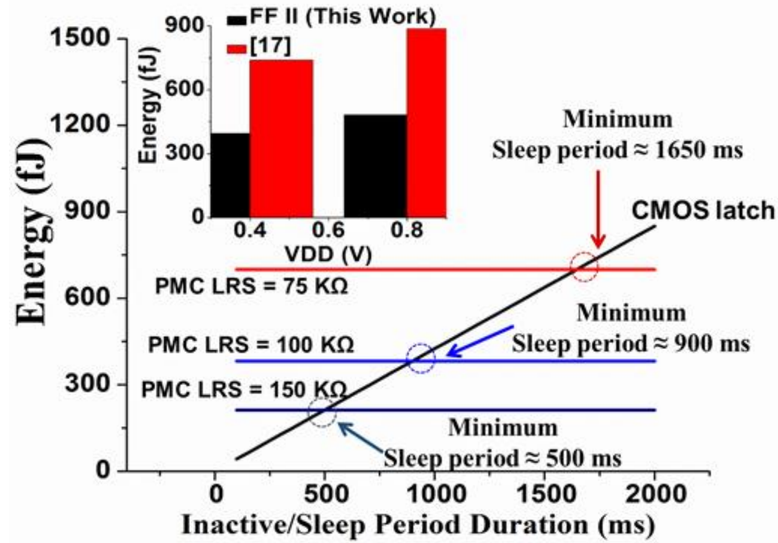


Fig.5.10 Estimation of minimum sleep period in order for proposed non-volatile flip-flop to be more energy efficient than a custom low leakage CMOS latch (black) [88]. Inset plot compares the total energy consumed for the proposed design (black) and non-volatile flip flop from [81] (red) for VDD= 0.8 V and VDD = 0.4 V.

5.4 Application for Reconfigurable Threshold Logic Computation

In this chapter, a non-volatile sense amplifier flip-flop design was proposed for ultra-low power applications that will utilize CBRAM resistive memory devices to store data during inactive period in absence of power supply, and then restore the data to the flip flop outputs when system resumes operation. Two flip-flop topologies with different data restore (read) circuits were presented and compared for robustness in noise margin. A Verilog-A compact model that fits experimental CBRAM switching behavior was presented and used to perform simulations for functionality, energy and reliability analysis of the proposed design. Since CBRAMs offer multi-level low resistance state (LRS) programmability, the effect of the choice of the LRS on the energy consumption and reliability of operation was investigated. It was concluded that even a sub-nominal HRS to LRS ratio of 10 can be sufficient to ensure reliable ultra-low voltage read operation at 250 mV power supply which is lower than previously reported non-volatile designs employing

resistive memory devices. The design proposed in this thesis has a number of advantages which can be summarized as follows. Firstly, the sense amplifier based design offers improvements in area and speed over previous master-slave latch based designs. Secondly, the proposed read circuit in this work (FF II) offers improved static noise margin and reliability in presence of process variations which may cause failure, especially at lower voltage supplies (VDD). An additional factor in improving read reliability is the much wider sense margin of the CBRAM devices used here, compared to other resistive memory devices considered in previous works.

The proposed sense amplifier based design can be extended in future for computing applications as well. As an example, we consider threshold logic functions which are a subset of Boolean logic functions (e.g. NOR, NAND) that can be computed by comparing a weighted sum of the logic inputs against a threshold value T [89]. The following inequality defines an n- input threshold logic function f,

$$f(x_1, x_2 \dots x_n) = \begin{cases} 0, & if \ \sum_{i=1}^{n} w_i . x_i < T, \\ 1, & if \ \sum_{i=1}^{n} w_i x_i \geq T, \end{cases} \tag{5.1}$$

where w_i represents the weight associated with input x_i. Such an implementation can enable more efficient computation of complex logic functions, reducing area and power consumption compared to traditional static CMOS designs [90].

Although a number of different physical implementation approaches have been used to realize threshold logic functions in hardware, the most popular implementations use a differential sense amplifier to compute and compare the two sides of the inequality of (5.1) for a given threshold function, as differential implementations remove static power dissipation and can quickly sense small differences in voltages or currents. The general

concept of a differential threshold logic gate [90] is shown in Fig. 5.11. It consists of the differential sense amplifier connected on each side to two banks of parallel transistors which receive the logic inputs of the threshold function. These sets of parallel transistors are thus commonly termed input network. Since a sense amplifier circuit cannot distinguish between two equal quantities, the inequality of (5.1) needs to be modified to remove the equality condition for physical implementation. Moreover, the number n should also be an odd integer and not even to prevent equality. Depending on the value of the logic inputs (i.e. high or low), one side of the input network will offer a lower impedance path for discharge to the sense amplifier output nodes than the other when clock (CLK) is high Thus in Fig. 5.11 if total resistance of the left hand side of input network transistors (R_{LHS}) is less than that of the right hand side of the network (R_{RHS}), node N_x will be able to discharge faster than node N_y and prevent latter from discharging to ground, and as a result output f will go high while $\bar{f}$ stays low. In the reverse case, $\bar{f}$ goes high and f stays low. Hence the threshold function output and its complement will hence be generated by the differential sense amplifier. The respective weights (w_i) of the inputs are implemented as the sizes of the transistors of the input network i.e. $w_i \propto (W/L)_i$, $i = 1,..., 2n$. The greater the value of n, the more are the number of complex functions that can be implemented, but reliability also degrades with higher n.

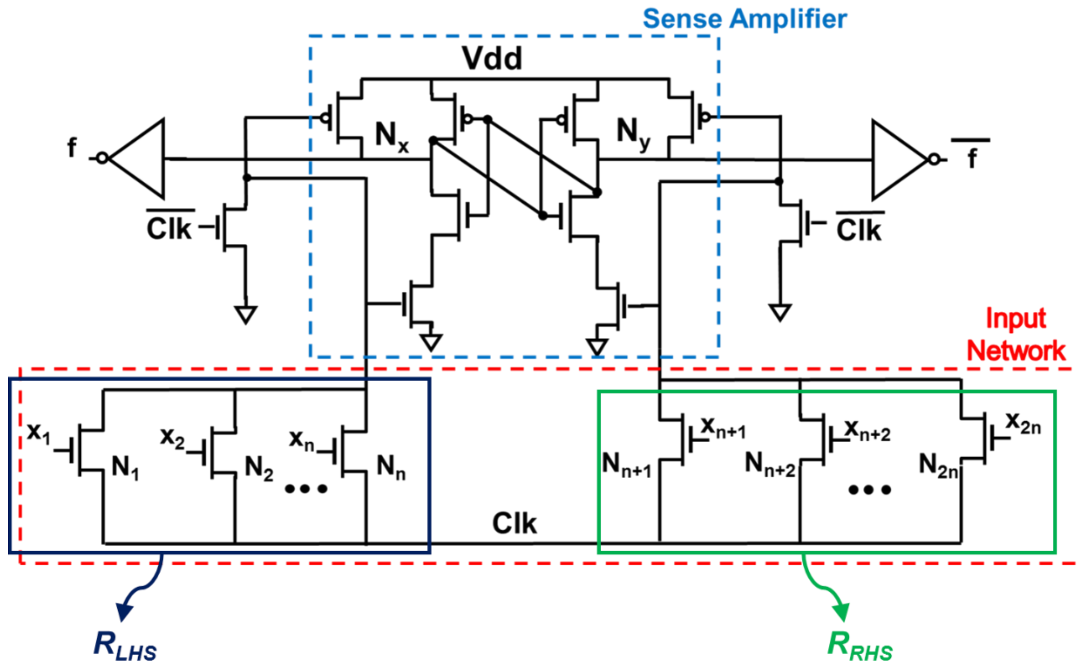


Fig.5.11 A general threshold logic gate architecture with sense amplifier to detect difference in impedance between two sets of input transistors (N_1-N_n and N_{n+1}-N_{2n}) referred to as input network.

Despite their advantages over static CMOS logic implementations such as more power and area efficient implementations of complex multi-input functions, CMOS based threshold logic gates are vulnerable to process variations and mismatch between transistors which can severely affect reliability of operation in practical hardware implementations, especially in modern processes with the scaling down of device sizes. Alternative implementations employing the programmable resistance of emerging resistive memory devices such as CBRAM as reconfigurable non-volatile weights can potentially lead to lower power computation. For this purpose, the sense amplifier based design with two CBRAMs described earlier can potentially be extended to design a multi-input threshold logic gate in which each logic input is associated with a CBRAM device. Perhaps more importantly, such hybrid implementations can potentially be more reliable, as a larger

difference in impedance can be achieved due to higher resistance levels of such devices (LRS levels reaching upto 100 KΩ or more) compared to the on resistances of modern MOSFETs (only a few KΩs). This is because larger differences in impedance can make the sensing operation more robust (larger difference in charging/discharging delay between differential branches) and help mask CMOS mismatches and variations more effectively. Lastly, using these devices as reconfigurable non-volatile weights can enable using minimum size transistors in the input network (since the input weights can be implemented using the programmable CBRAM resistance instead of the input transistor widths) , thus reducing planar CMOS footprint, as these devices can be fabricated as vertical stacks above the CMOS layer. Fig. 5.12 shows a possible implementation of such a hybrid threshold logic gate architecture based on the threshold logic latch (TLL) architecture proposed in [90]. The Clock (clk) input in this circuit is used to program the CBRAM weights to the desired LRS levels first. Different LRS levels in each CBRAM can be implemented by either controlling the gate voltage ($x_1, x_2, \ldots x_{2n}$) amplitude of the respective input network transistor or by controlling the applied pulse widths, based on the multi-level CBRAM programming behavior discussed in previous chapters. Once the CBRAM devices are programmed based on the threshold function to be implemented, logic evaluation can be performed by applying short pulses at the clock input during each evaluation cycle similar to the original TLL architecture. Thus in such an architecture, the input weights are implemented by the programmed LRS of each associated CBRAM device i.e. weight $w_i \propto 1/(LRS)_i$, $i = 1,..., 2n$.

The Program input remains high during CBRAM programming step, so as to turn on precharge PMOS devices and ensure both the complementary outputs remain at 0. During

evaluation cycle, Program is set low to enable discharging of the output inverter input nodes through the pull down NMOS transistors (N_{D1} or N_{D2}). Depending on the applied logic inputs and the programmed CBRAM resistances, one of N_{D1} or N_{D2} will turn on before the other one and hence one of the outputs Out or $\overline{Out}$ will go high and the other will stay low, thus performing the weighted sum computation. The reset input is set to ground during programming and evaluation but can be set high to erase the programmed CBRAMs before implementing a new function.

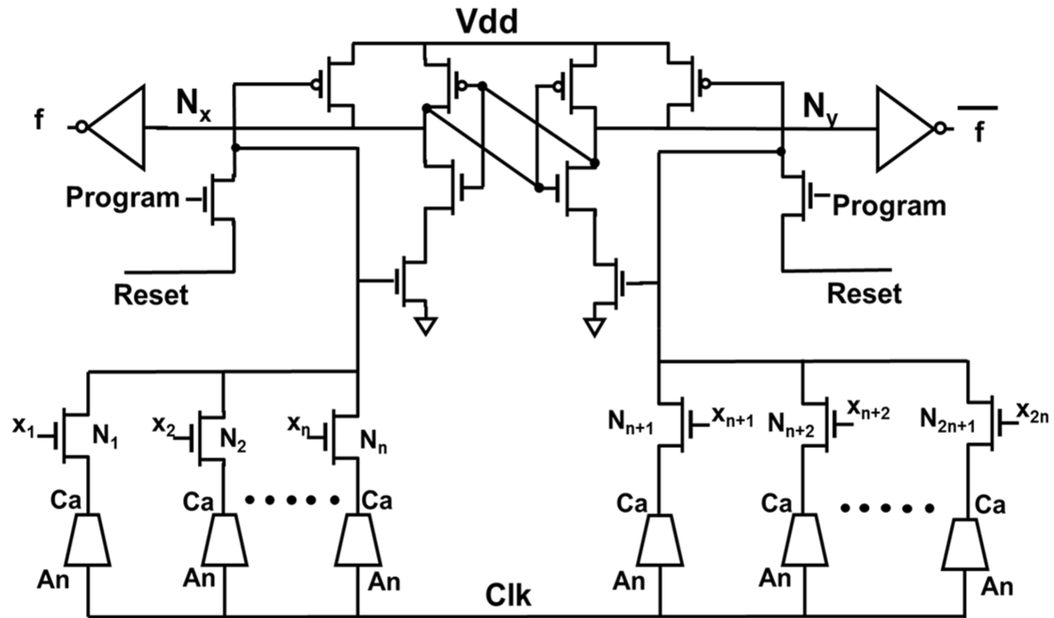


Fig. 5.12. Circuit implementation of an N-input reconfigurable threshold logic gate using CBRAMs as programmable weights. Two phases of operation exist - precharge and evaluation, similar to the flip-flop design proposed in this work. The CBRAMs are programmed to different LRS values to implement input weights.

The signal assignment in the input transistor network is an important factor in determining performance of the threshold logic gate. As mentioned earlier, the equality condition in (5.1) cannot be implemented in physical hardware since the sense amplifier cannot distinguish if both sides have the same impedance for any given input pattern. Also,

the threshold and weight values should preferably be integers for ease of implementation of CBRAM resistances. Hence for a given function, the equality condition has to be removed by mathematical manipulation of the threshold and weight values. As an example of signal assignment, let us consider the three input majority threshold function $f = ab + bc + ca$, which is the three input majority function, which is true if two or more of its inputs are true. For more complex threshold functions, such signal assignments can be performed by logic synthesis tools for design automation. Moreover there is no one unique signal assignment for a given threshold function. The corresponding inequality condition for the three input majority function can be derived from its truth table, as shown in Table 5.1 and is given by,

$$f(a,b,c) = \begin{cases} 0, & if\ a+b+c < 2, \\ 1, & if\ a+b+c \geq 2 \end{cases} . \tag{5.2}$$

In order to remove the equality condition for physical implementation, we can simply set the threshold value to 1.5 instead of 2 and then multiply both sides by a factor of 2 to make the threshold value an integer (i.e. 3), as shown below,

$$a + b + c \geq 2 \tag{5.3.1}$$

$$\Rightarrow\ 2a + 2b + 2c > 1.5 \tag{5.3.2}$$

$$\Rightarrow\ 2a + 2b + 2c > 3 \tag{5.3.3}$$

Since it is desirable to have same number of devices on each side of the input network so as to minimize mismatch [90], the above inequality can be modified to have the same number of terms (5) on each side as shown below,

$$a + b + c > (1 - a) + (1 - b) + (1 - c)\ , \tag{5.4.1}$$

$$\Rightarrow\ a + b + c > \bar{a} + \bar{b} + \bar{c} \tag{5.4.2}$$

TABLE 5.1 Truth Table For Threshold Function $f = ab + bc + ca$

a	*b*	*c*	*f(a, b, c)*
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

The above function representation can be implemented with 3 CBRAM devices on each side of input network of Fig. 5.12, along with their respective input NMOS transistors. If the minimum weight of 1 is represented by a CBRAM LRS of 60 KΩ, then since each input or it's complement has a weight of 1, CBRAM devices for each input should be programmed to the maximum LRS of 60 KΩ chosen for the design. The choice of the maximum LRS will depend on various design considerations such as operating voltage and resistance switching properties of the integrated CBRAM devices. For weights greater than 1, the corresponding CBRAM needs to be programmed to a corresponding lower value. Thus to implement an input weight of 3, the CBRAM should be programmed to 20 KΩ for a maximum LRS of 60 KΩ. A well-defined CBRAM multi-level switching behavior will thus be needed for designing robust hybrid threshold logic gates, which seems feasible based on current progress of the technology.

Fig. 5.13 shows the simulation results using the above mentioned LRS values to demonstrate the correct evaluation of outputs for the function $f = ab + bc + ca$ for different input combinations. Here output Out represents *f*, while output $\overline{Out}$ represents the

compliment of f i.e. $\overline{ab + bc + ca}$. It should be noted here that there can be many possible signal assignments for a given threshold function.

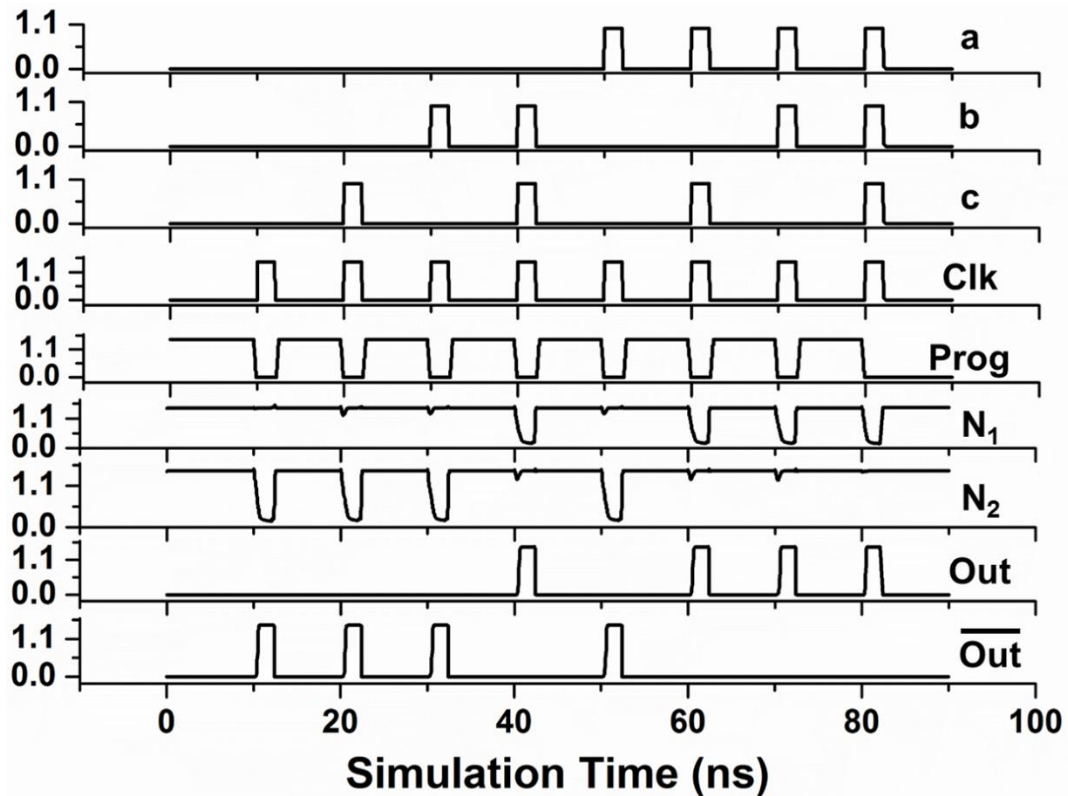


Fig. 5.13 Transient simulation showing input, output and control signals of the hybrid threshold logic gate of Fig. 5.12 during evaluation operation for the threshold function $f = ab + bc + ca$. The circuit evaluates every time clock (Clk) is high and Prog is low. Both f and its complement is implemented here at Out and $\overline{Out}$ respectively by programming the resistance of all the CBRAM devices to 60 KΩ (corresponding to a weight of 1 in (5.4.2), which is chosen as the maximum LRS level.

The critical case for assessing reliability of the differential threshold logic gate architecture discussed here occurs when the two sides of input network differ by a weight of unity. Thus in a 3input gate, the critical case corresponds to two inputs being high in one side and one input being high in the other side, and can be referred to as the [2:1] case. Fig. 5.14 shows a comparison of the number of failures in evaluating the correct outputs of a 5 input majority gate threshold function for the CMOS implementation (from Fig. 5.11) and

for the hybrid implementation (from Fig. 5.12). For worst case analysis, only the critical case [3:2] is simulated i.e. when 3 inputs are high on one side vs. two inputs being high on the other side. It can be seen from the figure that as VDD is lowered, the CMOS design begins to fail earlier and more often than the hybrid design for 1000 Monte-Carlo simulations.

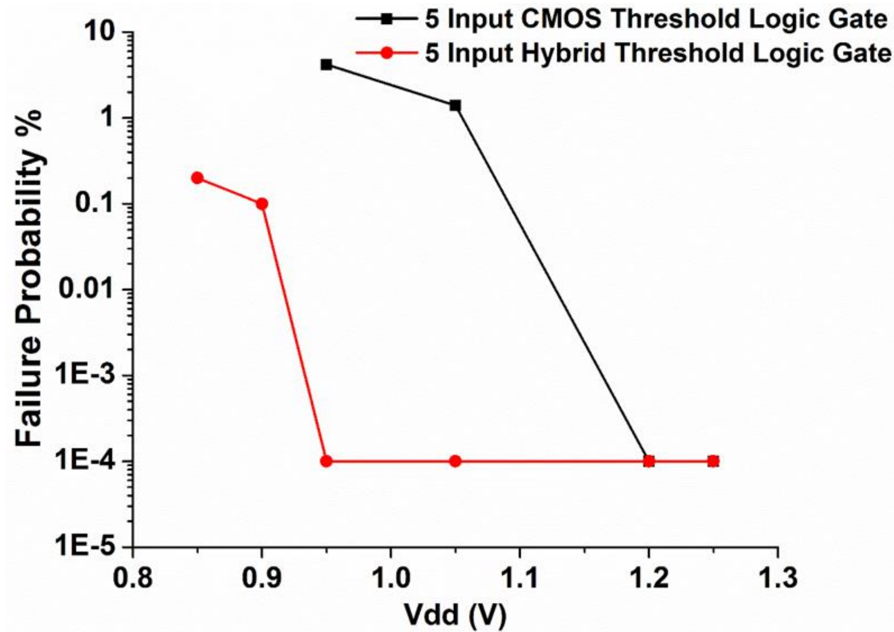


Fig.5.14 Monte-Carlo Simulation comparing the number of failures (incorrect evaluations) occurring in a CMOS only implementation vs. those occurring in a hybrid implementation of a 5 input majority logic function as power supply (VDD) is lowered. CMOS design is found to fail at higher values of VDD than the hybrid implementation.

Despite the potential benefits of the hybrid threshold logic gate implementation over the traditional CMOS implementation, the following issues need careful consideration for ensuring feasibility of using such implementations.

1) Such logic computation would require well-controlled multi-level programmability in BEOL CBRAM devices. The number of LRS levels would determine the complexity of functions that can be implemented reliably.

2) Careful analysis and modification of CBRAM parameters such as switching voltage threshold will be necessary, as the operating voltage during evaluation should be below the switching threshold (1.2- 1.4 V for pulse voltages) of CBRAM devices to avoid disturbing their programmed states

3) The CBRAM programming approach may requires further investigation to determine whether the optimal programming method lies with controlling the gate voltage amplitude during programming or a series of voltage pulses to gradually set the LRS.

4) Finally, the overhead of multi-level programming control circuitry (dependent on (3)) will need to be taken into account.

CHAPTER 6

SPIKE TIMING DEPENDENT PLASTICITY IN CBRAM DEVICES

6.1 Neuromorphic computing: Overview

Modern day computing systems are almost exclusively based on the Von Neumann computing paradigm. The key features of the Von Neumann computing architecture are sequential operation driven by a high speed clock and separation of memory and computing elements. The continuous scaling down of MOSFETs which has aided the popularity of this architecture by lowering costs while delivering more and more complex functionality is currently approaching its scaling limit, whereby quantum mechanics may begin to affect MOSFET performance and the performance benefit to scaling (Moore's law) may well reach saturation [91]. Moreover, the speed of memory access has not grown in proportion with the increase in efficiency of the computing unit, which leads to the well-known memory bottleneck problem [92]. Currently the additional latency due to memory access is compensated for with architectural modifications such as hierarchical memory system with larger on-chip cache memories and multi- core architectures among others. However, this leads to increased power consumption and also does not ensure increase in computing performance with scaling beyond the immediate future. The other concern is with the increased variability and lower reliability in smaller devices which can also limit performance scaling [93]. These concerns have led to renewed interest in the field of neuromorphic computing which aims to explore new computing paradigms by designing processor architectures that closely mimic biological neural networks [94]-[96].

Interest in neuromorphic computing goes back to early 1950's when Hodgkin and Huxley proposed the first model of a spiking neuron [97]. Since 1980s, interest in

developing electronic hardware inspired by biological neural networks increased following pioneering research by Carver Mead and others [98]. The human brain is the most complex example of a neuro-biological network which can process various kinds of sensory information in real time, performing computation intensive operations such as pattern recognition and image processing at a fraction of the power that would be consumed by state of the art processor chips. Moreover, the brain has the ability to learn and produce appropriate response by adapting to new information and unfamiliar patterns, unlike the fixed instruction set computers work from. While the Von Neumann architecture uses sequential linear processing, the brain is a massively parallel network of billions of elementary processing units called neurons. This parallel architecture has been shown to be much more tolerant to variations and defects than traditional Von-Neumann architecture [99]. Each neuron of such an architecture may be connected to thousands of other neurons via conductive pathways called synapses which are regulated based on the interaction between the neurons. The biological ability to regulate the strength of the synaptic connections between neurons is known as 'synaptic plasticity' and this is what enables brain to learn and perform complex tasks such audio-visual pattern recognition. The neurons remain inactive until they need to process sensory information unlike digital logic where nodes remain tied to a power supply voltage [100]. The low power consumption in brain is believed to be due to this event-driven analog computation in complete contrast to the clock-driven digital computation in computer processors. Another factor contributing to low power consumption and fast signal processing capability of brain is the co-existence of processing elements and memory unlike the Von-Neumann architecture. The synaptic

connections behave similar to non-volatile memory i.e. they retain their state unless modified by new neuronal interactions.

6.2 Synaptic Plasticity in Neural Networks

Fig. 6.1(a) shows the conceptual view of the elementary computation unit of an artificial neural network: the perceptron which computes a weighted sum of all its inputs (inputs x^1, $x^2 \ldots x^m$ and their respective weights w^1, $w^2 \ldots w^m$ in figure) and compares the result to a threshold value to produce a high or low output. The biological counterpart of perceptron is the neuron, as shown in Fig. 6.1(b), consisting of a bulbous cell body (soma) containing the cell nucleus, a set of filamentary outgrowths originating from the cell body in a branched fashion called dendrites and a single long branched filament called the axon. The axon serves as the output of the neuron propagating sensory information to other neurons, while the dendrites act as receivers of sensory information from axons of other neurons via ion-conducting pathways called synapses. Neurons maintain a voltage across its membrane (referred to as membrane potential), resulting from difference in concentration of various ions such as sodium, potassium, chloride and calcium within the cell. This membrane voltage can change from its resting value while receiving signals from other neurons via dendrites through the synaptic connections. If this potential increases significantly (i.e above a threshold value), then a temporary voltage spike (known as action potential) is generated within the neuron, which then travels along the axon and is transmitted to other neurons via the respective synaptic pathways. This process is often referred to as 'firing' of a neuron. Thus the synapses are analogous to the input weights of a perceptron. The action potential releases chemical compounds known as neurotransmitters that travel through synaptic pathways, get bound to chemical receptors in dendrites of connecting

neurons and modify the cell membrane permeability of these neurons, allowing influx of various ions (sodium, potassium, chloride and calcium). This leads to change in membrane potentials of all the neurons connected to the original firing neuron and the same process of generation of electrical action potentials is repeated. Through this process of electro-chemical signal transmission, the strength of synaptic connections between neurons gets modified based on their respective temporal spike generation patterns, i.e., more or less ion-conducting channels are opened or closed between the neurons. This change in synaptic strength is the basis of synaptic plasticity.

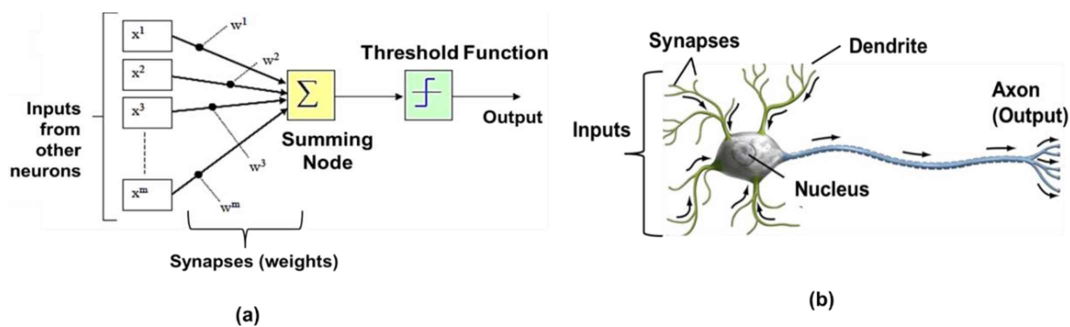


Fig. 6.1. (a) Schematic view of a perceptron, which forms the elementary unit of an artificial neural network. (b) Schematic of the biological counterpart of perceptron - a neuron with its dendrites being analogous to perceptron inputs, synapses being analogous to input weights and axon being analogous to perceptron output.

6.3 Spike Timing Dependent Plasticity (STDP) of Synapses

The dependency of synaptic plasticity on neuronal spike patterns has been observed extensively in biological neural networks and forms the basis of the learning process called spike timing dependent plasticity observed in biological neural networks, which will be discussed next. In 1949, Hebb proposed a learning rule to explain the adaptive nature of biological neural networks [101]. The Hebbian learning rule states that,

“When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.”

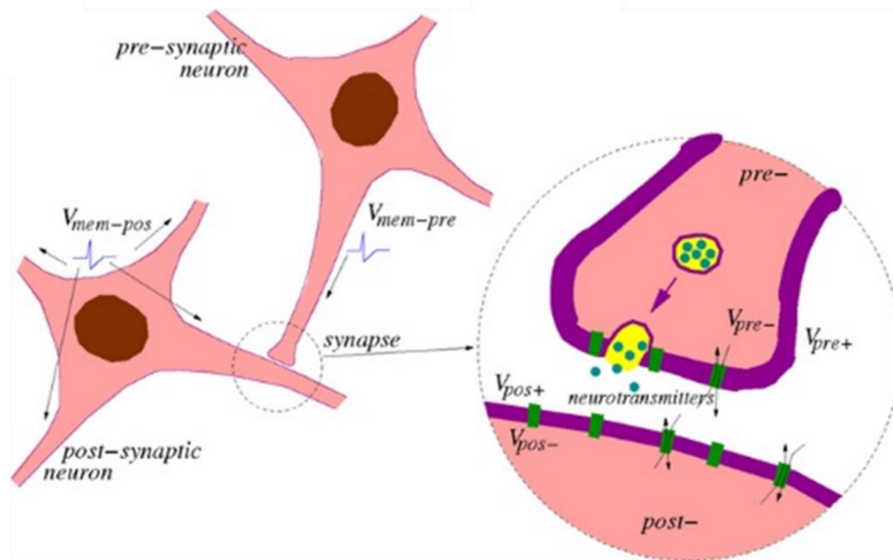
For practical implementations, the Hebbian rule can be interpreted as follows: the strength of synaptic connection between two neurons will increase if the spiking of neuron before the synapse (referred to as pre-synaptic neuron) leads to subsequent spiking of the neuron after the synapse (referred to as post-synaptic neuron) and will decrease if post-synaptic neuron spike occurs before the pre-synaptic neuron spike. In this way if a pre-synaptic neuron keeps contributing to spiking of the post-synaptic neuron, then that synaptic connection will keep on strengthening (known as long term potentiation or LTP) and if the pre-synaptic neuron remains unable to cause post-synaptic neuron to spike and the latter keeps firing before the former, then the synaptic connection weakens over time (known as long term depression or LTD). Spike timing dependent plasticity (STDP) [102]-[104] is a neuronal learning mechanism that follows from this Hebbian rule with the added emphasis on a temporal ‘learning window’, which is a period of time within which the pre and post-synaptic neurons need to fire in order for synaptic strength modification to happen. If the relative timing difference of pre and post neuron spikes exceed this timing window duration, synaptic strength remains unchanged. Whereas synaptic plasticity is based on based on relative spiking frequencies of connecting neurons according to the classical Hebbian rule, it is based on precise temporal relationship between spiking of neurons. To summarize, based on the STDP rule, the synaptic strength can change gradually. Every time pre-neuron spikes before the post-neuron, their synaptic connection strength increases and every time the post-neuron spikes before the pre-neuron, the

synaptic connection strength decreases, provided that the spikes occur within a certain time period of one another. The time difference between the neuron spiking events also determines the amount of change in synaptic strength. Synaptic plasticity can therefore be expressed as a complex function of the time difference between spiking events. Physiological evidence of STDP has been reported by several researchers in neuroscience [105] and it is now recognized as a major learning mechanism in biological neural networks [106].

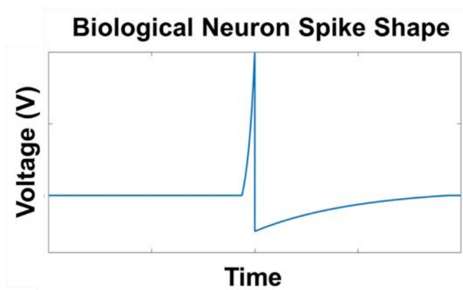
Fig. 6.2(a), reproduced from [107] shows a schematic of a synaptic pathway connecting a pre-synaptic and post-synaptic neuron. Although the exact shape of a neuron spike can vary for different types of neurons, in general such spikes are characterized by a sharp rise to a peak positive amplitude, followed by an abrupt transition to a peak negative amplitude and then a more gradual return to resting potential before the spike, as shown in Fig. 6.2(b). The synaptic strength (or weight) update function for STDP is typically denoted by the function $\xi\ (\Delta T)$, ΔT being the time difference in spiking of pre-synaptic neuron and post-synaptic neuron. Based on the experimentally obtained data for biological STDP rule for synaptic plasticity shown in Fig. 6.2(c), $\xi\ (\Delta T)$ has been mathematically modeled with the following equation [107],

$$\xi(\Delta T) = \begin{cases} a_{p.} e^{\frac{-\Delta T}{\tau_p}}, & if\ \Delta T = t_{post} - t_{pre} > 0, \\ -a_{n.} e^{\frac{\Delta T}{\tau_n}}, & if\ \Delta T = t_{post} - t_{pre} < 0, \end{cases} \tag{6.1}$$

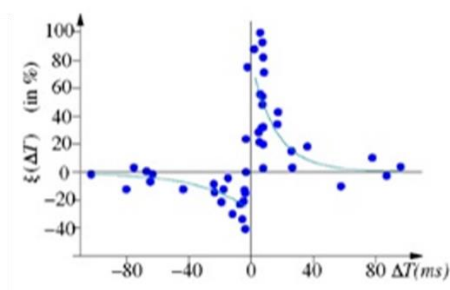
where t_{post} and t_{pre} are the time instants of occurrence of the post-synaptic and pre-synaptic neuron spikes respectively.



(a)



(b)



(c)

Fig. 6.2. (a) Two neurons connected via a synaptic pathway form the fundamental computational block of biological neural networks, reproduced from [107]. (b) A general shape of neuron spike (action potential) modeled from [107]. (c) Experimental data (points) for biological STDP learning function [107] and corresponding analytical approximation (solid line).

6.4 Resistive Memory as Electronic Synapses: Previous Work

The shortcomings of traditional silicon neuromorphic hardware implementations [108] such as large area and power consumption can be potentially overcome by using emerging resistive memory devices as electronic synapses along with silicon spiking neuron circuits [109]. For such a hybrid architecture to work, each synaptic element, i.e., each ReRAM

device should possess analog resistance change capability instead of only binary resistance switching. Various types of emerging ReRAM devices have been investigated for such incremental resistance switching capability. Metal Oxide based RRAM, [109]-[111] Ag/ amorphous Si based [112], TiO2 based [113], and Amorphous InGaZnO based memristive nanodevices [114], as well as phase change memory (PCM) devices [115] have been investigated for gradual resistance change behavior. However almost all such results were obtained from standalone devices either without any interaction with CMOS circuits or used a simplified version of the STDP. Although previously shown in [45], the possibility of incremental resistance programming was demonstrated for Ag/chalcogenide based CBRAM devices, such devices have not been explored for analog STDP based learning capability. A recent work proposed the possibility of implementing a stochastic version of the STDP rule by exploiting the binary resistance switching probability of CBRAM devices under weak programming conditions [116]. However this work did not explore the analog STDP characteristics of CBRAM devices and no post-silicon results were presented. In this chapter, compact model simulation results as well as experimental post-silicon results from CBRAM device integrated with CMOS spiking neuron circuits are presented that demonstrate potential for analog STDP characteristics in CBRAM memory for the first time.

6.5 STDP & Associative Memory Simulation

As discussed previously, CBRAM devices undergo an abrupt resistance transition from a HRS typically in the MΩ range to a LRS in the tens of KΩ range, due to the formation of a filament bridging the anode and cathode. However once in the LRS, the filament size can be controlled with depending on the bias conditions such as magnitude and duration of

voltage across the device or the current allowed through the device. This leads to presence of multiple LRS levels and possibility of incremental change in resistance once the device is in LRS as shown experimentally in previous work [45] and presented in Chapter 2. Thus in order to demonstrate STDP based learning, a CBRAM device in LRS can be well-suited. Based on the simulation results for compact model presented earlier in Chapter 3, the timing difference in arrival of pre and post neuron spikes can generate voltage differences of varying magnitudes and durations across the CBRAM device in LRS, which determines the change in filament dimension and hence the amount of deviation from initial LRS level. If the two spikes are generated too far apart in time, then the voltage generated across the device will be too small to cause any resistance change. This behavior is ideal for implementing the STDP rule which also requires spikes to be generated within a 'timing window' outside of which little or no change in synaptic plasticity occurs. Thus the CBRAM compact model used earlier can be used to simulate the expected STDP behavior.

The compact model is first programmed to a LRS value and then subjected to voltage spikes at anode and cathode to mimic spikes generated by post-synaptic neuron and pre-synaptic neuron respectively, as shown in the inset of Fig. 6.3. We can define the relative timing difference in arrival of the two spikes as the parameter ΔT used in equation 6.1 as ΔT = post neuron spike arrival time – pre neuron spike arrival time. If the post-synaptic neuron spikes before the pre-synaptic neuron ($\Delta T < 0$), then a net negative voltage will be applied across the CBRAM which can cause its initial LRS to increase (i.e. its conductance decreases) and if $\Delta T > 0$, then the CBRAM will see a positive voltage across it and its initial LRS can further decrease (i.e., its conductance increases). It should be noted that neuron spike magnitudes should be chosen such that spiking by only one of the neurons is not

sufficient to exceed the switching threshold and alter the CBRAM conductance significantly. To express the experimentally observed CBRAM STDP behavior, a the relative conductance change ΔG can be defined in terms of CBRAM conductance after application of pre and post neuron spikes (G_{post}) and the initial programmed conductance ($G_{initial}$) as,

$$\Delta G = \frac{G_{post} - G_{initial}}{G_{initial}} \quad . \tag{6.2}$$

Fig. 6.3 plots the simulated percentage values of ΔG vs. after application of each set of pre and post spikes for a set of different LRS values (50 KΩ, 25 KΩ and 10 KΩ). It can be seen that the conductance change follows the biological STDP rule. For positive and negative ΔT values within a certain range, the two neuron spikes overlap sufficiently to increase or decrease the CBRAM conductance (known as long term potentiation (LTP) or long term depression (LTD) respectively). For high values of ΔT in both positive and negative direction, the two spikes do not overlap sufficiently to create enough voltage drop to affect the conductance, which is desirable for mimicking the biological STDP rule.

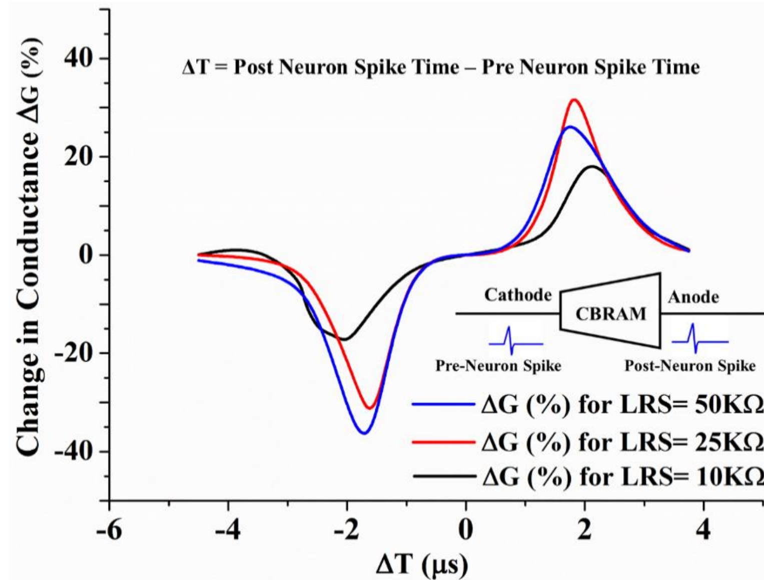


Fig. 6.3. Simulated STDP behavior with CBRAM compact model plotting percentage change in conductance vs. relative spike timing difference for three different initial LRS values.

A three neuron network was also simulated in Spectre using the CBRAM compact model to demonstrate an example of the associative learning behavior observed in biological neural networks. As shown in Fig. 6.4, this network consists of two input neurons (N_1 and N_2) connected to an output neuron (N_3) through their respective CBRAM synapses (S_1 and S_2). The neuron blocks are based on the leaky integrate-and-fire neuron model. For simulation, CBRAM synapses S_1 and S_2 are first set to LRS levels of 18 KΩ and 31 KΩ respectively. This is followed by application of spikes from the neurons to S_1 and S_2. From the simulation results shown in Fig. 6.5, it can be seen that spiking of only N_1 neuron causes the output neuron N_{out} to also fire. This causes a sufficient positive bias across S_1 so as to further decrease its resistance. However when only N_2 neuron fires later (at 75 μs), it is unable to cause output neuron spike due to much higher resistance of its synapse S_2. But if both N_1 and N_2 fire simultaneously (as at 125 μs), then N_{out} spikes again (due to N_1) and this generates sufficient voltage across S_2 so as to decrease its resistance

significantly through STDP rule. Subsequently, when only N_2 neuron fires (at 155 μs), it is now able to cause N_{out} to fire as well. This example demonstrates an example of associative learning, i.e., how an unconditioned stimulus (Neuron N_2 spike) can produce a response (output neuron spike) by being associated with a conditioned stimulus (Neuron N_1 spike). These simulation results indicate CBRAM devices may be suitable for use as electronic synapses.

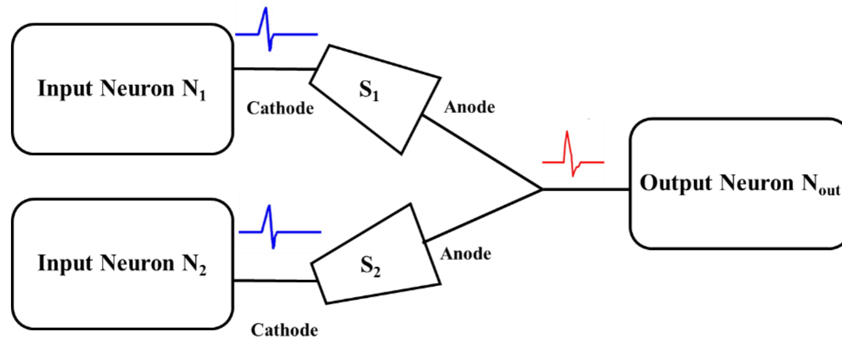


Fig. 6.4 Block diagram of the simulated three neuron network block diagram using CBRAM synapses.

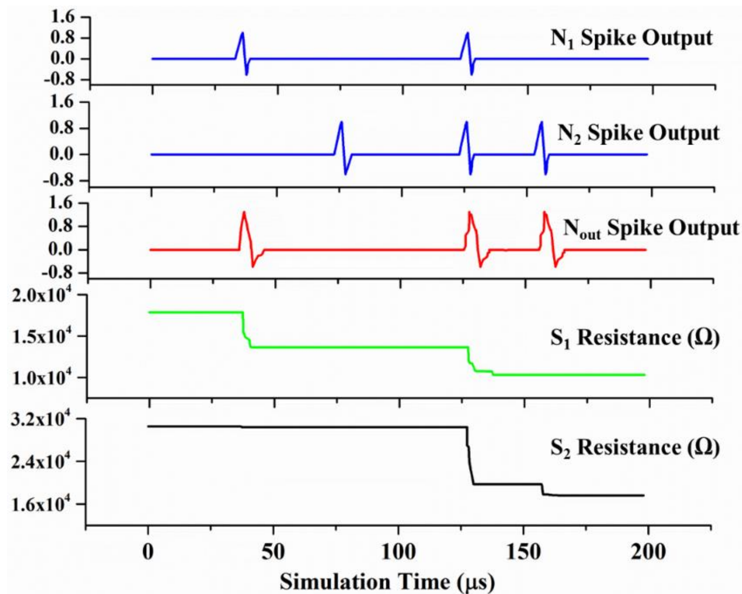


Fig. 6.5 Transient simulation results of the network of Fig. 6.4 showing how modification of CBRAM resistances based on input and output spike timing can be used for associative learning.

6.6 Neuron Hardware Implementation

Since there are many different types of biological neurons, there are many different ways to implement neuron circuits in silicon [117], [118]. In general digital neuron implementations tend to consume more power due to need for adders and multipliers compared to analog neuron implementations. The leaky integrate and fire (I & F) model based on the original Axon-Hillock neuron circuit proposed by Mead [119] is an example of a time-dependent neuron where the membrane potential increases due to integration of the input current onto a capacitor (referred to as the membrane capacitor). Once the membrane potential reaches a threshold of a comparator (in its simplest form an inverter), it causes comparator output to go high which in turn initiates the discharge of the membrane potential and this process continues in time. Such a model can be implemented using analog circuit design techniques and is well-suited to mimicking biological spiking neurons for asynchronous event based computation. The specific leaky I & F circuit used in this work for CBRAM STDP demonstration is based on the neuron design proposed in [120] was designed in a 0.6 µm process. Fig. 6.6(a) shows the transistor level schematic of the neuron circuit. When the Enable (En) input to PMOS P_2 is high, the neuron receives no input and it's membrane voltage (V_{mem} in Fig. 6.6(a)) remains at resting potential. When En is turned low, charge is injected into membrane capacitor C_{mem}. This causes V_{mem} to increase over time, until it exceeds the comparator threshold V_{th}, which causes the comparator output to go high. Output voltage of the second inverter (V_{out}) also rises and eventually turns on N_2, thereby initiating discharge of C_{mem}. Eventually V_{mem} goes below V_{th} and returns to its initial resting potential. The advantage of this circuit is the ability to control the threshold voltage (V_{th}) and also the membrane leakage current (through the V_{leak} bias input). The

comparator is chosen to be a compact five-transistor transconductance amplifier as shown in Fig. 6.6(b). Bias voltages for the neuron circuit (V_{bp} and V_{bn}) are generated using a constant g_m–bias circuit, also shown in Fig. 6.6(b). For the purpose of generating a voltage spike of sufficient magnitude to change the CBRAM state, a power supply voltage of 3.6 V was used. Since CBRAM write speeds can be less than 1 MHz, the membrane capacitance C_{mem} was chosen to be 9 pF so as to generate voltage spike widths in the µs range.

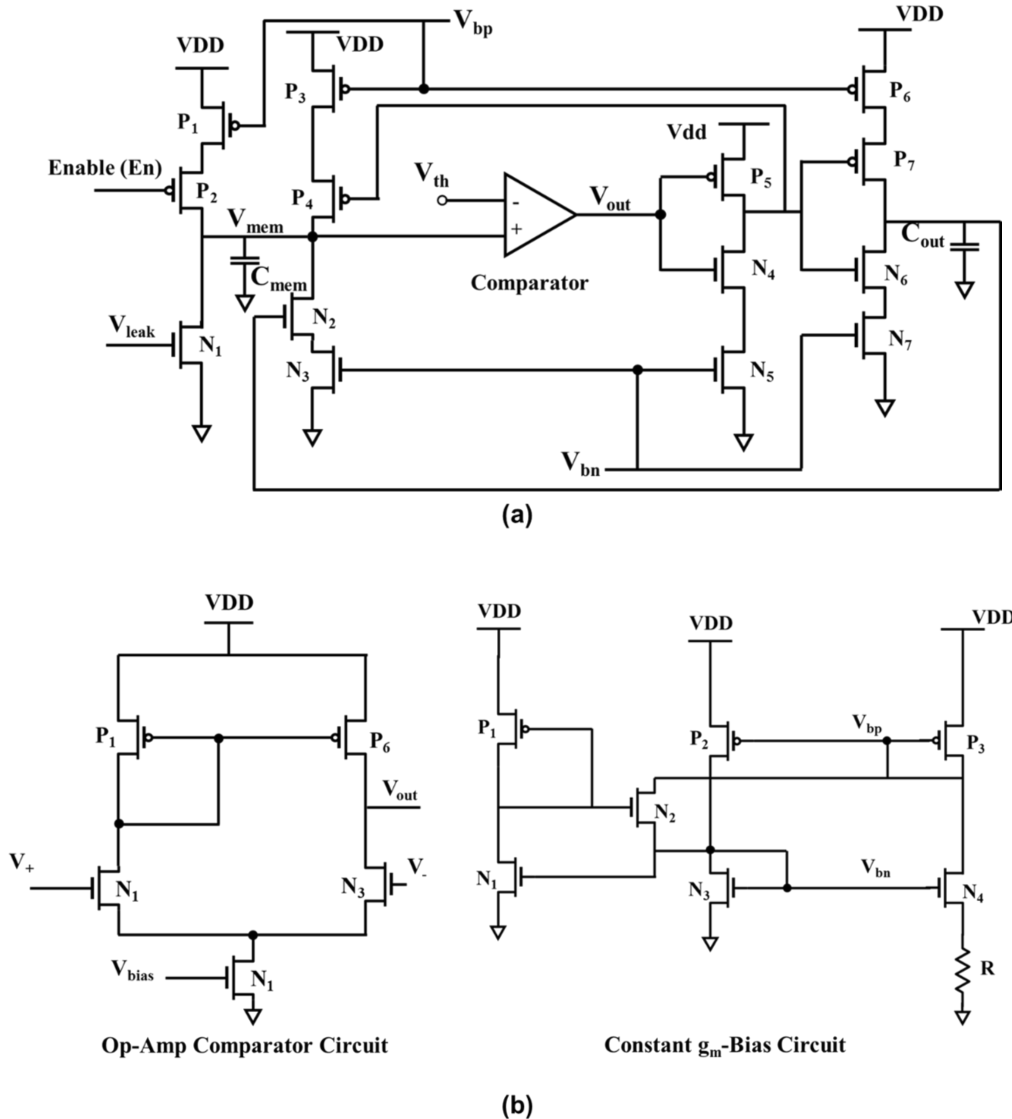


Fig. 6.6 (a) Leaky integrate and fire CMOS neuron circuit based on [120] used in this work. (b) Sub-blocks used for the neuron circuit - a transconductance amplifier for the comparator (left) and a constant-g_m bias circuit to generate the necessary voltage bias levels (right).

6.7 Silicon Neuron Post-Processing for CBRAM Synapse Integration

The CBRAM devices were fabricated as a vertical stack (Ag/ $Ge_{0.3}Se_{0.7}$) on top of metal pads on the die containing the CMOS neuron circuits at Arizona State University. As such the area of each device was same as the area of the pads, 30 µm x 30 µm. For each CBRAM device, a 100nm layer of chalcogenide glass ($Ge_{0.3}Se_{0.7}$) was first deposited by thermal

evaporation. A 400 nm Ag layer was then deposited also by thermal evaporation over the $Ge_{0.3}Se_{0.7}$ layer. Chalcogenide layer is then photo-doped with Ag by exposure to ultraviolet light for 20 minutes, in a similar way to other fabricated test devices used in this work. Another 100 nm layer of Ag is then deposited by thermal evaporation. Finally, a 1 micron layer of aluminum and 1 Additional 1 μm thick layers of Aluminum and Copper were deposited on top of Ag anode as contact for probing. Fig. 6.7(a) shows a micro-photograph of the top view of the die after post-processing and wire-bonding and Fig. 6.7(b) shows a schematic cross-section of the BEOL on-chip CBRAM device fabricated during post—processing.

Fig. 6.7. (a) Micro-Photograph of wire-bonded die after post-processing containing CMOS neurons and BEOL fabricated CBRAM devices. (b) Schematic of BEOL on-chip CBRAM cross-section.

6.8 Post-Silicon Validation & On-Chip STDP Demonstration

In order to demonstrate on-chip STDP learning behavior in CBRAM devices, two of the fabricated CMOS neurons (from Fig. 6.6(a)) are connected via a BEOL CBRAM device. The detailed test setup is shown in Fig. 6.8 below. The neuron circuit connected to the CBRAM cathode is referred to as the pre-synaptic neuron while the one connected to the CBRAM anode is referred to as the post-synaptic neuron, similar to Fig. 6.3 earlier. Hence, any spikes generated at the membrane potential nodes ($V_{mem,pre}$ and $V_{mem,post}$) of the

two neurons will be applied across the CBRAM device. The CBRAM cathode contact pad is internally connected in series to an on-chip access NMOS, which is used to program the CBRAM to a known LRS before spike generation and also to measure the change in CBRAM resistance after spiking event. As shown in the figure, an external switch is used to disconnect the CBRAM anode from the post-synaptic neuron membrane node during programming or reading so as to prevent any unwanted loading effect. Both the pre and post synaptic neurons have the same voltage bias and threshold voltage levels and hence produce identical spiking outputs. In order to demonstrate STDP learning, a relative timing difference (ΔT, as defined in section II) must be created between the pre and post neuron spikes arriving at the CBRAM terminals. For this purpose, enable inputs of the two neurons (Enable_pre and Enable_post in Fig. 6.8) are synchronized with a known phase difference by using a dual channel signal generator. Thus if one enable signal is delayed by a certain ΔT value with respect to the other, then the neuron receiving the enable signal later will fire after a delay ΔT with respect to the other neuron.

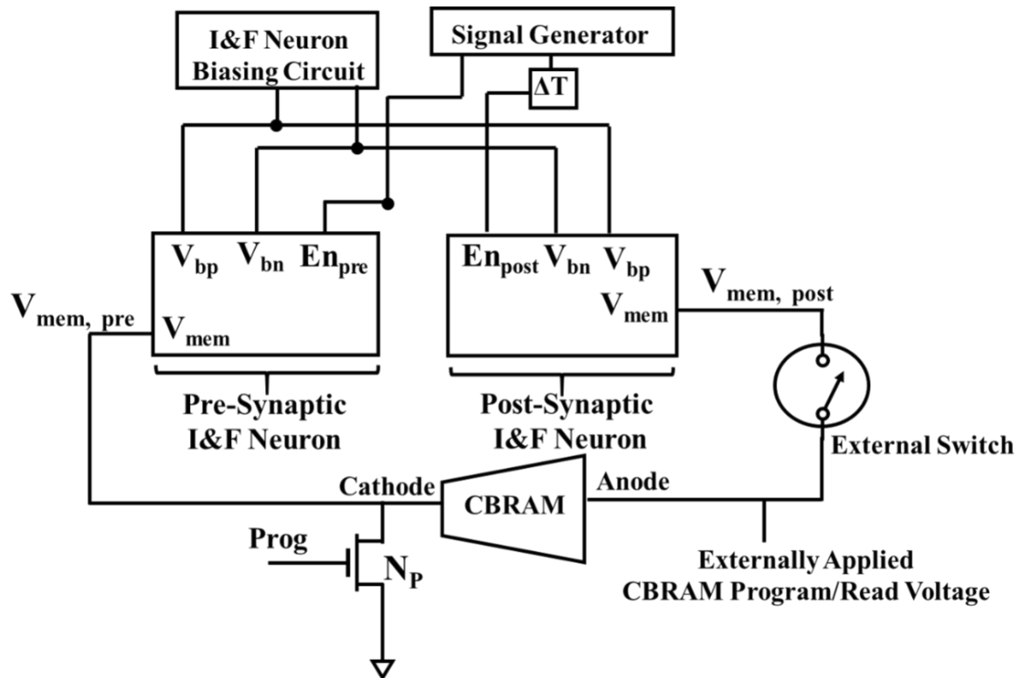


Fig. 6.8. Test setup for STDP demonstration consisting of two on-chip CMOS neurons connected via a single CBRAM synapse (also fabricated on chip in BEOL processing).

The STDP testing procedure used was as follows. First, the CBRAM is programmed to LRS using the access NMOS by applying a positive voltage (1.4 V) at the anode. Then the programmed LRS is sensed by applying a small read voltage at the anode. The Prog input is then made low to turn off the access NMOS and the external switch is closed to connect the CBRAM anode to the post-synaptic membrane output. Now the two enable signals are asserted with a certain ΔT phase difference, which causes two neurons to spike at time interval ΔT. The external switch is opened and the CBRAM state is sensed to detect the resistance change. This procedure is repeated by changing the timing difference ΔT in steps of 0.5 μs starting from zero.

Fig. 6.9 shows sample oscilloscope outputs of pre-synaptic and post-synaptic neuron spikes generated with a time delay (ΔT) of 1 μs using the test setup of Fig. 6.8. The experimentally observed percentage change in CBRAM conductance (ΔG) vs. ΔT for three

different LRS levels (10 KΩ, 5 KΩ and 1 KΩ) is shown for a spike width of 7 μs in Fig. 6.10(a), and for a spike width of 15 μs in Fig. 6.10 (b). It can be observed that a peak conductance change of about 20% occurs for 7 μs spike width, whereas a larger peak conductance change of about 40 % for LTP and -30% for LTD occurs for 15 μs spike width. This is expected as a longer spike width will enable larger change in filament dimension. Finally, it can be noted here that in chapter 2, device level transient characterization results demonstrated that the CBRAM LRS changes gradually with the magnitude and duration of the voltage across it. The results in Fig. 6.10 demonstrates that such bias dependent incremental LRS programming of CBRM LRS can be extended to BEOL devices at a circuit level as well and shows that CBRAM devices can be viable candidates for electronic synapse implementations in neuromorphic hardware.

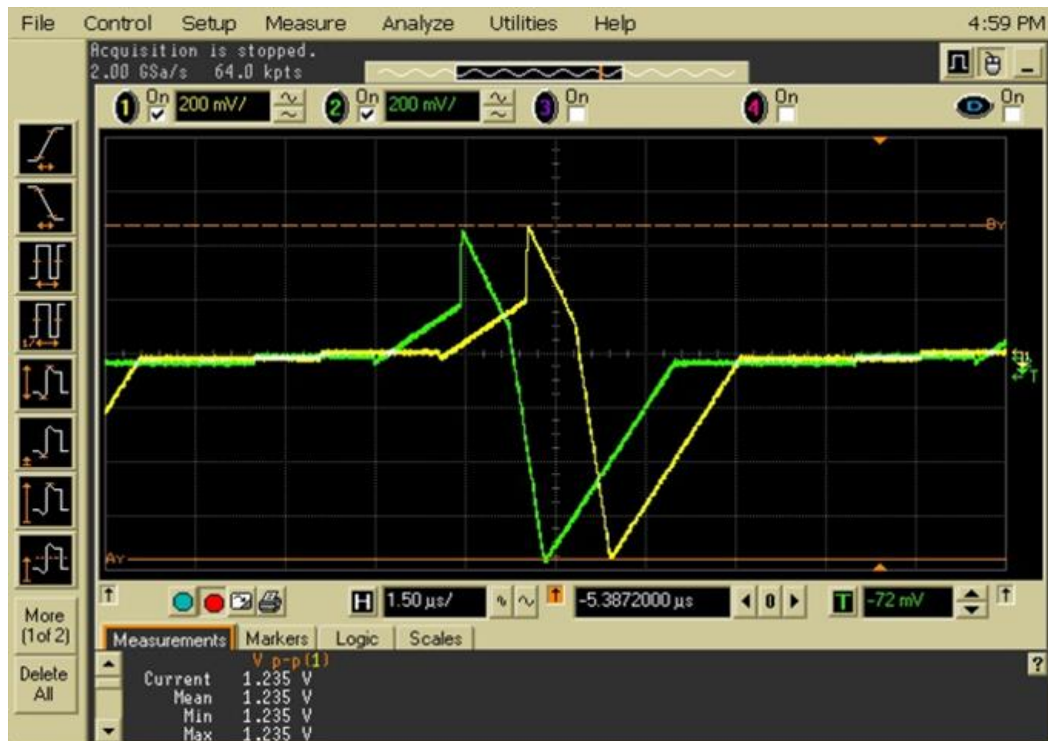


Fig. 6.9. Sample spike output waveforms from on-chip pre and post neuron using the test setup of Fig.6.8.

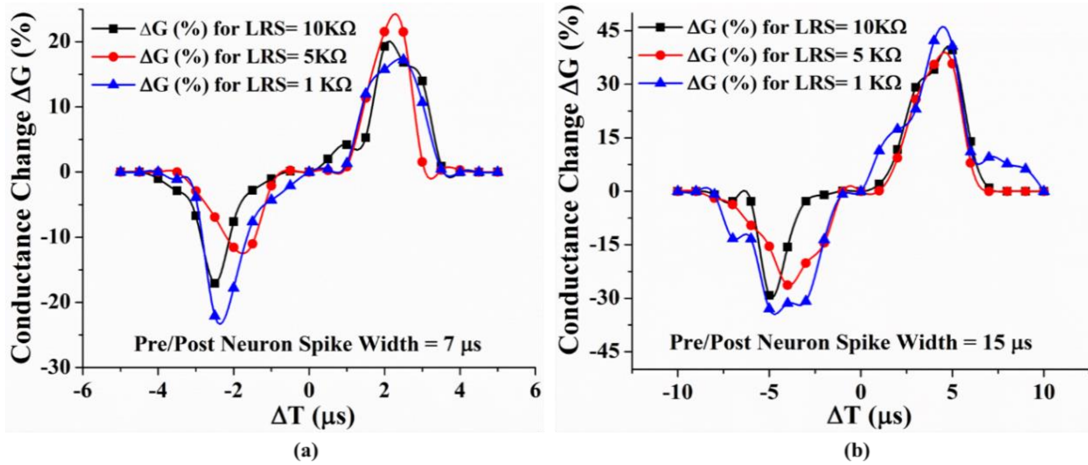


Fig. 6.10. Experimentally measured STDP behavior with on-chip CBRAM device, showing change in CBRAM conductance vs. spike timing difference, with pre and post neuron spike widths of (a) 7 μs and (b) 15 μs. Three different initial LRS values are used to exploit the multi-level programmability in CBRAM device.

6.9 Conclusions & Future Work

In this chapter we investigated analog spike timing dependent plasticity based learning in CBRAM resistive memory devices for neuromorphic applications. CMOS circuits for a leaky integrate and fire neuron with adjustable output spike characteristics were designed and fabricated. CBRAM devices were then fabricated on top of the CMOS layer of the die in back end of line post-processing. Experimental results obtained from the post-processed die demonstrated for the first time STDP based analog resistance change behavior in CBRAM devices that follows the biological STDP learning function. Effect of different factors such as initial programmed resistance level and neuron spike width are also presented through the experimental results. This property makes the CBRAM devices well-suited for use as low power nanoscale electronic synapses for neuromorphic applications. Since the feasibility of on-chip STDP learning is shown with these results, in the future

larger arrays of silicon neurons connected via BEOL CBRAM synapses can be fabricated to demonstrate complex pattern recognition or other learning behaviors.

CHAPTER 7

SUMMARY & CONCLUSIONS

Conductive bridge resistive memory technology is still in its developmental phase and therefore understanding of its physical mechanisms and determination of its application areas remains a subject of research. In this work, both these aspects were explored in detail. Additionally, multi-level resistance switching, being one of the more promising and unique characteristics of CBRAM technology, formed the primary focus of the results regarding experimental characterization, compact modeling and circuit applications of CBRAM devices presented in this work.

The key contributions of this work are summarized below:

- Chapter 1 outlines the various types of emerging non-volatile memory technologies and presents a well-defined physical model of resistance switching behavior observed in CBRAM devices.
- Chapter 2 presents experimental data obtained from electrical and material characterization of Silver-Chalcogenide glass based CBRAM devices fabricated at ASU. The small signal impedance measurement results are used to develop a RC equivalent circuit of CBRAM devices specifically with regard to the photo-doping step performed on such devices which is critical for ensuring uniform resistance switching without the need for the electro-forming step. Data from quasi-static and transient characterization is also presented to demonstrate the ability to program CBRAM devices into multiple different on states. Although compliance current based multi-level programmability has been demonstrated previously, the results in this work demonstrate the possibility of the LRS

undergoing gradual increase or decrease based on the magnitude, duration and polarity of voltage bias applied across the device. This has promising implications for use of CBRAM technology in neuromorphic circuit applications.

- Chapter 3 presents a CBRAM compact model developed in Verilog-a language based on the resistance switching model presented in chapter 1, that can be used to simulate the dynamic bias-dependent multi-level resistance programmable characteristics observed experimentally and can be used for circuit simulations. The model parameters can be adjusted to fit CBRAM devices composed of different materials.
- Chapter 4 analyzes the susceptibility of CBRAM devices under bias within arrays to transient radiation in the form of single event effects through experimental heavy ion testing of 1T-1R CBRAM devices and compact model simulations. Although effect of total ionizing radiation has been explored previously for such devices, this work explores their susceptibility to transient single event effects and results obtained demonstrate the possibility of both HRS to LRS and LRS to HRS upsets occurring due to single event effects in access transistor of 1T-1R cells or in peripheral CMOS circuit in crossbar CBRAM arrays.
- Chapter 5 presents a novel sense-amplifier based non-volatile flip-flop architecture using CBRAM devices to store flip-flop data during idle period to potentially completely eliminate leakage power consumption unlike traditional CMOS circuits. The tradeoffs involved between flip-flop energy efficiency and

reliability and the impact of the choice of CBRAM LRS levels on flip-flop performance are analyzed through compact model based circuit simulations. The extension of such an architecture to perform reconfigurable logic computation for a special class of Boolean logic functions known as threshold logic functions by exploiting the multi-level programmability of the CBRAM devices is also discussed.

- Chapter 6 explores the potential use of the analog (gradual) resistance change capability shown by CBRAM devices in experimental data from chapter 2 for neuromorphic applications. BEOL fabricated CBRAM devices are shown to demonstrate resistance change behavior in accordance with the important neuromorphic learning rule known as spike timing dependent plasticity (STDP). Details of a test chip fabricated with CMOS spiking neuron circuits and BEOL CBRAM devices are discussed. Such experimental demonstration of biological STDP behavior in CBRAM technology for the first time shows the suitability of CBRAM devices for potential use as electronic synapses in complex hybrid neuromorphic computing architectures.

In conclusion, it can be pointed out that although the physical mechanisms outlined and modeled here as well as the key circuit applications explored in this work focus on Silver-Chalcogenide based CBRAM devices, the same concepts should also be broadly applicable to the general class of electrochemical resistance switching memory devices for future research.

REFERENCES

[1] G. R. Fox, F. Chu, and T. Davenport,"Current and future ferroelectric nonvolatile memory technology," *J. Vacuum Sci. & Technol. B*19.5 (2001): 1967-1971.

[2] R. Bailey, G. Fox, J. Eliason, M. Depner, D. Kim, E. Jabillo, J. Groat, J. Walbert, T. Moise, S. Summerfelt, K. R. Udayakumar, J. Rodriquez, K. Remack, K. Boku and J. Gertas, "FRAM memory technology - advantages for low power, fast write, high endurance applications," *IEEE ICCD: VLSI in Computers and Processors,* pp. 485, 2-5 Oct. 2005.

[3] S. Tehrani, J. M. Slaughter, E. Chen, M. Durlam, J. Shi, M. DeHerren, "Progress and outlook for MRAM technology," *IEEE Trans. Magnetics*, vol.35, no.5, pp.2814-2819, Sep. 1999.

[4] M. Hosomi, H. Yamagishi, T. Yamamoto, K. Bessho, Y. Higo, K. Yamane, H. Yamada, M. Shoji, H. Hachino, H.; Fukumoto, C.; Nagao, H.; Kano, H., "A novel nonvolatile memory with spin torque transfer magnetization switching: spin-ram," *IEDM Tech. Digest,* pp. 459-462, 5-5 Dec. 2005.

[5] A. L. Lacaita, A. Redaelli, D. Ielmini, F. Pellizzer, A. Pirovano, A. Benvenuti, R. Bez, "Electrothermal and phase-change dynamics in chalcogenide-based memories," *IEDM Tech. Digest*, pp.911-914, 13-15 Dec. 2004.

[6] H.-S.P. Wong, S. Raoux, K. SangBum, Jiale Liang, J. P. Reifenberg, B. Rajendran, M. Asheghi, K. E. Goodson, "Phase Change Memory," *IEEE Proc.*, vol.98, no.12, pp.2201-2227, Dec. 2010.

[7] S. Lai, "Current status of the phase change memory and its future," *IEDM Tech. Digest*, pp.10.1.1-10.1.4, 8-10 Dec. 2003.

[8] S. Raoux, G. W. Burr, M. J. Breitwisch, C. T. Rettner, Y. C. Chen, R. M. Shelby, M. Salinga, D. Krebs, S.-H. Chen, H. L. Lung and C. H. Lam, "Phase-change random access memory: A scalable technology," *IBM J. R. & D.*, vol.52, no.4.5, pp.465-479, July 2008.

[9] H.-S.P. Wong, H. Lee, S. Yu, Y. Chen; Y. Wu, P. Chen; B. Lee, F. T. Chen and M. Tsai, "Metal–Oxide RRAM," *IEEE Proc.*, vol.100, no.6, pp.1951-1970, June 2012.

[10] H. Y. Lee, P. S. Chen, T. Y. Wu, Y. S. Chen, C. C. Wang, P. J. Tzeng, C. H. Lin, F. Chen, C. H. Lien and M. J. Tsai, " Low power and high speed bipolar switching with a thin reactive Ti buffer layer in robust HfO2 based RRAM," *IEDM Tech. Digest*, pp.1-4, 15-17 Dec. 2008.

[11] C. Chen, C. Song, J. Yang, F. Zeng, F. and F. Pan, “Oxygen migration induced resistive switching effect and its thermal stability in W/TaOx/Pt structure," *Appl. Phys. Lett.,* vol. 100, no. 25, pp. 2012.

[12] M.N. Kozicki, M. Park, M. Mitkova, “Nanoscale memory elements based on solid state electrolytes,” *IEEE Trans. Nanotech.*, vol. 4, no. 3, pp. 331–338, May 2005.

[13] I. Valov, M.N. Kozicki, “Cation-based resistance change memory,” *J. Phys. D: Appl. Phys*., volume 46 074005, 2013.

[14] R. Waser, M. Aono, “Nanoionics-based resistive switching memories,” *Nat. Mater.,* vol. 6, pp. 833–840, 2007.

[15] P. Blanchard et.al, Adesto Technologies, “First Commercial Demonstration of an Emerging Memory Technology for Embedded flash using CBRAM,” *Innovative Memory Technologies Workshop* MINATEC, Grenoble France, 2011.

[16] I. Valov, R. Waser , J.R .Jameson, M.N. Kozicki ,“Electrochemical metallization memories-fundamentals, applications, prospects*,” Nanotechnology*, vol. 22 254003, 2011.

[17] N.E. Gilbert, M.N. Kozicki-*“*An Embeddable Multilevel-Cell Solid Electrolyte Memory Array*”, IEEE Journal of Solid-State Circuits*, vol. 42, no. 6, pp. 1383-1391, June 2007.

[18] S. Wald, J. Baker, M. Mitkova, N. Rafla, "A non-volatile memory array based on nano-ionic Conductive Bridge Memristors," *IEEE Workshop on Microelectronics and Electron Devices (WMED)*, vol. 1, no. 4, pp.22 , April 2011.

[19] R. Bruchhaus, M. Honal, R. Symanczyk, and M. Kund, “Selection of optimized materials for CBRAM based on HT-XRD and electrical test results,” *J. Electrochem. Soc.*, vol. 156, issue 9, pp. 729-733, 2009.

[20] S. Manhart, "Memory switching in SiO films with Ag and Co electrodes," *J. Physics D: Appl. Phys.*, Vol. 6, no.1, (1973).

[21] M.N. Kozicki, "A low-power nonvolatile switching element based on copper-tungsten oxide solid electrolyte." *IEEE Trans. Nanotech.,* vol. 5, no. 5, pp. 535-544, 2006.

[22] M. N. Kozicki, M. Mitkova, M. Park, M. Balakrishnan, C. Gopalan, "Information storage using nanoscale electrodeposition of metal in solid electrolytes," *Superlattices and Microstructures*, vol. 34, pp. 459-465, December 2003.

[23] Pradeep Dandamudi, "Resistance Switching in Chalcogenide based Programmable Metallization Cells (CBRAM) and Sensors under Gamma-Rays," PhD. Dissertation, Arizona State University, 2013.

[24] A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications.* New York: Wiley.

[25] N. F. Mott and R.W. Gurney, "Electronic Processes in Ionic Crystals," *J. Phys. Chem.*, vol. 45, issue 7, pp 1142–1142, 1941.

[26] M.N. Kozicki, C. Ratnakumar, M. Mitkova,"Electrodeposit Formation in Solid Electrolytes," *NVMTS*, vol. 111, no. 115, pp. 5-8, Nov. 2006.

[27] J. R. McDonald and W.B. Johnson, "Fundamentals of Impedance Spectroscopy," in *Impedance Spectroscopy: Theory, Experiment, and Applications***,** *2nd* ed. New Jersey: John Wiley & Sons, Inc., 2005, ch. *1*.

[28] A. J. Bard and L. R. Faulkner, "Techniques Based on Concepts of Impedance," in *Electrochemical Methods: Fundamentals and Applications***,** *2nd* ed. New York: John Wiley & Sons, Inc., 2005, ch. *10*.

[29] D.C**.** Sinclair, "Characterization of Electro-materials using ac Impedance Spectroscopy." *Boletín de la Sociedad Española de Cerámica y Vidrio* 34.2, pp. 55-65, 1995.

[30] M.E. Orazem and B. Tribollet, “Preliminary Graphical Methods,” in *Electrochemical Impedance Spectroscopy*, New Jersey: John Wiley & Sons, Inc., ch.17, 2008.

[31] J. D. Greenlee, W. L. Calley, M.W. Moseley and W. A. Doolittle *“*Comparison of Interfacial and Bulk Ionic Motion in Analog Memristors,*” IEEE Trans. on Electron Devices*, vol. 60, no. 1, pp. 427-432, Jan. 2013.

[32] Y. Huang,, Y. Huang and T. Hsieh, "A study of phase transition behaviors of chalcogenide layers using in situ alternative-current impedance spectroscopy," *Journal of Applied Physics* , vol.111, no.12, pp.123706,123706-7, Jun 2012.

[33] I. Snorri, M. Arikan, M. Carter, S. Weifeng and X. Gang, "Impedance spectroscopy of micron sized magnetic tunnel junctions with MgO tunnel barrier," *Appl. Phys. Lett.,* vol. 96, issue 23, 2010.

[34] D. S. Jeong,_ H. Schroeder, R. Waser, “Impedance spectroscopy of TiO2 thin films showing resistive switching,,” *Appl. Phys. Lett.,* vol. 89, issue 8, 2006.

[35] D. Mahalanabis, Y. Gonzalez-Velo, H.J. Barnaby, M.N. Kozicki, P. Dandamudi and S. Vrudhula, "Impedance Measurement and Characterization of Ag-$Ge_{30}Se_{70}$-Based Programmable Metallization Cells," *IEEE Trans. Electron Devices*, vol.61, no.11, pp.3723-3730, Nov. 2014.

[36] M. T. Kostyshin, E. V. Mikhailovskaya, and P. F. Romanenko, *Sov. Phys. Solid St.,* vol. 8, pp. 451-453, 1966.

[37] A.V. Kolobov, S.R. Elliott, “Photodoping of amorphous chalcogenides by metals,” *Advances in Physics*, vol. 40, Issue 5, p.625-684, 1991.

[38] M. Mitkova, and M.N. Kozicki, “Silver incorporation in Ge–Se glasses used in programmable metallization cell devices,” *J. Non-Crystalline Solids*, vol. 299-302, part 2, pp. 1023-1027, April 2002.

[39] M. Mitkova, M.N. Kozicki, H.C. Kim, T.L. Alford, “Crystallization Effects in Annealed thin Ge-Se films Photo diffused with Ag,” J. *Non-Crystalline Solids*, 352 (2006) 1986-1990.

[40] M.N. Kozicki, M. Mitkova, J. Zhu, M. Park, “Nanoscale phase separation in Ag-Ge-Se glasses”, *Microelectronic Engineering* vol. 63, issues 1-3, pp. 155-159, Aug 2002.

[41] M. Mitkova, M.N. Kozicki, H.C. Kim, and T. Alford, “Local structure resulting from photo and thermal diffusion of Ag in Ge-Se films,” J. *Non-Crystalline Solids*, vol. 338-340, pp. 552-556, Jun 2004.

[42] C. Schindler, “Resistive switching in electrochemical metallization memory cells,” Ph.D. dissertation, RWTH Aachen Univ., Aachen, N. Rhine-Westphalia, 2009.

[43] U. Russo, D. Kamalanathan, D. Ielmini, A. L. Lacaita and M. N. Kozicki, “Study of Multilevel Programming in Programmable Metallization Cell (CBRAM) Memory,” *IEEE Trans. Electron Devices*, vol. 56, no. 5, pp. 1040-1047, May 2009.

[44] C. Schindler, G. Staikov, R. Waser, “Electrode kinetics of Cu–SiO2-based resistive switching cells: Overcoming the voltage–time dilemma of electrochemical metallization memories,” *Applied Physics Letters*, vol. 94, 2009.

[45] D. Mahalanabis, H.J. Barnaby, Y. Gonzalez-Velo, M.N. Kozicki, S. Vrudhula and P. Dandamudi, “Incremental resistance programming of programmable metallization cells for use as electronic synapses,” *Solid-State Electron.*, vol. 100, pp. 39-44, Oct. 2014.

[46] S. Yu, Yi Wu, R. Jeyasingh , D. Kuzum and H.-S.P. Wong, "An Electronic Synapse Device Based on Metal Oxide Resistive Switching Memory for Neuromorphic Computation," *IEEE Trans. on Electron Devices*, vol.58, no.8, pp.2729-2737, Aug. 2011.

[47] B. Linares-Barranco, T. Serrano-Gotarredona, “Memristance can explain spike time-dependent-plasticity in neural synapses,” Nature Precedings, 2009.

[48] S. Yu, H.-S.P. Wong, “Modeling the Switching Dynamics of Programmable-Metallization-Cell (CBRAM) Memory and Its Application as Synapse Device for a Neuromorphic Computation System,” *IEEE International Electron Devices Meeting (IEDM),* vol. 22.1.1, no. 22.1.4, pp. 6-8, Dec. 2010.

[49] C. Gopalan, Y. Ma, T. Gallo, J. Wang, E. Runnion, J. Saenz, F. Koushan, P. Blanchard, and S. Hollmer, “Demonstration of Conductive Bridging Random Access Memory (CBRAM) in logic CMOS process,” Solid-State Electron., vol. 58, pp. 54-61, 2011.

[50] F.W. Sexton, "Destructive single-event effects in semiconductor devices and ICs," *IEEE Transactions on Nuclear Science,* vol.50, no.3, pp.603-621, June 2003.

[51] P. E. Dodd, M. R. Shaneyfelt, D. S. Walsh, J. R. Schwank, G. L. Hash, R. A. Loemker, B. L. Draper, and P. S.Winokur, “Single-event upset and snapback in silicon-on-insulator devices and integrated circuits,” *IEEE Trans. Nucl. Sci.*, vol. 47, pp. 2165–2174, Dec. 2000.

[52] J. Maiz and R. Baumann, “Radiation induced soft errors in silicon components and computer systems,” in *Tutorial, Int. Reliability Physics Symp.*, Dallas, April 2002.

[53] F. Faccio, K. Kloukinas, A. Marchioro, T. Calin, J. Cosculluela, M. Nicolaidis, and R. Velazco, “Single event effects in static and dynamic registers in a 0.25 μm CMOS technology,” *IEEE Trans. Nucl. Sci.*, vol. 46, pp. 1434–1439, Dec. 1999.

[54] S. Duzellier, R. Ecoffet, D. Falguère, T. Nuns, L. Guibert, W. Hajdas, and M. C. Calvet, “Low energy proton induced SEE in memories,” *IEEE Trans. Nucl. Sci.*, vol. 44, pp. 2306–2310, Dec. 1997.

[55] R. Ronen, A. Mendelson, K. Lai, L. Shih-Lien, F. Pollack, and J. P. Shen, “Coming challenges in microarchitecture and architecture,” *Proc. IEEE*, vol. 89, p. 325, Mar. 2001.

[56] *Semiconductor Research Corp. Nat. Technology Roadmap*, 1999.

[57] P.E. Dodd and L.W. Massengill, "Basic mechanisms and modeling of single-event upset in digital microelectronics," in *IEEE Trans. Nucl. Sci.*, vol.50, no.3, pp.583-602, June 2003.

[58] P. E. Dodd, M. R. Shaneyfelt, J. A. Felix, and J. R. Schwank, “Production and Propagation of Single-Event Transients in High-Speed Digital Logic ICs,” *IEEE Trans. Nucl. Sci.*, vol. 51, no. 6, pp.3278-3284, Dec. 2004.

[59] Y. Gonzalez-Velo, H.J. Barnaby, M.N. Kozicki, P. Dandamudi, A. Chandran, K.E. Holbert, M. Mitkova and M. Ailavajhala, "Total-Ionizing-Dose Effects on the Resistance Switching Characteristics of Chalcogenide Programmable Metallization Cells," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 6, pp. 4563 – 4569, Dec. 2013.

[60] P. Dandamudi, H. J. Barnaby, M. N. Kozicki, Y. Gonzalez-Velo, and K. E. Holbert, "Total Ionizing Dose Tolerance of the Resistance Switching of Ag-$Ge_{40}S_{60}$ based Programmable Metallization Cells,", *IEEE Trans. Nucl. Sci.,* vol.61, no.4, pp.1726-1731, Aug. 2014.

[61] Y. Gonzalez-Velo, H. J. Barnaby, M. N. Kozicki, C. Gopalan, and K. Holbert, "Total Ionizing Dose Retention Capability of Conductive Bridging Random Access Memory," *IEEE Electron Device Lett.,* vol. 35, no. 2, pp. 205-207, Feb. 2014.

[62] J.L. Taggart, Y. Gonzalez-Velo, D. Mahalanabis, A.Mahmud, A. H.J. Barnaby, M.N. Kozicki, K.E. Holbert, et al., "Ionizing Radiation Effects on Nonvolatile Memory Properties of Programmable Metallization Cells," *IEEE Nucl. Sci.*, vol.61, no.6, pp.2985-2990, Dec. 2014.

[63] D. Chen, H. Kim, A. Phan, E. Wilcox, K. LaBel, S. Buchner, A. Khachatrian and N. Roche, "Single-Event Effect Performance of a Commercial Embedded ReRAM," *IEEE Nucl. Sci.,* vol.61, no.6, pp.3088-3094, Dec. 2014.

[64] D. Mahalanabis, H. J. Barnaby, M. N. Kozicki, V. Bharadwaj and S. Rajabi, "Investigation of Single Event Induced Soft Errors in Programmable Metallization Cell Memory," *IEEE Trans. Nucl. Sci.*, vol.61, no.6, pp.3557-3563, Dec. 2014.

[65] D.Mahalanabis, R. Liu, H. J. Barnaby, S. Yu, M. N. Kozicki, A. Mahmud and E. Deionno, "Single Event Susceptibility Analysis in CBRAM Resistive Memory Arrays," Accepted for publication, *IEEE Trans. Nucl. Sci.*, Sept. 2014.

[66] P. E. Dodd, M. R. Shaneyfelt, K. M. Horn, D. S. Walsh, G. L. Hash, T. A. Hill, B. L. Draper, J. R. Schwank, F. W. Sexton, and P. S. Winokur, "SEU-sensitive volumes in bulk and SOI SRAMs from first-principles calculations and experiments," *IEEE Trans. Nucl. Sci.*, vol. 48, pp.1893–1903, Dec. 2001.

[67] H.T.Weaver, "Soft error stability of p-well versus n-well CMOS latches derived from 2D, transient simulations," *IEDM Tech. Dig.*, Dec. 1988, pp. 512–515.

[68] D. Kobayashi, H. Saito, K. Hirose, "Estimation of Single Event Transient Voltage Pulses in VLSI Circuits From Heavy-Ion-Induced Transient Currents Measured in a Single MOSFET," *IEEE Transactions on Nuclear Science,* vol.54, no.4, pp.1037-1041, Aug. 2007.

[69] J.L. Wirth and S.C. Rogers, "The Transient Response of Transistors and Diode to Ionizing Radiation," *IEEE Trans. Nucl. Sci.*, vol. 11, no. 5, pp. 24-38, Nov. 1964.

[70] P. Oldiges, R. Dennard, D. Heidel, B. Klaasen, F. Assaderaghi, and M. Ieong, "Theoretical Determination of the Temporal and Spatial Structure of α -Particle Induced Electron–Hole Pair Generation in Silicon", *IEEE Trans. Nucl. Sci*., vol. 47, pp. 2575–2579, Dec. 2000.

[71] W. G. Bennett, N. C. Hooten, R. D. Schrimpf, J. Bi, E. Zhang, M. L. Alles, R.t A. Reed. D. Linten, M. Jurczak and A. Fantini. "Single Event Induced Upsets in HfO2/Hf 1T-1R RRAM," *IEEE Trans. Nucl. Sci.*, vol. 61, no. 4, pp 1717-1725, 2014.

[72] Y. Wang, Y. Hao Yu and W. Zhang, "Nonvolatile CBRAM-Crossbar-Based 3-D-Integrated Hybrid Memory for Data Retention," *IEEE Trans.VLSI Systems,* vol.22, no.5, pp.957-970, May 2014.

[73] K.H. Kim, S. Gaba, D. Wheeler, J.M. Cruz-Albrecht, T. Hussain, N. Srinivasa and W. Lu, "A functional hybrid memristor crossbar-array/CMOS system for data storage and neuromorphic applications," *Nano Lett.*, vol. 12, issue 1, pp. 389-395, Jan. 2012.

[74] R. Liu, D. Mahalanabis, H.J. Barnaby and S. Yu, "Investigation of Single-Bit and Multiple-Bit Upsets in Oxide RRAM-Based 1T1R and Crossbar Memory Arrays," in *IEEE Trans. Nucl. Sci.,* vol.62, no.5, pp.2294-2301, Oct. 2015.

[75] Y. Deng, P. Huang, B. Chen, X. Yang, B. Gao, J. Wang, L. Zeng, G. Du, J. Kang and X. Liu, "RRAM Crossbar Array With Cell Selection Device: A Device and Circuit Interaction Study," *IEEE Trans. Electron Devices*, vol.60, no.2, pp.719-726, Feb. 2013.

[76] R. Vullers, R. Schajk, H. Visser, J. Penders, and C. Hoof, "Energy harvesting for autonomous wireless sensor networks," *IEEE Solid-State Circuits Mag.*, vol. 2, no. 2, pp. 29-38, Spring 2010.

[77] J. Constantin, A. Dogan, O. Andersson, P. Meinerzhagen, J. N. Rodrigues, D. Atienza and A. Burg, "TamaRISC-CS: An ultra-low-power application-specific processor for compressed sensing," *IEEE/IFIP 20th International Conference on VLSI-SoC*, vol. 159, no. 164, pp.7-10, Oct 2012.

[78] Y. Jung, K. Jisu, R. Kyungho, J. Seong-Ook, J.P. Kim and S.H. Kang, "MTJ based non-volatile flip-flop in deep submicron technology," *International SoC Design Conference (ISOCC),* vol. 424, no. 427, pp.17-18, 2011.

[79] J. Chul-Moon, J. Kwan-Hee, Eun-Sub Lee, Huan Minh Vo and Kyeong-Sik Min, "Zero-Sleep-Leakage Flip-Flop Circuit With Conditional-Storing Memristor Retention Latch," *IEEE Transactions on Nanotechnology,* vol.11, no.2, pp.360-366, Mar. 2012.

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

[81] I. Kazi, P. Meinerzhagen, P.-E. Gaillardon, D. Sacchetto, Y. Leblebici, A. Burg, G. De Micheli,"Energy/Reliability Trade-Offs in Low-Voltage ReRAM-Based Non-Volatile Flip-Flop Design," *IEEE Trans. Circuits and Systems I: Regular Papers,* vol.61, no.11, pp.3155-3164, Nov. 2014.

[82] D. Mahalanabis, V. Bharadwaj, H. J. Barnaby, S. Vrudhula and M.N. Kozicki, "A Nonvolatile Sense Amplifier Flip-Flop Using Programmable Metallization Cells," *IEEE JETCAS*, vol.5, issue 2, pp 205-213, June 2015.

[83] B. Nikolic and V. G. Oklobdzija, "Design and Optimization of Sense-Amplifier-Based Flip-Flops," *ESSCIRC*, pp. 410-413, 1999.

[84] B. Nikolic, G. Vojin Oklobdzija, V. Stojanovic, J. Wenyan, James Kar-Shing Chiu and M. Ming-Tak Leung, "Improved sense-amplifier-based flip-flop: design and measurements," IEEE Journal of Solid-State Circuits, vol.35, no.6, pp.876-884, June 2000.

[85] A.G.M. Strollo, D. De Caro, E. Napoli, and N. Petra, "A novel high-speed sense-amplifier-based flip-flop," IEEE Trans. on Very Large Scale Integration (VLSI) Systems, , vol.13, no.11, pp.1266-1274, Nov. 2005.

[86] H.P. Alstad, S. Aunet, "Three Subthreshold Flip-Flop Cells Characterized in 90 nm and 65 nm CMOS Technology," IEEE Workshop on Design and Diagnostics of Electronic Circuits and Systems, 2008. DDECS 2008. 11th, vol. 1, no.4, pp.16-18, Apr. 2008.

[87] B. Wicht, T. Nirschl, D. Schmitt-Landsiedel,"Yield and speed optimization of a latch-type voltage sense amplifier," IEEE Journal of Solid-State Circuits, vol.39, no.7, pp.1148-1158, July 2004.

[88] P. Meinerzhagen, O. Andersson, B. Mohammadi, Y. Sherazi, A. Burg, and J. N. Rodrigues, "A 500 fW/bit 14 fJ/bit-access 4kb standard-cell based sub-VT memory in 65nm CMOS," *Proceedings of ESSCIRC,* vol. 321, no. 324, pp. 17-21, Sept. 2012.

[89] V. Beiu, J. M. Quintana, and M. J. Avedillo, "VLSI implementations of threshold logic - a comprehensive survey," *IEEE Trans. on Neural Networks*, vol. 14, no. 5, pp. 1217-1243, 2003.

[90] S. Leshner, N. Kulkarni, S. Vrudhula and K. Berezowski, "Design of a robust, high performance standard cell threshold logic family for DSM technology," *International Conference on Microelectronics (ICM),* vol. 52, no. 55, pp.19-22 Dec. 2010.

[91] "International Technology Roadmap for Semiconductors, Executive Summary, Emerging Research Devices Grand Challenges, p.30; Available: www.itrs.net/Links/2011ITRS/Home2011.htm.

[92] "Can Programming Be Liberated from the von Neumann Style? A Functional Style and Its Algebra of Programs J. Bachus", *Commun. ACM*, pp 613, 1978.

[93] P.Carrasco, Z.Ramos, Gotarredona, L.Barranco, "On Neuromorphic Spiking Architectures for asynchronous memristive STDP systems", Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS).

[94] J. Schemmel, D. Brüderle, A. Grübl, M. Hock, K. Meier and S. Millner, "A wafer-scale neuromorphic hardware system for large-scale neural modeling," *IEEE (ISCAS), Proc.*, pp.1947-1950, May 30-June 2, 2010.

[95] H. Markram, “The Blue Brain Project,” *Nat. Rev. Neurosci.,* vol. **7**, issue 2, pp 153-160, February 2006.

[96] J. Lin, P. Merolla, J. Arthur and K. Boahen, "Programmable Connections in Neuromorphic Grids," *IEEE MWSCAS*, vol.1, pp.80-84, Aug. 6-9, 2006.

[97] E. M. Izhikevich, "Simple model of spiking neurons." *IEEE Trans. Neural Networks,* vol. 14, no. 6, pp. 1569-1572, 2003.

[98] C. Mead, "Neuromorphic electronic systems," *IEEE Proc.*, vol.78, no.10, pp.1629-1636, Oct 1990.

[99] G. S. Snider, "Self-organized computation with unreliable, memristive nanodevices." *Nanotechnology*, vol. 18, no. 36, pp 365202, 2007.

[100] J. Seo, B. Brezzo, B. Yong Liu, B.D. Parker, S.K. Esser, R.K. Montoye, B. Rajendran, J.A. Tierno, L. Chang, D. S. Modha and D.J. Friedman, "A 45nm CMOS neuromorphic chip with a scalable architecture for learning in networks of spiking neurons," in *IEEE CICC,* pp.1-4, Sept. 19-21, 2011.

[101] T. J. Sejnowski, and G. Tesauro, "The Hebb rule for synaptic plasticity: algorithms and implementations," *Neural models of plasticity: Experimental and theoretical approaches*, pp. 94-103, Academic Press, 1989.

[102] S. Song, K. D. Miller, and L. F. Abbott, "Competitive Hebbian learning through spike-timing-dependent synaptic plasticity," *Nature neuroscience* vol. 3, no. 9 pp. 919-926, 2000.

[103] Y. Dan and M. Poo, "Spike timing-dependent plasticity of neural circuits," *Neuron* vol. 44, issue 1, pp. 23-30, Sept. 2004.

[104] N. Caporale and Y. Dan, "Spike timing-dependent plasticity: a Hebbian learning rule," *Annu. Rev. Neurosci.,* vol. 31, pp. 25-46, 2008.

[105] G. Bi and M. Poo, "Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type." *J. neuroscience,* vol. 18, no. 24, pp.10464-10472, 1998.

[106] T. Masquelier, R. Guyonneau, and S. J. Thorpe, "Competitive STDP-based spike pattern learning," *Neural computation*, vol. 21, no. 5, pp.1259-1276, 2009.

[107] C. Zamarreño-Ramos, L. A. Camuñas-Mesa, J. A. Pérez-Carrasco, T. Masquelier, T. Serrano-Gotarredona and B. Linares-Barranco, "On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex," *Frontiers in neuroscience,* vol. 5, Mar. 2011.

[108] G. Indiveri, E. Chicca and R. Douglas, "A VLSI array of low-power spiking neurons and bistable synapses with spike-timing dependent plasticity," *IEEE Trans. Neural Networks,* vol.17, no.1, pp.211-221, Jan. 2006.

[109] A. Afifi, A. Ayatollahi and F. Raissi, "CMOL implementation of spiking neurons and spike-timing dependent plasticity", *Int. J. Circ. Theor. Appl.*, vol. 39, no. 4, pp. 357-372, April 2011.

[110] Y. Wu, S. Yu; H.-S.P. Wong, Y. Chen, H. Lee, S. Wang, P. Gu, F. Chen and M. Tsai, "AlOx-Based Resistive Switching Device with Gradual Resistance Modulation for Neuromorphic Device Application," in *IEEE IMW*, pp.1-4, May 20-23, 2012.

[111] S. Yu, B. Gao, Z. Fang, H. Yu, J. Kang, H.-S.P. Wong, "A neuromorphic visual system using RRAM synaptic devices with Sub-pJ energy and tolerance to variability: Experimental characterization and large-scale modeling," *IEEE IEDM,* vol. 10, no.4, pp.1-4, Dec. 10-13, 2012.

[112] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder and W. Lu, "Nanoscale memristor device as synapse in neuromorphic systems." *Nano let.,* vol. 10, no. 4 pp.1297-1301, 2010.

[113] G. S. Snider, "Spike-timing-dependent learning in memristive nanodevices," *IEEE Nanoscale Architectures,* pp.85-92, Jun. 12-13, 2008.

[114] Z. Q. Wang, H. Y. Xu, X. H. Li, H. Yu, Y. C. Liu and X. J. Zhu, "Synaptic learning and memory functions achieved using oxygen ion migration/diffusion in an amorphous InGaZnO memristor,"*Adv. Funct. Mater.,* vol. 22, issue 13, pp. 2759-2765, 2012.

[115] Y. Li, Y. Zhong, L. Xu, J. Zhang, X. Xu, H. Sun and X. Miao, "Ultrafast synaptic events in a chalcogenide memristor," *Scientific Reports*, vol. 3, pp. 1619-1626, April 2013.

[116] M. Suri, M. O. Bichler, D. Querlioz, G. Palma, E. Vianello, D. Vuillaume, C. Gamrat, B. DeSalvo, "CBRAM devices as binary synapses for low-power stochastic neuromorphic systems: Auditory (Cochlea) and visual (Retina) cognitive processing applications," *IEEE IEDM,* vol. 10, no. 3, pp. 1-4, Dec. 10-13, 2012.

[117] L. S. Smith, "Implementing neural models in silicon," *Handbook of nature-inspired and innovative computing*, pp. 433-475, Springer US, 2006.

[118] G. Indiveri, "Neuromorphic silicon neuron circuits," *Frontiers in neuroscience,* vol. 5, 2011.

[119] C. Mead and M. Ismail, *Analog VLSI implementation of neural systems*. Vol. 80. Springer Science & Business Media, 2012.

[120] A. Van Schaik, "Building blocks for electronic spiking neural networks," *Neural networks,* vol. 14, no. 6, pp. 617-628, 2001.

APPENDIX A

VERILOG-A CBRAM COMPACT MODEL

```
module cbram_model (an,ca);
inout an,ca;
electrical an, ca;

//structure parameters
parameter real rho_f = 5e-4;
parameter real rho_e = 8e2;
parameter real l = 60e-9;
parameter real rcell = 2.5e-6;
parameter real kb = 1.38e-23;
parameter real kb1 = 8.617e-5;
parameter real T0 = 300.15;

//filament parameters

parameter real q = 1.602e-19;
parameter real h0 = 10e-9;
parameter real rmin = 0.1e-9;
parameter real vh = 0.5;
parameter real vr = 0.1;
parameter real alpha = 0.4;
parameter real beta = 0.25;
parameter real beta1 = 0.22;
parameter real ea = 0.46;
parameter real c = 0.2;
parameter real delta = 0;
parameter real icomp = 1e-6;
parameter real roff = 10e6;
parameter real vwrite = 0.1;
parameter real verase = -0.05;
parameter real zecaf = 18.826e3;
parameter real mag2se = 294.7;
parameter real na = 6.022e23;
parameter real rho_ag2se = 8.216;
parameter real a = 240e-10;
parameter real rth = 1e5;
real h,vol,r,store,r_limit, wa,jhop,jhop_on,e,dhdt,drdt,rsf,rse,re,rf,r,m,n,sgn,state,
    dh,dr,rd,t, vol;

//set initial conditions
analog begin@(initial_step)
begin
vol = 0;
m = 0;
i_dummy = 0;
n = rmin;
```

```
r = rmin;
h= h0;
store =0;
r_limit =0;
dhdt = 0;
drdt=0;
state = 0;
vol = 0;
T = T0;
end

// filament growth rate calculation during write/read

dhdt = vh*exp(-Ea/(Kb1*T))*sinh((alpha*q*(V(an,ca)))/(Kb*T));

        if (h == L) begin
                drdt = vr*exp(-Ea/(Kb1*T))*sinh((beta*q*(V(an,ca)-delta))/(Kb*T));
                dr = drdt;
                store = 1;
                end
        else begin
                store = 0;
                dh = dhdt;
        end
        end
        if (I(an,ca) >= Icomp) begin
                dr = 0;
                dh = 0;
       end

// calculate filament dissolution rate during erase

else if (V(an,ca) < Verase) begin
        drdt = vr*exp(-Ea/(Kb1*T))*sinh((beta1*q*(V(an,ca)))/(Kb*T));
 if (r > rmin) begin
       dr = drdt;
  end
 else begin
    dr = 0;
    dhdt = vh*exp(-Ea/(Kb1*T))*sinh((alpha*q*(V(an,ca)))/(Kb*T));
  if (h>h0) begin
       dh = dhdt;
 end
 else begin
    store = 0;
 end
```

```
 end
 end

else begin

        dhdt = 0;
       drdt = 0;
end

//time integration of filament dimension parameters

n = idt(dr,rmin);
m = idt(dh,h0);

Rse = ((rho_e*L)/(3.14*((rcell*rcell)- (r*r)))); //overall device resistace

Rsf = ((rho_f*h + rho_e*(L-h))/(3.14*r*r)); // conductive filament resistance

Rd = 1/((1/Rsf)+(1/Rse));

// set CBRAM resistance to LRS or HRS
if(r >  rmin) begin
R = Rd;
end
else begin
//R = Roff;
R = (rho_f*h0)/(3.14*rmin*rmin);
end

// signum function
if (V(an,ca) >= 0)
        begin
               sgn = 1;
         end
else begin
               sgn = -1;
end
// set minimum and maximum bounds for filament height
if (m >= L)
      begin
                 h=L;
       end
if (L > m > h0)
      begin
               h=m;
       end
```

```
if (m <= h0)
      begin
              h = h0;
      end

if (v(an,ca) >= 0) begin
sgn = 1;
end
else begin
sgn = -1;
end

// set minimum bound for filament radius
if (rmin < n)
begin
r = n;
end
if (n < rmin)begin
r = rmin;
n= rmin;
end

//Set compliance current check
if (V(an,ca)/R >= Icomp)
        begin
        I(an,ca) <+ Icomp;
        vol = Icomp*R;
        end

else
        begin
        I(an,ca) <+ V(an,ca)/R;
        vol = V(an,ca);
        end

end

endmodule
```

BIOGRAPHICAL SKETCH

Debayan Mahalanabis received the B.E. degree in Electrical Engineering from Jadavpur University, Kolkata, India, in 2009, and the M.S. degree in Electrical Engineering from the University of Southern California, Los Angeles, CA, USA, in 2011. He is currently pursuing the Ph.D. degree in electrical engineering with Arizona State University, Tempe, AZ, USA. His current research interests include characterization, modeling and circuit applications of emerging resistive memory devices and CMOS logic reliability analysis.