Resistive Switching Memory and Reconfigurable Devices

by

Jiantao Zhou

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering)
in the University of Michigan

2016

Doctoral Committee:

Professor Wei Lu, Chair
Assistant Professor Emmanouil Kioupakis
Assistant Professor Becky L. Peterson
Associate Professor Zhaohui Zhong

Table of Contents

List of Figures ..... **v**
List of Tables ..... i**x**
Abstract ..... **x**
Chapter 1 Overview **1**
1.1 Background 1
1.2 Emerging Nonvolatile Memory 2
1.3 Resistive Random Access Memory 5
1.4 Crossbar Array RRAM Architecture 8
1.4.1 Integration Configuration 8
1.4.2 Sneak-Path Issue and Solutions 10
1.5 Organizations of Thesis 11
1.6 References 12
Chapter 2 Simulation Study of Selector Devices within Crossbar RRAM Arrays **16**
2.1 Introduction 16
2.2 Simulation Framework of Crossbar RRAM Array 17
2.3 Selector Parameter Dependence 23
2.4 Bias Configurations 28
2.5 Influence of Interconnect Resistance 33
2.6 Optimal Sense Resistance 34
2.7 Storage Node Parameters 35
2.8 Conclusions 37
2.9 References 38
Chapter 3 Conduction Mechanism and Array Application of Tantalum Oxide Selector **41**
3.1 Introduction 41
3.2 Device Fabrication and Measurement Setup 42
3.3 Conduction Mechanism of Tantalum Oxide Selectors 43
3.4 Multilayer Oxide Based Selector-less RRAM Devices 49

3.5 Crossbar Array Simulation for TaO_x/HfO_2 RRAM Devices 52
3.6 Conclusions 54
3.7 References 55
Chapter 4 Low Programming-Current Resistive Switching Memory 58
4.1 Introduction 58
4.2 Copper Based RRAM with Polycrystalline Silicon In-cell Resistor 59
4.2.1 Device Structure and Fabrication 59
4.2.2 Device Characteristics 60
4.3 Copper Based RRAM with Self-Rectifying Characteristics 64
4.3.1 Device Concept 64
4.3.2 Amorphous Silicon Based Selector Devices 66
4.3.3 Device Structure and Fabrication 68
4.3.4 Low-current Operation and Self-Rectifying Characteristics 68
4.3.5 Retention and Endurance 73
4.3.6 Multilevel Capability 74
4.3.7 Device Variability and Optimization 75
4.4 Conclusions 76
4.5 References 78
Chapter 5 Reconfigurable Devices Based on Complex Oxide Heterojunctions 81
5.1 Introduction 81
5.2 Sample Preparation and Device Fabrication 84
5.3 Electrical Measurement 85
5.3.1 Virgin State and Forming Process 85
5.3.2 Reconfigurable Modulation 88
5.3.3 Cycling Operations 92
5.3.4 Switching Dynamics: Stress Test 93
5.3.5 Switching Dynamics: Pulse Test 95
5.4 Microscopic Origins and Simulation Study 98
5.5 Conclusions 105
5.6 References 106
Chapter 6 In-Memory Computing Based on Crossbar RRAM Arrays 110
6.1 Introduction 110
6.2 Computing Architecture and Design Principle 112
6.3 Proof-of-Concept Demonstration 115

6.3.1 Device Characterization and Array Fabrication 115
6.3.2 Prototype Circuit Setup 116
6.3.3 Implementation of NOR Logic and 1-bit Full Adder 117
6.3.4 Implementation of 4-bit Multiplier 121
6.4 Performance Analysis 124
6.5 Conclusions 125
6.6 References 125
Chapter 7 Summary and Future Work 127
7.1 Simulation Study of Selector Element 127
7.2 Tantalum Oxide Selector Device 128
7.3 Ultra-low Power Conductive Bridge RAM (CBRAM) 128
7.4 Field-Induced Conductance Switching in the $LaAlO_3/SrTiO_3$ Interface 129
7.5 In-Memory Computing Using Crossbar RRAM 129
7.6 Future Work 130
7.7 References 132

List of Figures

Figure 1.1. (a) Internal temperature profile of PCRAM during SET/RESET. PCRAM cell in the amorphous (b) and crystalline (c) states. …..……....3
Figure 1.2. Typical architecture for a 1transistor-1MTJ (1T-1R) STTRAM cell. …..4
Figure 1.3. Two-terminal sandwiched RRAM structure…….....5
Figure 1.4. Classification of RRAM Devices according to the switching polarity.….6
Figure 1.5. Schematic of a 2D crossbar RRAM array.…...……..8
Figure 1.6. Horizontal stacked 3D crossbar RRAM array. …….....8
Figure 1.7. Schematic of a vertical 3D crossbar architecture……..9
Figure 1.8. Schematic of the sneak-path issue in a crossbar RRAM array……10
Figure 2.1. Schematic of a crossbar memory architecture with different read schemes. …….....18
Figure 2.2. *I-V* curves of different selectors (a) and 1S1R memory cells (b). ……...20
Figure 2.3. (a) Read margin as a function of the crossbar array size at different selector on-currents and a fixed selector on/off ratio.(b) Read margin as a function of the crossbar array size at different on/off ratios and a fixed selector on-current. (c) 3D and 2D contour plots of the read margin as a function of the array size and the selector on/off ratio. ……...24
Figure 2.4. (a) Effect of selector on/off ratio on read margin; (b) Effect of selector on-state current on read margin. Bias (c) and current (d) of the selector device in the target cell. ……..25
Figure 2.5. Read margins in crossbar arrays with different selectors. ……...28
Figure 2.6. Comparison of read margins using the four read schemes and in the ideal case without sneak currents. ……..29
Figure 2.7. The sneak-path coefficients defined for different read schemes at LRS (a) and HRS (b). ……..30
Figure 2.8. Overall power consumption of the entire crossbar array for different read schemes……..32
Figure 2.9. 2D contour plots showing the read margin affected by interconnect resistance using the GN-FT scheme. ……..33
Figure 2.10. Read margin's dependence on sense resistor. …….34
Figure 2.11. Relationship between the read margin and the storage node parameters.36
Figure 3.1. (a) SEM image of the as-fabricated Pd/TaO_x/Ta/Pd devices in a crossbar structure. (b) High resolution cross-sectional TEM image showing different layers. (c) EDS spectrum of the TEM specimen tested in different locations (d) I-V characteristics of a Pd/TaO_x/Ta/Pd selector device measured from 100 consecutive tests. …….45
Figure 3.2. Band diagram of the Pd/TaO_x/Ta/Pd selector device under positive (a) and negative bias (b). (c) Temperature dependence of

device *I-V* (150 K to 300 K). (d) ln(I_1) vs. $(V-\Delta\phi)^{1/2}$ plot for the voltage range from 1.75 V to 2.5V. (e) ln(I_2/V^2) vs. $1/V$ plot for the voltage range from -1.6 V to -2 V. (f) Simulated *I-V* curves at different temperatures. ……46

Figure 3.3. (a) Bipolar resistive switching observed in the selector-less $Pd/TaO_x/HfO_2/Pd$ RRAM device. (b) Endurance of the selector-less device. (c) Cumulative distributions of V_{set} and V_{reset}. (d) Room temperature (RT) retention of HRS and LRS. ……50

Figure 3.4. (a) Simulated *I-V* curve of the $Pd/TaO_x/HfO_2/Pd$ device. (b) Calculated read margin under different read voltages (V_D=2.1 V and 1.8 V) for three read schemes: V/2, V/3 and Ground……...53

Figure 4.1. (a) Linear and log (inset) scale *I–V* curve showing sub-1nA current operation. (b) Device structure schematic. ……60

Figure 4.2. Linear fitting for LRS current vs. V^2 indicating SCLC as the conduction mechanism……..61

Figure 4.3. (a) Elevated temperature retention test (6 hours at 85 °C). (b) 10000 cycle pulse data indicating robust endurance. ……..…62

Figure 4.4. (a) Simulated read margin for an N×N crossbar array using the grounding scheme. (b) Read current as a function of line resistance for N = 512.....63

Figure 4.5. (a) *I-V* plots for the MSM selectors (40nm×40nm) with different silicon thicknesses. (b)(c) Cross-sectional TEM images for MSM selector. …..66

Figure 4.6. (a) Band diagram of the MSM selector with ultrathin undoped amorphous silicon. (b) Trap assisted tunneling (TAT) and defect-free tunneling models that consistently explain the device *I-V*……...67

Figure 4.7. Conceptual device schematic of a $Cu/Al_2O_3/aSi/Ta$ RRAM cell at LRS….....67

Figure 4.8. SEM image of two $Cu/Al_2O_3/aSi/Ta$ devices sharing a Ta bottom electrode…….68

Figure 4.9. Typical *I-V* curve of a $Cu/Al_2O_3/aSi/Ta$ device. Upper inset: data plotted in linear scale. Lower inset: resistance variations of HRS/LRS over 50 DC cycles……...69

Figure 4.10. Typical *I-V* curve of a Cu/aSi/Ta device in log and linear scale (inset)…70

Figure 4.11. The electrode effect on the aSi rectifying layer performance……..…71

Figure 4.12. (a) *log(I)-V* plots of the Cu/aSi/Ta control device (before/after scaling) and the $Cu/Al_2O_3/aSi/Ta$ memory device in LRS. (b) Temperature dependence of the Cu/aSi/Ta device. The *I-V* of control samples can be modeled by F-N tunneling at high bias (inset)……73

Figure 4.13. (a) HRS/LRS retention over 10000 sec measured at 100 °C. (b) Pulse endurance test without current compliance. ……...74

Figure 4.14. (a) Different compliance currents I_{cc} employed during SET. (b) Multibit data retention measured from single device at room temperature……...75

Figure 4.15. (a) cycle-to-cycle variation of V_{set} and device-to-device variation of V_{form} for cells with 4.5nm aSi. (b) V_{set} and V_{form} as a function of the aSi thickness…….76

Figure 5.1. Band diagram of the $LaAlO_3/SrTiO_3$ heterostructure……..82

Figure 5.2. Possible mechanisms for 2DEG formation in oxide heterostructure.....83

Figure 5.3. Schematic and optical microscopic image of the device structure……...84
Figure 5.4. Output curves of a virgin $LaAlO_3/SrTiO_3$ transistor……86
Figure 5.5. Drain current and gate current measured during the forming process of the $LaAlO_3/SrTiO_3$ transistor……..87
Figure 5.6. Stabilized output curves of the transistor after the forming process.....87
Figure 5.7. Extracted electron gas (a) carrier concentration and (b) mobility at different temperatures…….88
Figure 5.8. Output curves of (a) the ON state programmed by 300sec 5V stress; and (b) the OFF state erased by 300sec -5V stress……89
Figure 5.9. Transfer curves and current window of the ON and OFF states in the $LaAlO_3/SrTiO_3$ transistor. ……..90
Figure 5.10. Capacitance-voltage (*C-V*) curve for the gate capacitance. ……....90
Figure 5.11. Data retention of the ON state and the OFF state……91
Figure 5.12. (a) Drain current and (b) threshold voltage measured over 50 P/E cycle.92
Figure 5.13. (a) Normalized modulation of drain current with 5V/-5V stress at 280K (b) Temperature dependence of current modulation with 5V stress……93
Figure 5.14. (a) Arrhenius plot of the required stress time as a function of temperature under 5V stress. (b) Required stress time as a function of the applied positive bias at a fixed temperature (295K)…….94
Figure 5.15. Drain current I_{ds} change measured after repeated positive (4V, #1-100) and negative (-4V, #101-200) gate pulses at different temperatures……96
Figure 5.16. The accumulated I_{ds} change measured after 100 positive pulses (100ms) as a function of temperature and pulse amplitude……...97
Figure 5.17. Pulse-width dependency of I_{ds} change measured at 220K. …….98
Figure 5.18. Band diagrams of oxide heterointerfaces in the proposed device under positive (a) and negative V_g (b)……..99
Figure 5.19. Schematic illustrations of (a) the ON state and (b) the OFF state in the $LaAlO_3/SrTiO_3$ transistor. (c) 1D mesh setup for the numerical simulation across the gate direction……102
Figure 5.20. Workflow of the MATLAB based simulations……103
Figure 5.21. Simulated density of oxygen vacancies n_{Vo} in LAO/Al_2O_3 and electrons n_e in STO at different moments. ……...104
Figure 5.22. Measured and simulated I_{ds} change as a function of stress time at 280K…….105
Figure 6.1. (a) The von Neumann Bottleneck results from the frequent data flow across the entire memory hierarchy. (b) In-memory/near-data processing may reduce the data movement and alleviate the related issues concerning energy, bandwidth and delay…….111
Figure 6.2. Schematic block diagram of the proposed architecture…….112
Figure 6.3. Schematic of the NOR logic implementation using the half V_{dd} scheme……..113
Figure 6.4. Proposed procedures for parallel search of input pattern and direct read-out of function result…….114
Figure 6.5. (a) Optical microscopic image of the as-fabricated RRAM array. (b) Structure Schematic of single device and crossbar array……...115
Figure 6.6. (a) Measured and simulated device *I-V*. (b) HRS and LRS resistance

distribution measured from 20 randomly chosen devices……..116
Figure 6.7. Circuit demonstration of the in-memory computing based on RRAM array……..117
Figure 6.8. Block diagram of the designed PCB and its working setup……...117
Figure 6.9. Measured current response of NOR logic while applying ½V_{dd} on Inputs A/B and V_{dd} on Output C. ……118
Figure 6.10. Proposed programing process of a 1-bit full adder. (a) Decomposed NOR logics for the 1-bit FA. (b) Programming pulses on the WLs. (c) Resistance mapping after programming……119
Figure 6.11. Simulated read-out process of the 1-bit FA. ……..120
Figure 6.12. (a) Measured resistance mapping after writing the functions into the RRAM array for 4 input cases. (b) Measured CAM current from the BLs during data identification process and the output results directly read from the WLs… ……...……...……..121
Figure 6.13. Building a 4-bit MULT from a 2-bit MULT and a 2-bit FA.....121
Figure 6.14. Array resistance mappings of the pre-programmed 2-bit MULT (a) and 2-bit FA (b). ……..…122
Figure 6.15. Simulated read-out process of the 4-bit MULT: (a) Current response during the read of the 2-bit multiplication. (b) Current response during the read of the 2-bit summation. ……...123
Figure 6.16. Average energy cost of the proposed FA/MULT implementation. ……124
Figure 6.17. Read-out current ratio (“1” vs. “0”) as a function of R_{on}/R_{off} ratio and array size……...125

List of Tables

Table 2.1. Simulation parameters for read operations……22
Table 3.1. Simulation parameters for *I-V* curves of the Pd/TaO_x/Ta/Pd selector…...48

Abstract

The demand in the data storage – from mobile devices to enterprise applications - has been driving the explosive development of non-volatile memories (NVMs). Based on the field effect transistor (FET), flash memory has benefitted from the geometric scaling and dominated the electronics market over the past decades. However, as Moore's Law approaches its fundamental limit, there will be little reward from the process aspect. Flash memory is facing increasing challenges from technical issues as well as economic concerns. The emergence of Internet of Things (IoT) also brings up new challenges including faster access and low-voltage operation that are not compatible with flash memory. Therefore, for a feasible next-generation NVM solution with high storage density, high performance and low fabrication cost, a novel non-FET based replacement of flash memory is urgently desired. Among different candidates, resistive switching memory (RRAM) has attracted broad interest due to its simple structure, high speed, long retention, excellent endurance and energy efficiency.

In this work, we present studies on resistive switching memories and related reconfigurable devices. First, we systematically investigate the "sneak-path" issue of crossbar RRAM arrays and specify the device requirement of nonlinear selector element in the one-selector-one-resistor (1S1R) configuration. Through numerical simulations, we analyze the crossbar array from a perspective of device-circuit interaction and propose optimized benchmarks for the future improvement of RRAM and selector devices.

Next, we develop a tantalum oxide (TaO_x) based selector device which exhibits high

nonlinearity (~10^4) and good uniformity. Verified by experimental observation and theoretical model, the underlying conduction mechanism of this selector is attributed to thermionic emission and tunneling emission. A HfO_2 switching layer is integrated with the proposed selector to constitute a self-rectifying RRAM cell with high LRS selectivity (~5×10^3), which can potentially enable large-scale crossbar array (up to 1Mbit) with less than 4% degradation in read margin.

Further, we demonstrate sub-nA operation current in a Cu based conductive bridge RAM (CBRAM) device for the first time, which offers significant energy savings during program and read steps. An improved $Cu/Al_2O_3/aSi/Ta$ cell with a built-in barrier/rectifying layer is developed to enhance device reliability. Apart from low current, other attractive properties including high on/off ratio (>100x), retention (over 10^4 seconds at 100°C) and endurance (500 cycles without external current compliance) can be obtained. The proposed process is fully compatible with mainstream CMOS back-end-of-line (BEOL) integration.

Additionally, we explore coupling the ionic migration process in resistive switching devices with transistor operation. A reconfigurable top-gate transistor structure is developed for the $LaAlO_3/SrTiO_3$ heterojunction system which offers a two-dimensional electron gas (2DEG) at the oxide interface. By incorporating ionic processes in the gate stack, we show that the channel conductivity can be modulated in a non-volatile manner by an external electric field.

Finally, we propose a novel in-memory computing architecture using crossbar RRAM arrays, which breaks the boundary between computing and memory and offers high parallelism. We design the basic protocols and demonstrate a prototype circuit. As proof-of-concept verification, a 1-bit full adder and a 4-bit multiplier are designed and verified.

Chapter 1

Overview

1.1. Background

With surging storage demand driven by the big data and the Internet of Things (IoT), the market of non-volatile memory (NVM) has been growing rapidly and became a main driver for the entire semiconductor industry over the past decade. In particular, NAND flash has enjoyed great success and dominated the NVM market owing to the aggressive node shrinking and the reduced price from a dollar-per-bit perspective.

However, further scaling of NAND flash will inevitably face serious challenges in terms of technology and cost. On one hand, the operations of NAND will be susceptible to several major technical issues, including (1) crosstalk between adjacent cells; (2) device variability due to the limited trapped electrons; (3) compromised endurance from dielectric damage and (4) retention degradation caused by random telegraph noise (RTN) [1], [2]. On the other hand, the reduction in memory price from planar size scaling may be offset by the soaring process-development cost and the unsatisfactory yield [3].

To explore the scaling feasibility and extend the Moore's Law in the NVM segment, different 3D NAND solutions have been proposed [4]–[8]. By stacking memory layers and adopting less advanced process nodes, high-density NAND arrays can be constructed vertically, thus avoiding the process obstacles in purely lateral scaling. However, given the fundamental physics limitation associated with transistors, 3D NAND will eventually

come across the same bottlenecks in the next few generations.

Since critical applications are becoming more data-centric today, the design of modern large-scale high-performance system are required to overcome the giant gap between memory (SRAM and DRAM) and storage (flash and HDD disk). The concept of storage-class memory (SCM) has been proposed to bridge these two distinct hierarchies [9], [10]. An ideal SCM element is expect to possess fast access time of DRAM, high integration density of flash and nonvolatile features allowing low-power applications.

Therefore, not only the challenges in scaling flash memory, but also the future computing needs on SCM propel the development of new memory technology.

1.2. Emerging nonvolatile memory

To find a replacement for NAND flash and a candidate of future SCM element, several novel nonvolatile memories have been proposed, including Phase-Change Random Access Memory (PCRAM) [11]–[13], Ferroelectric Random Access Memory (FeRAM) [14], [15], Spin-Transfer Torque Random Access Memory (STTRAM) [16]–[18] and Resistive Random Access Memory (RRAM)[19]–[21]. In this section, PCRAM, FeRAM and STTRAM will be briefly introduced.

PCRAM exploits the resistance difference between crystalline (low resistance) and amorphous (high resistance) phases of the phase-change material (generally chalcogenide alloy, such as $Ge_2Sb_2Te_5$) for memory application [22]. The material phase is determined by the internal temperature, and its transition can be controlled by applying external pulses that provide necessary Joule heating. Once rapidly heated to the melting point (>600 °C) by a high current pulse, the phase-change material loses the crystallinity, and eventually

becomes amorphous and resistive after fast cooling down (quenching). If another medium pulse is applied to keep a temperature between the glass transition temperature and the melting point for long enough time (~100ns), the material will crystallize and transform back into the conductive state. Relying on the thermal process, PCRAM consumes large power and has endurance issues due to the thermal stress and related material degradations [23]. The minimum set pulse length also limits the speed of PCRAM.

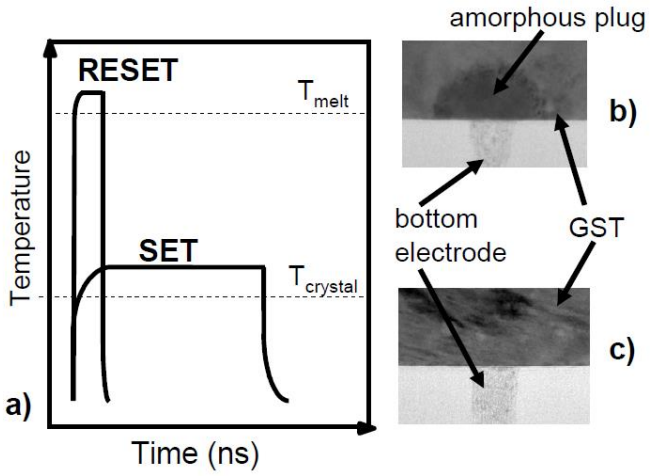


Figure 1.1. (a) Internal temperature profile of PCRAM during SET/RESET. PCRAM cell in the amorphous (b) and crystalline (c) states. Reproduced from [12]

FeRAM achieves the non-volatile memory functionality by storing different dipole moment states in a variable ferroelectric capacitor which typically consists of a lead zirconate titanate (PZT) layer [14], [15]. With an external electric field applied across the dielectric, the atoms/ions will shift in the direction of the field and align the dipoles accordingly. The atomic positions and the corresponding polarization states can be used to store “0” and “1” in the binary logic. FeRAM offers an advantage of extremely low energy consumption (<0.1pJ/bit). However, the scaling issue and the destructive read operation constrain the wide application of FeRAM.

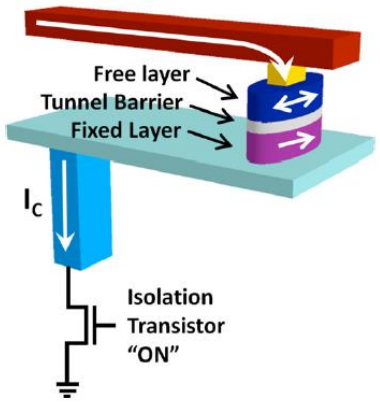


Figure 1.2. Typical architecture for a 1transistor-1MTJ (1T-1R) STTRAM cell. Compatible with the CMOS technology, the STTRAM device can be integrated during the back-end-of-line (BEOL) process. Reproduced from [18]

STTRAM is an emerging subset of magnetic random access memory (MRAM). The memory element of STTRAM is typically comprised of a MgO-based magnetic tunnel junction (MTJ) where the storage layer and the reference magnetic layers are separated by a tunnel barrier [17]. Due to the magnetoresistance effect, the MTJ shows high (low) resistance when the storage layer exhibits the magnetizations antiparallel (parallel) to the reference layer. Different from conventional toggle MRAMs, STTRAM uses spin-polarized current, instead of an external magnetic field, to switch the magnetization of the storage layer. As a result, STTRAM requires a substantially reduced switching energy (a few pJ) and overcomes several challenges facing conventional MRAMs such as the crosstalk issue. In the meantime, STTRAM inherits the merits of long endurance (10^{16}) and fast switching (~10ns) from MRAM. However, the process reliability of STTRAM has to be proven prior to large-scale manufacturing due to the use of many (typically > 20) non-conventional material layers. Scalability and cost are the other challenges facing STTRAM applications.

1.3. Resistive Random Access Memory

Resistive Random Access Memory (RRAM) is based on a simple two-terminal MIM (metal-insulator-metal) structure, with a switching medium sandwiched by the top and bottom electrodes, as show in Figure1.3. By applying external electric signals on the electrodes, the resistance of a RRAM device can be modulated and used to store different states for memory and computing application. The change in resistance value can be understood by the formation/rupture of conductive filaments within the (much more resistive) switching medium. Once a filament is formed (ruptured), the device exhibits low (high) resistance state.

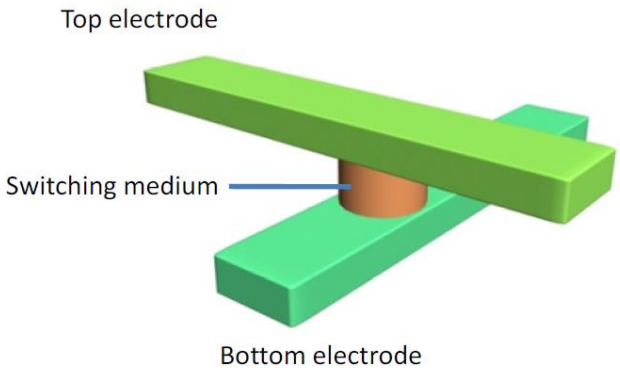


Figure 1.3. Two-terminal sandwiched RRAM structure. Reproduced from [24]

Typically, the resistive switching process in a RRAM device includes three operations: Forming, SET and RESET. Forming (or electroforming) is performed to initialize the as-fabricated RRAM cell and create the conductive filament for the first time. Afterwards, the filament material typically will not be fully removed and subsequent SET and RESET processes are based on the remaining filament structure. SET refers to the completion of the filament and the switching process from high resistance state (HRS) to low resistance state (LRS). RESET corresponds to the rupture of the filament and the associated process

from LRS to HRS. To prevent the device from over-programming, a compliance current is usually applied during the forming and SET/RET operations.

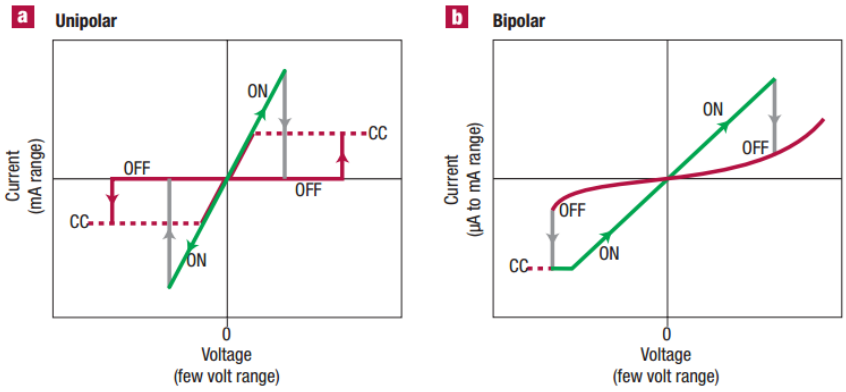


Figure 1.4. Classification of RRAM Devices according to the switching polarity: (a) Unipolar RRAM; (b) Bipolar RRAM. Reproduced from [25]

RRAM can be classified into various categories according to different criteria, such as switching polarity (bipolar vs. unipolar) and switching mechanism (cation vs. anion).

As show in Figure 1.4, the switching of a unipolar RRAM device is not dependent on the polarity of the applied bias. With either positive or negative voltage, the unipolar RRAM device can be SET and RESET. Unipolar RRAM typically requires high RESET current, and its resistive switching is essentially believed as a thermal process determined by the internal temperature profile of device [21]. On the other hand, the SET and RESET transitions must occur at different polarities in a bipolar RRAM cell. Bipolar switching is mainly a consequence of ionic motion driven by the electric field.

Up to now, more than 50 types of materials have been investigated for RRAM applications. In most cases, the related resistive switching process originates from the migration of either cations or anions. Specifically, cation based RRAM devices are usually named as conductive bridge random access memory (CBRAM) or electrochemical

metallization memory (ECM); while anion based devices are termed as valence-change memory (VCM) or oxide-RRAM [19]. In CBRAM or ECM, active metal atoms (Cu, Ag, et al) are ionized and injected into the switching medium (SiO_2, Al_2O_3, et al). Driven by the electric field, these metal cations migrate through the switching layer, and eventually become reduced to neutral atoms again. The accumulation of the metal atoms finally forms the metallic filaments during the SET and Forming processes [26], [27]. VCM devices are typically based on perovskite oxides and transitional metal oxides (TMO) [28]–[34]. The resistance states of VCM are tuned by the migration of oxygen vacancies (V_O). The filament formation during the SET process of VCM can be explained by the creation of an oxygen-deficient (Vo-rich) oxide region, which is more conducting than the stoichiometric host oxide material. During RESET it is generally believed that Joule heat assists the rupture of the filament in unipolar VCMs; while bipolar VCMs rely mostly on the V_O drifting under electric field but can still be assisted by Joule heating. It is worth noting that apart from the filamentary resistive switching characteristics similar to ECM, VCM can also exhibit non-filamentary switching behaviors decided by the modulation of the effective Schottky barrier height at the electrode/oxide interface [35].

Compatible to the mainstream CMOS technology, high performance RRAM devices have been demonstrated with large on/off ratios [26], excellent endurance [31], long retention [36], fast switching speed [37] and high integration density [24]. The potential commercialization of RRAM is subject to improvement on process yield and device variability as well as careful verifications on large-scale array implementation [38].

1.4. Crossbar Array RRAM Architecture

1.4.1. Integration Configuration

Due to its simple two-terminal structure, RRAM can be easily integrated into a crossbar array architecture. A 2D crossbar array consists two groups of parallel electrodes (top electrodes and bottom electrodes) orthogonally crossing each other and sandwiching the switching medium between the electrodes (Fig. 1.5). With one memory device defined at each cross-point, the crossbar array provides a minimal device area of $4F^2$, which F corresponds to the smallest features size. Once several crossbar arrays are stacked on top of each other (Fig. 1.6), the effective device area can be further scaled to $4F^2/n$, where n stands for the stack number. The high packing density makes crossbar RRAM array highly attractive as one of the most promising candidates for post-NAND memory applications.

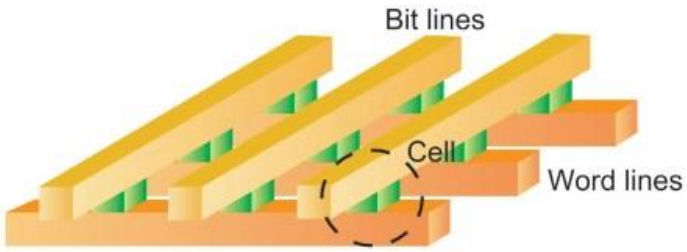


Figure 1.5. Schematic of a 2D crossbar RRAM array. 200nm. Reproduced from [35].

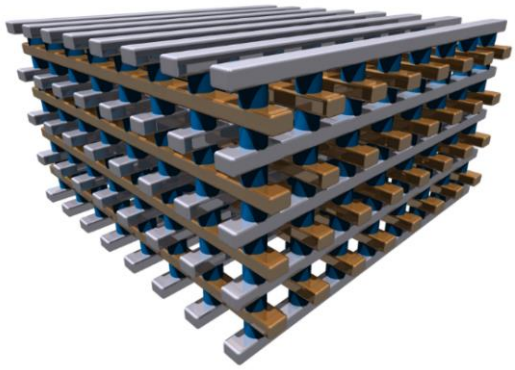


Figure 1.6. Horizontal stacked 3D crossbar RRAM array. Reproduced from [39]

Based on the 2D crossbar structure, two types of 3D crossbar RRAM array configurations have been proposed. As shown in Figure 1.6, the stacked 3D crossbar architecture defines the memory cell horizontally and enhances the storage density by simply stacking memory layers. This design hides the area of selection devices and allows high cell efficiency. However, a possible negative factor that limits this architecture from replacing the NAND technology is the potential higher cost, since the number of fabrication steps and masks increases when more memory layers are stacked.

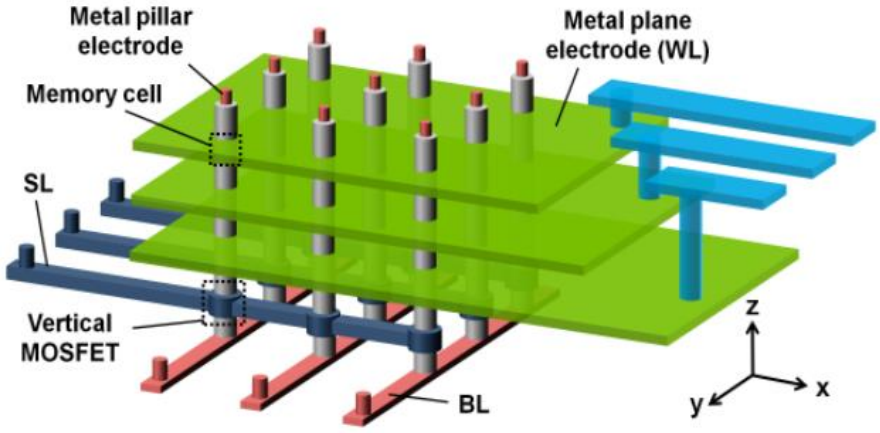


Figure1.7. Schematic view of a vertical 3D crossbar architecture (a) using the sidewall RRAM cell and a vertical MOSFET (b) as the bit-line selector. Reproduced from [40].

Compared to the horizontal stacked configuration, the vertical RRAM (VRRAM) solution adopts an integration scheme similar to the VNAND, and is potentially more cost effective [40], [41]. In the VRRAM design the RRAM cells are formed at the sidewalls of the vertical electrodes. The number of masks is relatively independent of the stack number, since different stacks can share the same step of lithography and etching. However, the VRRAM structure involves increased process complexity, such as high aspect ratio etching and deposition with excellent step coverage.

1.4.2. Sneak-Path Issue and Solutions

A fundamental challenge of crossbar architecture comes from the difficulty in device selection and isolation, or the so-called "sneak-path" issue. As illustrated in Figure 1.8, while accessing a specific memory cell, the current flows through not only the selected cell ($I_{element}$), but also numerous unselected cells (I_{sneak}). Since the peripheral sensing circuit cannot distinguish these two different current components, an incorrect cell state may be read for the selected cell. Besides affecting read, the sneak current also affects cell programming, by increasing the current the peripheral circuitry has to supply while reducing the voltage actually applied on the selected cell which makes SET/RESET difficult.

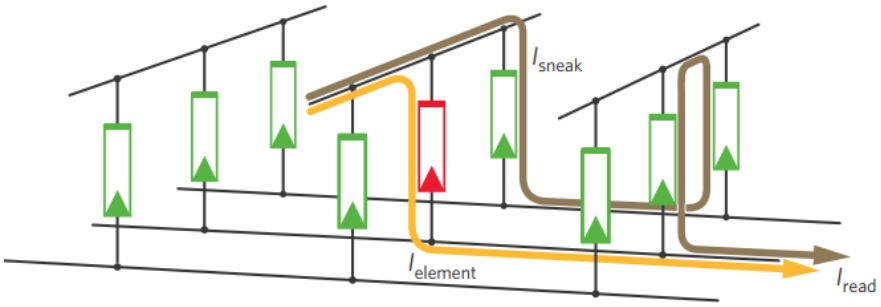


Figure 1.8. Schematic of sneak-path issue in a crossbar RRAM array: $I_{element}$ corresponds to the current through the selected cell (red) during read process. I_{sneak} corresponds to the potential sneak leakage current from unselected cells (green). Reproduced from [42].

To address the sneak-path issue and ensure correct array operations, memory cells with highly nonlinear current-voltage (I-V) characteristics are needed in the crossbar memory array. By suppressing the leakage current at low bias, the sneak currents from the unselected cells (which are biased at a lower voltage than the selected cell) can be minimized.

The I-V nonlinearity can be achieved either by internally tuning the memory element

and creating a self-rectifying cell, or by serially connecting an external nonlinear selector device and forming a one-selector-one-resistor (1S1R) configuration. Generally, both solutions introduce the nonlinearity by integrating selective/rectifying layer(s) with the switching medium. On a structure point of view, the main difference is that, the 1S1R system needs to insert a middle metal electrode, which is not required in a self-rectifying cell. On one hand, the middle electrode explicitly separates the selector and the RRAM element, and provides the 1S1R structure with more flexibility in optimizing both components independently. On the other hand, the middle electrode requires additional process steps that are difficult to adapt with the economic VRRAM scheme. Therefore, in general the 1S1R structure provides better device performance than the self-rectifying cell structure, such as nonlinearity and on/off ratios; while the self-rectifying cell is more integration friendly and less expensive.

1.5. Organizations of Thesis

This Ph.D. thesis is organized as follows.

In Chapter 1, the basic background of RRAM and crossbar memory array is briefly reviewed.

In Chapter 2, SPICE simulations are performed to investigate the selector device requirements for crossbar RRAM array operation. Theoretical guidelines on selector development and 1S1R systems are provided from a perspective of device-circuit interaction.

In Chapter 3, a tantalum-oxide selector device with high nonlinearity and excellent uniformity will be discussed. The conduction mechanism is systematically studied by

experiment and modeling. Integrated with this selector, a self-rectifying RRAM cell is further developed and evaluated for its potential crossbar array applications.

Chapter 4 focuses on the low-power RRAM devices. Two different novel CBRAM cell structures will be discussed to achieve low operating current without sacrificing other critical performance.

Chapter 5 explores integrating RRAM with transistors and building reconfigurable devices in the $LaAlO_3/SrTiO_3$ heterojunction. Nonvolatile modulation on transistor current is realized by controlling the ionic migration within the complex oxide systems.

Chapter 6 discusses a novel in-memory computing architecture using crossbar RRAM arrays. Prototype circuit implementation is demonstrated to verify the architecture concept.

Chapter 7 summarizes this thesis and proposes several directions for future research.

1.6. References

[1] C.-Y. Lu, K.-Y. Hsieh, and R. Liu, "Future challenges of flash memory technologies," *Microelectron. Eng.*, vol. 86, no. 3, pp. 283–286, 2009.

[2] S. Lai, "Non-volatile memory technologies: The quest for ever lower cost," in *2008 IEEE International Electron Devices Meeting*, pp. 1–6, 2008.

[3] A. Goda and K. Parat, "Scaling directions for 2D and 3D NAND cells," in *2012 International Electron Devices Meeting*, pp. 2.1.1–2.1.4, 2012.

[4] H. Tanaka, M. Kido, K. Yahashi, M. Oomura, R. Katsumata, M. Kito, Y. Fukuzumi, M. Sato, Y. Nagata, Y. Matsuoka, Y. Iwata, H. Aochi, and A. Nitayama, "Bit Cost Scalable Technology with Punch and Plug Process for Ultra High Density Flash Memory," in *2007 IEEE Symposium on VLSI Technology*, pp. 14–15, 2007.

[5] R. Katsumata, M. Kito, Y. Fukuzumi, M. Kido, H. Tanaka, Y. Komori, M. Ishiduki, J. Matsunami, T. Fujiwara, Y. Nagata, L. Z. L. Zhang, Y. Iwata, R. Kirisawa, H. Aochi, and A. Nitayama, "Pipe-Shaped BiCS Flash Memory with 16 Stacked Layers and Multi-Level-Cell Operation for Ultra High Density Storage Devices," in *2009 Symposium on VLSI Technology*, pp. 136–137, 2009.

[6] J. J. J. Jang, H.-S. K. H.-S. Kim, W. C. W. Cho, H. C. H. Cho, J. K. J. Kim, S. I. S. S. Il Shim, Y. J. Y. Jang, J.-H. J. J.-H. Jeong, B.-K. S. B.-K. Son, D. W. K. D. W. Kim, K. Kihyun, J.-J. S. J.-J. Shim, J. S. L. J. S. Lim, K.-H. K. K.-H. Kim, S. Y. Y.

S. Y. Yi, J.-Y. L. J.-Y. Lim, D. C. D. Chung, H.-C. M. H.-C. Moon, S. H. S. Hwang, J.-W. L. J.-W. Lee, Y.-H. S. Y.-H. Son, U.-I. C. U.-I. Chung, and W.-S. L. W.-S. Lee, "Vertical cell array using TCAT(Terabit Cell Array Transistor) technology for ultra high density NAND flash memory," *2009 Symp. VLSI Technol.*, pp. 14–15, 2009.

[7] SungJin Whang, KiHong Lee, DaeGyu Shin, BeomYong Kim, MinSoo Kim, JinHo Bin, JiHye Han, SungJun Kim, BoMi Lee, YoungKyun Jung, SungYoon Cho, ChangHee Shin, HyunSeung Yoo, SangMoo Choi, Kwon Hong, Seiichi Aritome, SungKi Park, and SungJoo Hong, "Novel 3-dimensional Dual Control-gate with Surrounding Floating-gate (DC-SF) NAND flash cell for 1Tb file storage application," in *2010 International Electron Devices Meeting*, vol. 79, pp. 29.7.1–29.7.4, 2010.

[8] H.-T. Lue, T.-H. Hsu, Y.-H. Hsiao, S. P. Hong, M. T. Wu, F. H. Hsu, N. Z. Lien, S.-Y. Wang, J.-Y. Hsieh, L.-W. Yang, T. Yang, K.-C. Chen, K.-Y. Hsieh, and C.-Y. Lu, "A highly scalable 8-layer 3D vertical-gate (VG) TFT NAND Flash using junction-free buried channel BE-SONOS device," in *2010 Symposium on VLSI Technology*, pp. 131–132, 2010.

[9] R. F. Freitas and W. W. Wilcke, "Storage-class memory: The next storage system technology," *IBM J. Res. Dev.*, vol. 52, no. 4.5, pp. 439–447, 2008.

[10] G. W. Burr, "Storage Class Memory," 2010.

[11] S. Lai, "Current status of the phase change memory and its future," in *2003 International Electron Devices Meeting*, pp. 10.1.1-10.1.4, 2003.

[12] M. J. Breitwisch, "Phase Change Memory," in *2008 IEEE International Interconnect Technology Conference*, vol. 23, no. 16, pp. 219–221, 2008.

[13] S. Raoux, F. Xiong, M. Wuttig, and E. Pop, "Phase change materials and phase change memory," *MRS Bull.*, vol. 39, no. 08, pp. 703–710, 2014.

[14] J. F. Scott, "Device physics of ferroelectric memories," *Ferroelectrics*, vol. 183, no. 1, pp. 51–63, 1996.

[15] H. Kohlstedt, Y. Mustafa, A. Gerber, A. Petraru, M. Fitsilis, R. Meyer, U. Böttger, and R. Waser, "Current status and challenges of ferroelectric memory devices," *Microelectron. Eng.*, vol. 80, pp. 296–304, 2005.

[16] S. Tehrani, J. M. Slaughter, M. Deherrera, B. N. Engel, N. D. Rizzo, J. Salter, M. Durlam, R. W. Dave, J. Janesky, B. Butcher, K. Smith, and G. Grynkewich, "Magnetoresistive random access memory using magnetic tunnel junctions," *Proc. IEEE*, vol. 91, no. 5, pp. 703–714, 2003.

[17] Jian-Gang Zhu, "Magnetoresistive Random Access Memory: The Path to Competitiveness and Scalability," *Proc. IEEE*, vol. 96, no. 11, pp. 1786–1798, 2008.

[18] K. L. Wang, J. G. Alzate, and P. Khalili Amiri, "Low-power non-volatile spintronic memory: STT-RAM and beyond," *J. Phys. D. Appl. Phys.*, vol. 46, no. 7, p. 074003, 2013.

[19] R. Waser, R. Dittmann, G. Staikov, and K. Szot, “Redox-Based Resistive Switching Memories - Nanoionic Mechanisms, Prospects, and Challenges,” *Adv. Mater.*, vol. 21, no. 25–26, pp. 2632–2663, 2009.

[20] H. S. P. Wong, H. Y. Lee, S. Yu, Y. S. Chen, Y. Wu, P. S. Chen, B. Lee, F. T. Chen, and M. J. Tsai, “Metal-oxide RRAM,” in *Proceedings of the IEEE*, vol. 100, no. 6, pp. 1951–1970, 2012.

[21] Y. Yang and W. Lu, “Nanoscale resistive switching devices: mechanisms and modeling,” *Nanoscale*, vol. 5, no. 21, p. 10076, 2013.

[22] H.-S. P. Wong, S. Raoux, S. Kim, J. Liang, J. P. Reifenberg, B. Rajendran, M. Asheghi, and K. E. Goodson, “Phase Change Memory,” *Proc. IEEE*, vol. 98, no. 12, pp. 2201–2227, 2010.

[23] L. Goux, D. Tio Castro, G. A. M. Hurkx, J. G. Lisoni, R. Delhougne, D. J. Gravesteijn, K. Attenborough, and D. J. Wouters, “Degradation of the Reset Switching During Endurance Testing of a Phase-Change Line Cell,” *IEEE Trans. Electron Devices*, vol. 56, no. 2, pp. 354–358, 2009.

[24] S. H. Jo, K. H. Kim, and W. Lu, “High-density crossbar arrays based on a Si memristive system,” *Nano Lett.*, vol. 9, no. 2, pp. 870–874, 2009.

[25] R. Waser and M. Aono, “Nanoionics-based resistive switching memories,” *Nat. Mater.*, vol. 6, no. 11, pp. 833–840, 2007.

[26] S. H. Jo, K.-H. Kim, and W. Lu, “Programmable Resistance Switching in Nanoscale Two-Terminal Devices,” *Nano Lett.*, vol. 9, no. 1, pp. 496–500, 2009.

[27] A. Belmonte, W. Kim, B. T. Chan, N. Heylen, A. Fantini, M. Houssa, M. Jurczak, and L. Goux, “A thermally stable and high-performance 90-nm Al_2O_3\Cu-based 1T1R CBRAM cell,” *IEEE Trans. Electron Devices*, vol. 60, no. 11, pp. 3690–3695, 2013.

[28] Y. B. Nian, J. Strozier, N. J. Wu, X. Chen, and A. Ignatiev, “Evidence for an Oxygen Diffusion Model for the Electric Pulse Induced Resistance Change Effect in Transition-Metal Oxides,” *Phys. Rev. Lett.*, vol. 98, no. 14, p. 146403, 2007.

[29] T. Fujii, M. Kawasaki, A. Sawa, Y. Kawazoe, H. Akoh, and Y. Tokura, “Electrical properties and colossal electroresistance of heteroepitaxial $SrRuO_3/SrTi_{1-x}Nb_xO_3$ ($0.0002 \leqslant x \leqslant 0.02$) Schottky junctions” *Phys. Rev. B*, vol. 75, no. 16, p. 165101, 2007.

[30] C.-Y. Lin, C.-Y. Wu, C.-Y. Wu, T.-C. Lee, F.-L. Yang, C. Hu, and T.-Y. Tseng, “Effect of Top Electrode Material on Resistive Switching Properties of ZrO_2 Memory Devices,” *IEEE Electron Device Lett.*, vol. 28, no. 5, pp. 366–368, 2007.

[31] M.-J. Lee, C. B. Lee, D. Lee, S. R. Lee, M. Chang, J. H. Hur, Y.-B. Kim, C.-J. Kim, D. H. Seo, S. Seo, U.-I. Chung, I.-K. Yoo, and K. Kim, “A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta_2O_{5-x}/TaO_{2-x} bilayer structures,” *Nat. Mater.*, vol. 10, no. 8, pp. 625–630, 2011.

[32] 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 HfO_2 based RRAM,” in *2008 IEEE International Electron Devices Meeting*, pp. 1–4, 2008.

[33] W. C. Chien, Y. R. Chen, Y. C. Chen, A. T. H. Chuang, F. M. Lee, Y. Y. Lin, E. K. Lai, Y. H. Shih, K. Y. Hsieh, and Chih-Yuan Lu, “A forming-free WO_x resistive memory using a novel self-aligned field enhancement feature with excellent reliability and scalability,” in *2010 International Electron Devices Meeting*, pp. 19.2.1–19.2.4, 2010.

[34] C. Yoshida, K. Tsunoda, H. Noshiro, and Y. Sugiyama, “High speed resistive switching in Pt/TiO_2/TiN film for nonvolatile memory application,” *Appl. Phys. Lett.*, vol. 91, no. 22, p. 223510, 2007.

[35] A. Sawa, “Resistive switching in transition metal oxides,” *Mater. Today*, vol. 11, no. 6, pp. 28–36, 2008.

[36] N. Banno, M. Tada, T. Sakamoto, M. Miyamura, K. Okamoto, N. Iguchi, T. Nohisa, and H. Hada, “A Fast and Low-Voltage Cu Complementary-Atom-Switch 1Mb Array with High-Temperature Retention,” vol. 35, no. 2011, pp. 202–203, 2014.

[37] S.-S. Sheu, P.-C. Chiang, W.-P. Lin, H.-Y. Lee, P.-S. Chen, Y.-S. Chen, T.-Y. Wu, F. T. Chen, K.-L. Su, M.-J. Kao, K.-H. Cheng, and M.-J. Tsai, “A 5ns fast write multi-level non-volatile 1 K bits RRAM memory with advance write scheme,” in *2009 Symposium on VLSI Circuits*, pp. 82–83, 2009.

[38] “International Technology Roadmap for Semiconductors (ITRS),” 2013.

[39] S. H. Jo, “Recent Progress in RRAM Materials and Devices,” 2015.

[40] H.-Y. Chen, S. Yu, B. Gao, P. Huang, J. Kang, and H.-S. P. Wong, “HfO_x based vertical resistive random access memory for cost-effective 3D cross-point architecture without cell selector,” in *2012 International Electron Devices Meeting*, vol. 2, pp. 20.7.1–20.7.4, 2012.

[41] I. G. Baek, C. J. Park, H. Ju, D. J. Seong, H. S. Ahn, J. H. Kim, M. K. Yang, S. H. Song, E. M. Kim, S. O. Park, C. H. Park, C. W. Song, G. T. Jeong, S. Choi, H. K. Kang, and C. Chung, “Realization of vertical resistive memory (VRRAM) using cost effective 3D process,” in *2011 International Electron Devices Meeting*, pp. 31.8.1–31.8.4, 2011.

[42] E. Linn, R. Rosezin, C. Kügeler, and R. Waser, “Complementary resistive switches for passive nanocrossbar memories,” *Nat. Mater.*, vol. 9, no. 5, pp. 403–406, 2010.

Chapter 2

Simulation Study of Selector Devices within Crossbar RRAM Arrays

2.1. Introduction

As discussed in the Chapter 1, resistive random access memory (RRAM) is inherently compatible to the crossbar array architecture, which potentially enables $4F^2$ scaling, vertical 3D stacking and low fabrication cost [1], [2]. However, with continuously scaled device and electrodes, the sneak path leakage and the interconnect resistance within such a purely passive crossbar array result in major concerns on device isolation and selection.

To mitigate these issues, the memory cell in a high-density crossbar array requires non-linear current-voltage characteristics. Different solutions have been proposed, including 1S1R serial configuration (with an explicit separate selector) [3], self-rectifying cell (with an implicit integrated selector layer) [4] and complementary resistive switching concepts[5]. Despite impressive progress on selector performance[6]–[17], a series of bottlenecks on large current density, high on/off ratios, fast switching speed and process reliability remain to be resolved.

There have been several reports on the relationship between cell performance and crossbar memory array [18]–[25]. However, due to the oversimplification of the system,

an accurate and systematic analysis on the performance metrics required for selector devices to achieve successful crossbar array operation was still missing, which poses a significant challenge for RRAM array development given the increased emphasis on selector device development and integration. For example, the recent ITRS report specified performance requirements of the selectors based mainly on the individual cell level [26]. However, we expect that a different set of figures-of-merits (FOA) are needed to ensure proper operation at the memory array level.

In this chapter, a comprehensive numerical study on crossbar RRAM array operation based on the 1S1R structure is presented. Based on the SPICE simulations, the selector device requirements and the device-circuit interaction are carefully investigated, in particularly for the read operation. The aim is to provide theoretical references for future selector design and optimization. Specifically, the remainder of chapter is organized as follows. Section 2.2 describes the simulation framework of crossbar RRAM arrays. The device model and bias schemes used for the simulations are discussed here. Section 2.3 provides specifications of the selector performance metrics at different array sizes to obtain optimal read margins. Section 2.4 discusses the choice of bias schemes to improve read margins and reduce power consumptions. The effects of the interconnect resistance, different sensing circuitry and resistive storage node parameters are analyzed in Section 2.5-2.7, respectively. The final conclusion of this chapter is summarize in Section 2.8.

2.2. Simulation Framework of Crossbar RRAM Array

Typically, a crossbar resistive memory array consists of switching cells sandwiched by two sets of parallel electrodes as word-lines and bit-lines, as shown in Figure 2.1. The transition of the cell element between the high-resistance state (HRS) and the low-

resistance state (LRS) occurs when proper voltages are applied on the electrodes. This simple design allows minimal cell size, but the interconnected passive network structure also leads to sneak leakage currents that can severely limit the output read margin. Figure 2.1 illustrates the cause of misreading in crossbar architecture due to sneak path currents through unselected cells. During read operation, parasitic leakage paths through unselected cells in the array lead to inaccurate output signal and can prevent proper identification of the HRS from the LRS for the target cell.

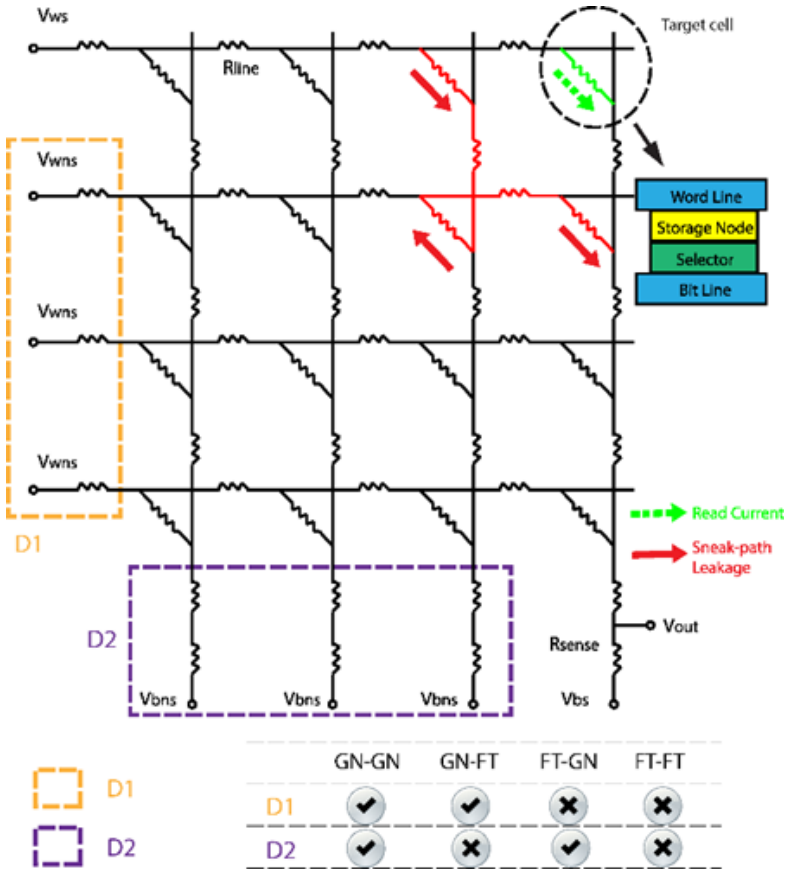


Figure 2.1. Schematic of a crossbar memory architecture with 4 different read schemes. (Inset) cell structure. (Green dashed arrows) The read current through the target cell. (Red solid arrows) Sneak-path currents through unselected cells may degrade the output voltage. (Box D1/D2) The circuitry inside the boxes will be removed to create a read scheme with floating word-lines or bit-lines, respectively. GN-GN, GN-FT, FT-GN and FT-FT stand for different read schemes (see in Section 2.4). Reproduced from [28].

The read operation of a crossbar array is achieved by sensing the output voltage, V_{out}. In practice, sense amplifiers are connected with all bit lines to convert the output current into voltage signal. In our simulation, the peripheral sensing circuit is simplified as a sense resistor R_{sense}. The read margin is defined as

$$RM = \frac{\Delta V_{out}}{V_{read}} = \frac{V_{out}(LRS) - V_{out}(HRS)}{V_{WS}}$$

where V_{ws} is the read voltage applied to the selected word-line, and V_{out}(LRS) and V_{out}(HRS) are the output voltages measured at R_{sense} when reading 1 (LRS) or 0 (HRS). The commonly used ground/ground read scheme was studied first to illustrate the effects of the selector and its performance requirements. In the ground/ground read scheme, to access a specific target cell within the array, we bias the selected word-line to $V_{ws}=V_{dd}$ and the selected bit-lines to $V_{BS}=0$; while leaving all the unselected word-lines and bit-lines grounded ($V_{WNS}=V_{BNS}=0$). This bias mode helps raise the read margin, though it has a major drawback of high power consumption[18]. Several different read schemes will be discussed later in Section 2.4.

Both the data storage pattern in the crossbar RRAM array and the target cell location influence the output voltage swing. Our simulations show that for the commonly used grounding scheme, the worst case for sneak current also results in the worst case for output voltage, and in both cases the worst case corresponds to the unselected cells being in LRS for reading “1” and “0”. Further, due to parasitic interconnect resistance ΔV_{out} will be smallest for the target cell at the farthest corner from the word/bit-line voltage sources. As a consequence, in this study the worst case is chosen as the target cell being at the farthest corner with the unselected cells in LRS.

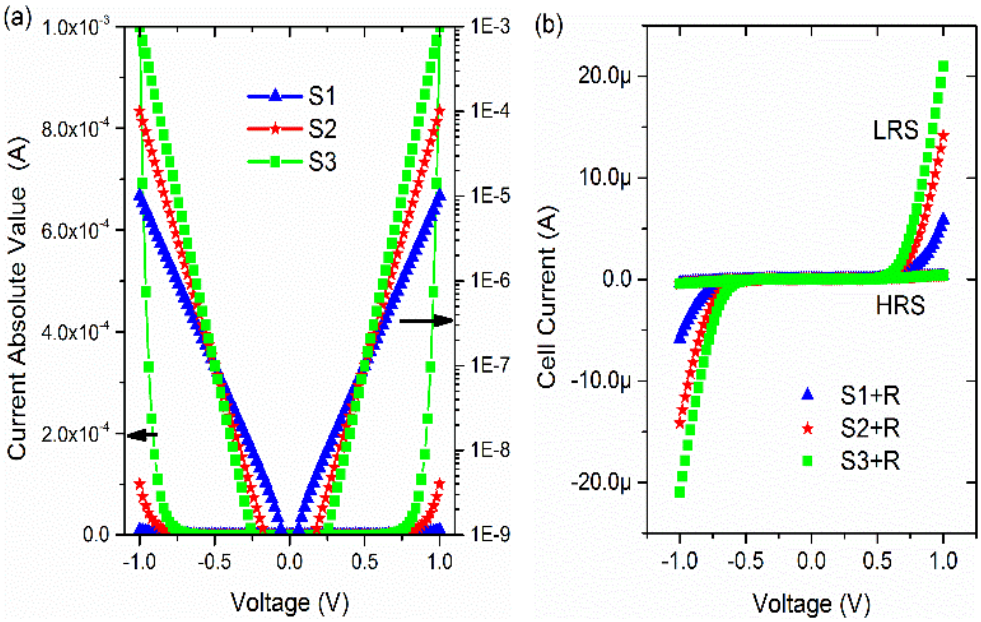


Figure 2.2. (a) I-V characteristics of three different selectors (S1-S3) in linear (solid lines) and log (symbols) scales. S1: $k=10^2$, $I_{sel}(on)=10\mu A$; S2: $k=10^3$, $I_{sel}(on)=100\mu A$; S3: $k=10^4$, $I_{sel}(on)=1000\mu A$. (b) I-V characteristics of memory cells serially connected with the selectors and storage elements in (a). Reproduced from [28].

In this work, we mainly focus on bipolar RRAMs in which set/reset processes occur at different bias polarities. Each memory cell in the crossbar array is composed of a non-volatile storage element serially connected with a symmetric selector device, forming the 1S1R structure schematically shown in Figure 2.1. The storage node is assumed to have fixed values of R_{on} and R_{off} in the LRS and HRS, respectively, i.e. dynamic switching processes are not studied here since read operation is expected not to disturb the cell states. The I-V characteristics of the selector device is modeled by the following hyperbolic sine function,

$$I_{sel} = \gamma \cdot \sinh(\alpha \cdot V)$$

where γ is a conductance parameter, and α represents the nonlinearity of select device. The I-V characteristics of the selector serially connected with the memory element forming the

1S1R structure are shown in Figure 2.2(a) (b). We note that at moderate voltages the hyperbolic sine function becomes the exponential function used in [24], and also agree reasonably with experimentally observed I-V curves for selector devices based on Schottky diode [9], tunnel barrier [10], [17], MIEC [15] and punch-through devices [16].

Considering the maximal voltage dropped on an unselected cell is roughly half of the read voltage on the target cell in the ground/ground read scheme, we can further define the on/off ratio for the selector (which describes the nonlinearity of the selector) as below,

$$k \equiv \frac{I_{sel}\ (on)}{I_{sel}(off)} = \frac{I_{sel}(V_{ws})}{I_{sel}\left(\frac{1}{2}V_{ws}\right)}.$$

Obviously, on/off ratio is independent of the selector conductance prefactor, γ. For simplicity, the definition above merely describes the nonlinear characteristic of the selector itself, measured at V_{WS} = 1 V, without taking the serially connected storage node into account. Furthermore, under idealized circumstances with negligible interconnect resistance, negligible sneak path leakages and negligible voltage divider effect from the select devices [19], the maximum read margin is obtained by setting the sense resistor to be

$$R_{sense} = \sqrt{R_{on}R_{off}}$$

In Section III we will first discuss the effects of selector performance parameters on the crossbar array performance. Optimizations by choosing appropriate R_{on}, R_{off} and R_{sense} will then be investigated. The default parameters used in our simulations are listed in Table 2.1.

Table 2.1. Simulation Parameters for Read Operations. Reproduced from [28].

Parameters	**Values**
High resistance state (R_{off})	1MΩ
Low resistance state (R_{on})	10kΩ
Sense resistor (R_{sense})	100kΩ
Interconnect resistance between neighboring cells (R_{line})	5Ω
Voltage at the selected word line (V_{WS})	1V
Voltage at unselected word lines (V_{WNS})	0V
Voltage at the selected bit line (V_{BS})	0V
Voltage at the selected bit line (V_{BNS})	0V
Selector On/Off ratio (k)	10^4
On-state current of selector (I_{sel}(on))	100μA
Nonlinear factor of selector (α)	18.4207 V^{-1}
Selector Conductance Factor (γ)	2×10^{-12}A

2.3. Selector Parameter Dependence

HSPICE simulations were conducted based on the model mentioned in Section 2.2. The read margin of passive crossbar arrays can be calculated numerically for selectors with different characteristics. Figure 2.3 summarizes the read margin at different array sizes when selectors with different on/off (non-linearity) and on-currents I_{sel}(on) are used.

Figure 2.3a indicates that for a specific selector on/off ratio of 10^4, the swing of read voltage remains relatively constant for small arrays but degrades significantly as the crossbar memory array size increases beyond 64×64. A significant result here is that selectors with larger I_{sel}(on) (hence lower effective resistance R_{sel}) leads to much better read margins (e.g. read margins is improved by more than 50% when I_{sel}(on) is increased from 10μA to 100μA) due to a smaller voltage divider effect from the selector in the serially-connected memory element/selector configuration. Detailed analysis on the HRL and LRS show that the LRS tends to benefit more from the enhanced output voltage compared to the HRS (since the voltage divider effect is more pronounced when the storage element is in the LRS), thus causing a wider read margin. On the other hand, the improvement by increasing I_{sel}(on) alone becomes limited in absolute terms as the array size is increased. As shown in Figure 2.3a, for selectors with $k=10^4$, the read margin drops rapidly below the minimum requirement of 10% when the array exceeds 256×256 due to the sharp increase in sneak currents at very large arrays, regardless of the 100-fold change in selector on-state current.

Introducing selectors with high on/off ratios can potentially eliminate the sneak-path leakage under low reverse bias, and provide more distinguishable output signals on the sense resistor, particularly for large arrays, as shown in Figure 2.3b. Figure 2.3c shows the

variation of the read margin when the selector on/off ratios are changed (while maintaining the same selector on-state current I_{sel}(on)) in crossbar arrays of different sizes. In practice, once the array size and the anticipated read margin are chosen, the allowed parameter range of on/off ratio can be predicted according to the 2-D contour in Figure 2.3c.

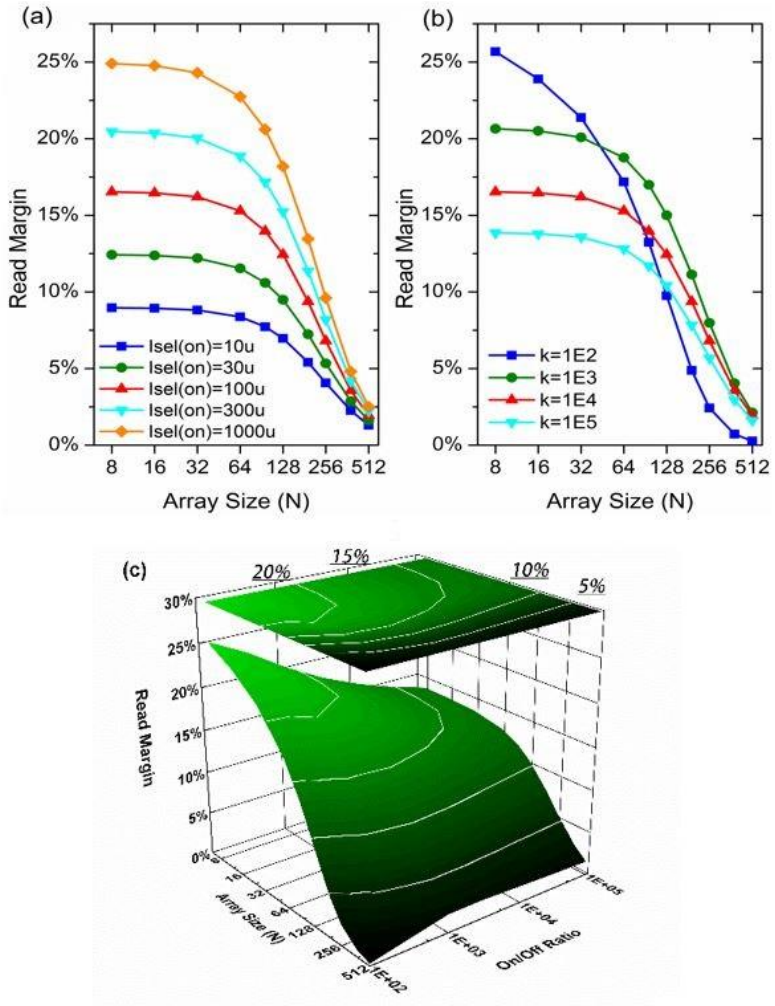


Figure 2.3. (a) Read margin as a function of the crossbar array size at different selector on-currents and a fixed selector on/off ratio k=10^4. (b) Read margin as a function of the crossbar array size at different on/off ratios and a fixed selector on-current I_{sel}(on) = 100μA. (c) 3D and 2D contour plots of the read margin as a function of the array size and the selector on/off ratio. The selector on-current is kept 100μA. Reproduced from [28].

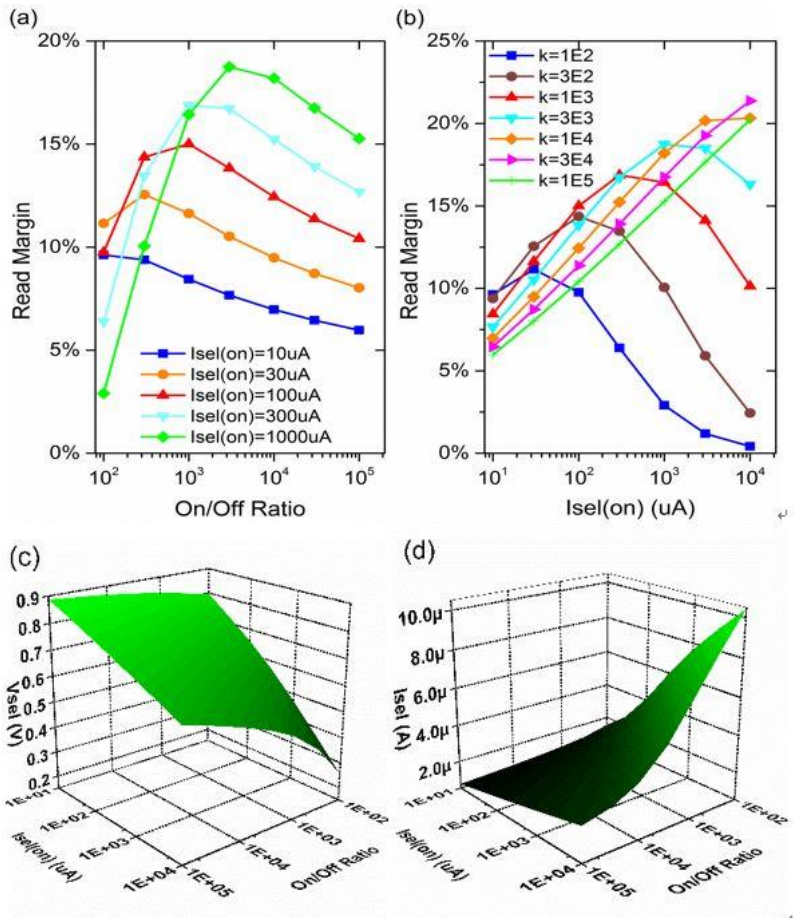


Figure 2.4. (a) Effect of selector on/off ratio on read margin; (b)Effect of selector on-state current on read margin. Bias (c) and current (d) of the selector device in the target cell as functions of selector on-current and on/off ratios during LRS reading. The size of crossbar array is fixed (128×128) in (a)-(d). Reproduced from [28].

One surprising result from Figure 2.3 b-c is that a higher selector on/off ratio (i.e. non-linearity) does not necessarily result in an improved read margin. This observation is counterintuitive and deserves further investigation. To analyze this effect further, we chose a fixed crossbar array size (128×128) as our model system and systematically studied the effects of selector on/off ratios and on-currents I_{sel}(on) on the read margin. The results are shown in Figure 2.4. As shown in Figure 2.4a, for a certain I_{sel}(on), there exists an optimal value of selector non-linearity and further increasing the selector non-linearity in fact leads to a decrease (albeit slowly) in read margin. The preferred position of the selector non-

linearity (on/off ratio) also depends on the selector on-current I_{sel}(on) and shifts to higher values with increasing peak read margin as the selector becomes more conducting.

This non-monotonic dependence on on/off ratios can be explained by two competing mechanisms: on one hand, increasing the on/off ratio alleviates the sneak-path leakage around the target cell; on the other hand, however, a very large on/off ratio means that a high voltage is needed for the selector to maintain the desired current as a small drop in voltage will severely reduce the current through the selector (and the target cell). As shown in Figure 2.4 c-d, with increasing on/off ratios, a larger V_{sel} is consumed by the selector even though the target cell current I_{sel} start to decrease due to the increase in selector nonlinearity (on/off) at a given I_{sel}(on). The reduction of output current accumulated by the sense resistor in turn leads to the reduced read margin at very large on/off.

At what point the non-linear behavior plays an adverse role depends on how conductive the selector is, and this problem is particularly pronounced for selectors having a low I_{sel}(on), as shown in Figure 2.4 a-b. For the sake of completeness, cases with extremely large on-state current are also evaluated in Fig. 4(b). In contrast to Figure 2.4a, not all of the curves in Figure 2.4b can reach their peaks within the simulated parameter range of I_{sel}(on). A selector with very high on/off ratio (e.g. $>10^4$) does not offer any benefit compared to a selector with a low on/off (e.g. 300) until the selector on-current is more than 200μA. This analysis suggests obtaining a high I_{sel}(on) is a key requirement for a nanoscale select device, provided the selection device is of sufficient nonlinearity.

Previous analyses focused mostly on the role of selector on/off while the requirement of selector on-current was normally relaxed. For example, the benchmark prediction of International Technology Roadmap for Semiconductors (ITRS) suggested the selector

should have an on-state current of 1 μA at 1V and an on/off ratio as high as 10^6 [26]. Another widely used reference value for selector current density is 1MA/cm2, corresponding to an on-state current of 4μA for a 20nm×20nm crossbar cell. However, as indicated in Figure 2.5, our analysis shows that with this level of on-state current, the optimal on/off selector ratio is in fact around 10^2, and on/off ratios beyond that will actually hurt the array performance. Also shown in Figure 2.5 is a further optimized selector with on/off = 10^3 and I_{sel}(on) = 100 μA, which produces much higher read margins compared with the ITRS values. Considering an I_{sel}(on) of 1MA/cm^2 is already very challenging to obtain, this study suggests an urgent need to develop selectors with high on-state current.

In a typical selector device two distinct regions can be observed in its I-V characteristics: "subthreshold region" at low bias with sharp increase (e.g. similar to (2)) in current, and "saturation region" at high bias with slowly increasing current. For most reported electrical selectors that are the focus of this study [8]–[10], [15]–[17], their finite "subthreshold slope" limits how fast the selectors can be switched. As a result, the requirements of high Isel(on) and high on/off are not decoupled but instead end up competing with each other. With a given subthreshold slope and Isel(off), a larger on/off means the selector will consume a larger voltage to reach the desired Isel(on) hence the selector will exhibit a more pronounced voltage divider effect. This explains the counterintuitive observation that a high selector on/off actually may lead to a deteriorated read margin. To resolve the dilemma, novel selector devices with very sharp subthreshold slopes will be highly desired [11]–[13], [27].

Additionally, characterizing a selector device with I_{sel}(on) measured at the read voltage V_{read} (e.g. 1V in this study) is not optimal. As shown in Fig.2.4 c-d, during operation the

actual voltage of the selector V_{sel} can be much lower than V_{read}. In other words, the selector device does not need to maintain the exponentially increasing current beyond the actual working bias V_{sel}. A better figure of merit will be the selector subthreshold slope and the saturation current (or saturation voltage). For example, a selector with $k=10^3$ and saturation current of 1.95μA at 0.72V will be equivalent to a selector with the same non-linearity factor and on-current of 100μA measured at 1 V (Optimized2 in Figure 2.5), and provides a read margin of 15% for a 128×128 crossbar array.

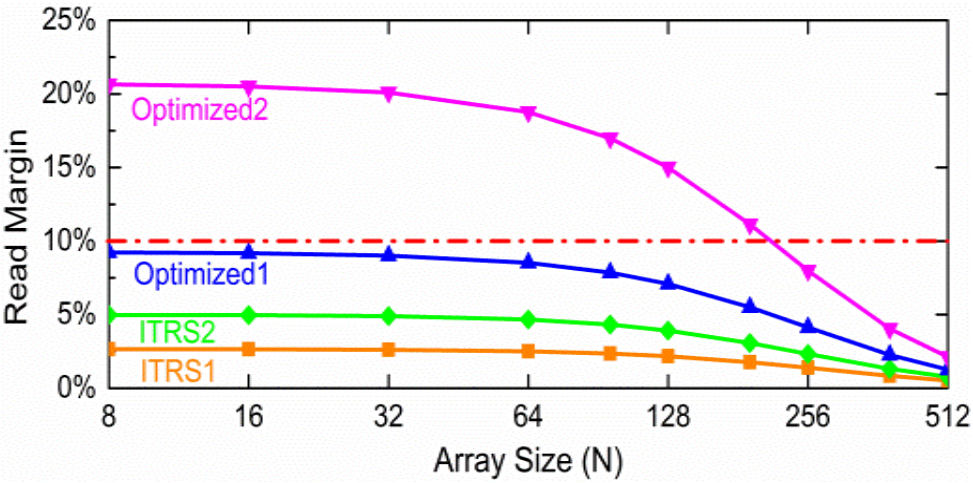


Figure 2.5. Read margins in crossbar arrays with different selectors. ITRS1: $k=10^6$, $I_{sel}(on)=1\mu A$; ITRS2: $k=10^6$, $I_{sel}(on)=4\mu A$; Optimized1: $k=10^2$, $I_{sel}(on)=4\mu A$; Optimized2: $k=10^3$, $I_{sel}(on)=100\mu A$. Reproduced from [28].

2.4. Bias Configurations

The analysis so far employed the commonly used "grounding" scheme with grounded unselected word lines and bit lines (GN-GN) during read [18], [19], [22], [24]. In this section, we examine possible performance improvements from optimizing the read schemes, e.g. by floating, grounding or biasing some of the address lines.

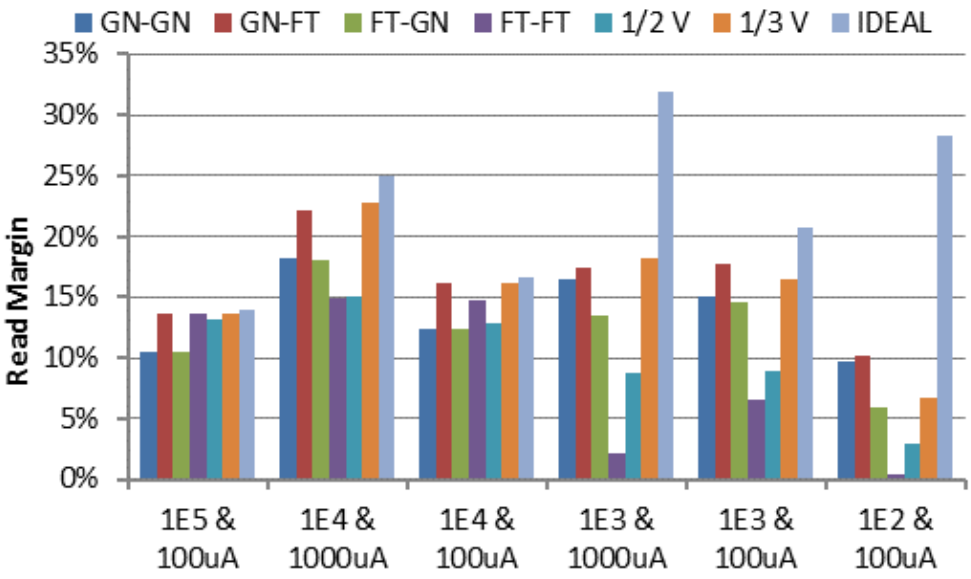


Figure 2.6. Comparison of read margins using the four read schemes and in the ideal case without sneak currents. Six different parameter combinations of selector are used. The crossbar array size is fixed at 128×128. Reproduced from [28].

Here, GN-FT is referred to the read scheme with grounded unselected word lines and floating unselected bit lines. Similar abbreviations apply to other schemes of FT-GN and FT-FT, as shown in Figure 2.1. Additionally, 1/2 V (1/3 V) schemes where the unselected word-lines are biased at 1/2 (1/3) V_{read} and the unselected bit-lines are biased at 1/2 (2/3) V_{read} are also considered. In order to provide clear insight into the different read schemes, we define the sneak-path coefficient as $\theta = I_{sense}/I_{sel}$, where I_{sense}, I_{sel} are the currents flowing through the sense resistor and the target memory cell, respectively. Obviously $\theta=1$ indicates the sneak current is negligible. θ larger than unity corresponds to a forward injection of current into the selected bit-line; while θ smaller than unity corresponds to a reverse flow of current into the unselected word-lines. Analyzing the sneak-path coefficient θ thus provides us information on how sneak current affects the read margin in the different read schemes. For a 128×128 crossbar array, the read margins and the sneak-path

coefficients for the four different read schemes are simulated with various combinations of selector parameters for the target cell in HRS and LRS, respectively (see in Figure 2.6 and Figure 2.7). Results of the ideal case without sneak currents are also included.

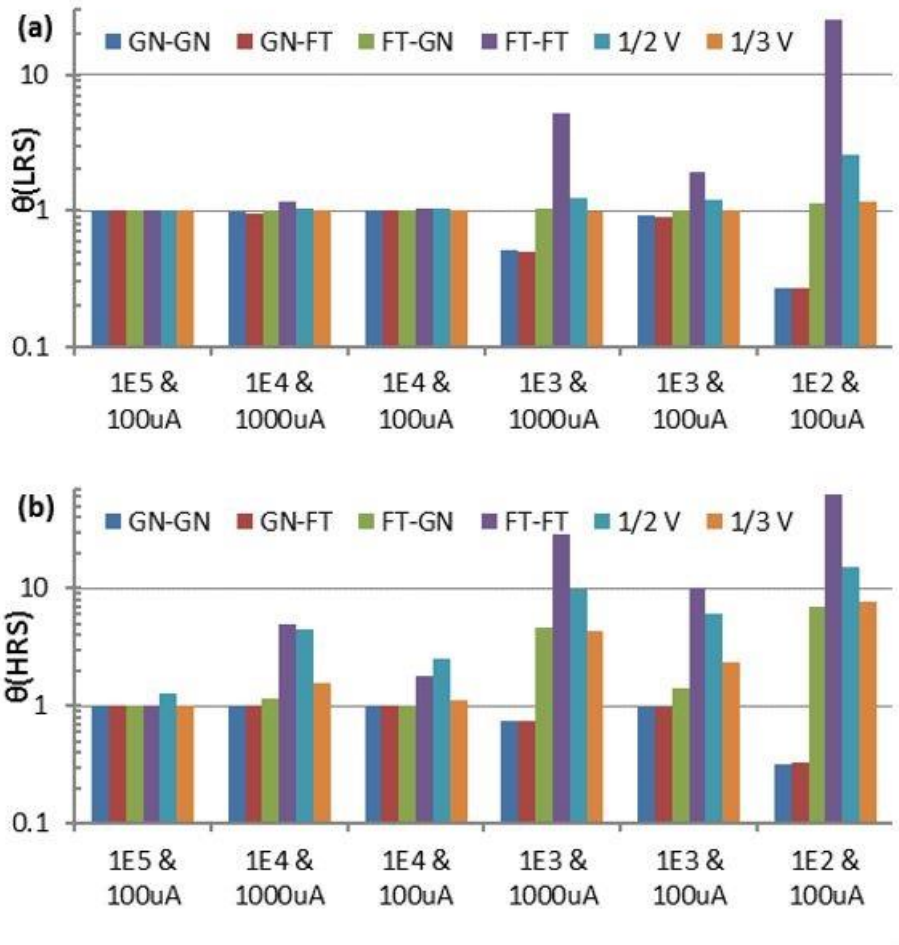


Figure 2.7. Sneak-path coefficients defined for different read schemes within a fixed RRAM array (128×128) at LRS (a) and HRS (b). Reproduced from [28].

In term of read margin, it is clear that for any given scheme the read margin will be improved as θ is made closer to 1. Figure 2.6 suggests that GN-FT and 1/3 V normally offer the best read margins. The lower read margins in FT-GN and GN-GN can be attributed to the downward current paths to ground along the unselected bit-lines in the FT-

GN and GN-GN setups. This leads to an increase in supply current which in turn raises the voltage drop on interconnect series resistance in the selected word-line and results in a reduced effective read voltage across the target cell.

We note the performances of the 1/2 V and FT-FT schemes strongly depend on the switching function of the select device. The potentials at the unselected word-lines are raised by biasing (1/2 V) or by the sneak currents through the unselected bit-lines (FT-FT). Consequently, at low on/off ratios, large amount of current is injected into the selected bit-line and gives rise to an unrecognizable output voltage swing, as shown in Figure 2.7. However, when highly nonlinear selectors are introduced, the sneak path leakage from unselected word-lines will be inhibited significantly and both schemes can exhibit a desirable read margin comparable with that of GN-FT.

Additionally, it remains true that in all configurations an enhanced selector conductance is preferred for better array performance, consistent with earlier conclusions. Further, analyses on the extreme cases illustrate that, when selectors are more "ideal", e.g. with on/off = 10^5 and $I_{sel}(on) = 100\mu A$, the read margins of the GN-FT, FT-FT, 1/2 V and 1/3 V configurations are very similar to the ideal case without any sneak-currents (e.g. Figure 2.6), but the GN-GN and FT-GN configurations still exhibit lower read margins due to downward current paths through the unselected bit-lines to ground in these two configurations.

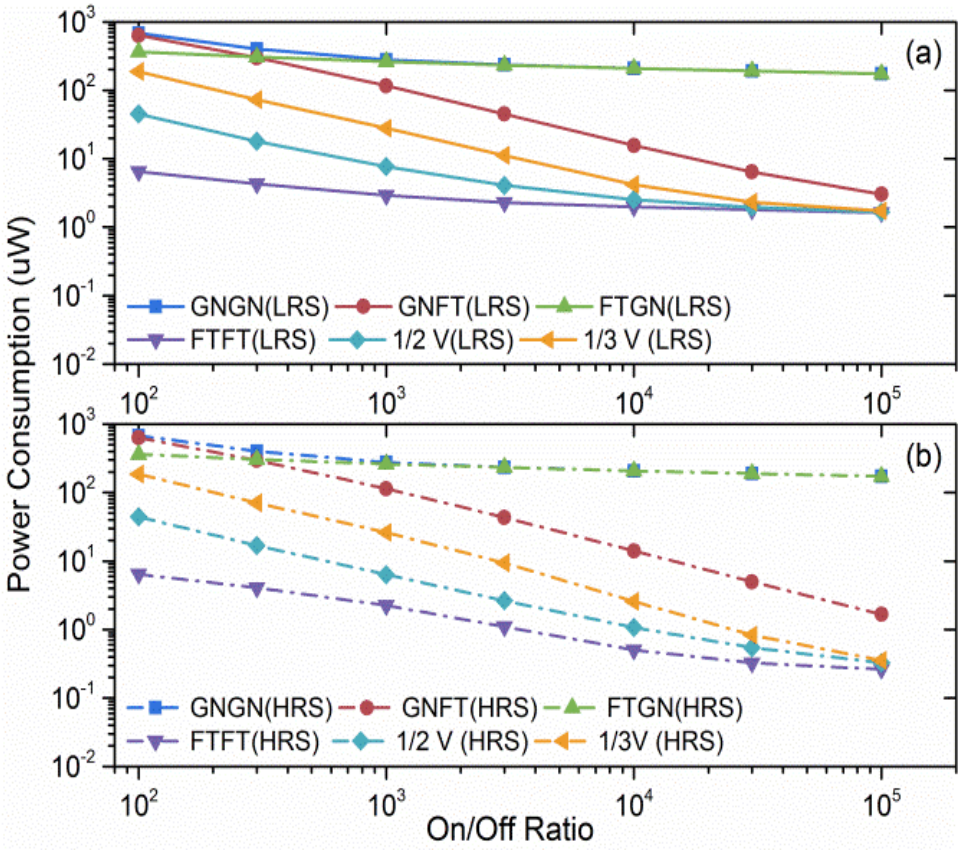


Figure 2.8. The overall power consumption of the entire crossbar array (128×128) with different read schemes for reading 1 (a) and 0 (b). The on-state current of selector is kept at 100μA. Reproduced from [28].

Power consumption during read operation can be another concern and is analyzed in Figure 2.8 for different read schemes. It can be seen that the FT-FT scheme offers by far the lowest power consumption and the GN-GN scheme offers the highest power consumption in general. The 1/2 V also offers improved power dissipation since it minimizes the current through the unselected cells due to its symmetric bias on unselected word-lines and bit-lines. The 1/3 V scheme offers another compromise since the power dissipation can be minimized as the selector becomes more non-linear.

2.5. Influence of Interconnect Resistance

Figure 2.9 shows the read margins influenced by different interconnect resistances using the GN-FT scheme. We note that when the array size is small the read margin is insensitive to the interconnect resistance. However, as the number of cells increases, the interconnect resistance can negatively affect the array performance.

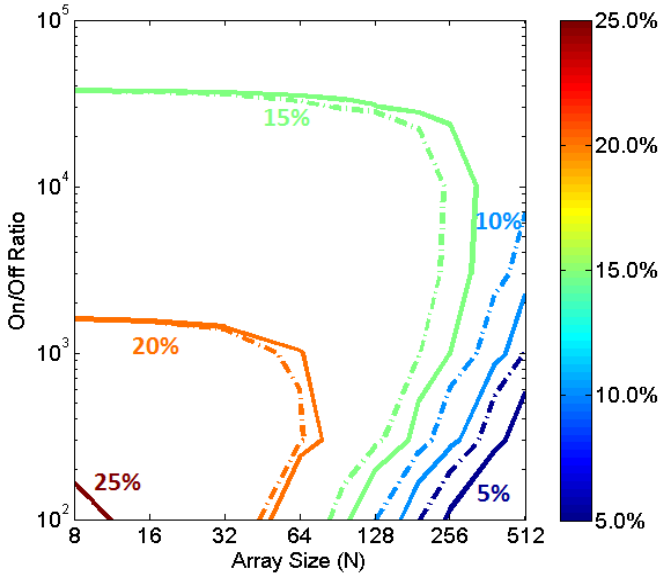


Figure 2.9. 2D contour plots showing the read margin affected by interconnect resistance using the GN-FT scheme. The selector on-state current is kept at 100μA. (Dashed line) R_{line}=5Ω. (Solid lines) R_{line}=2.5Ω. Reproduced from [28].

Three factors from interconnect resistance cause the degradation of read margin: (1) voltage dropped on the selected word-line; (2) voltage dropped on the selected bit-line; (3) the raising or lowering of potential on unselected word-lines that can drive parasitic current into (or out of) the sense resistor. Therefore, even though nonlinear selectors are incorporated in crossbar arrays, reducing the interconnect resistance is still crucial for large array operation. All three effects can be verified by analyzing the potentials on the word-lines and bit-lines at different nodes in the array during simulation.

As can be seen from Figure 2.9, for a fixed array dimension, electrode lines with lower R_{line} allows a larger range of selector on/off ratios for a given read margin. Equivalently, if a selector is already chosen, the maximum array size can be improved with more conducting electrodes for a given read margin.

2.6. Optimal Sense Resistance

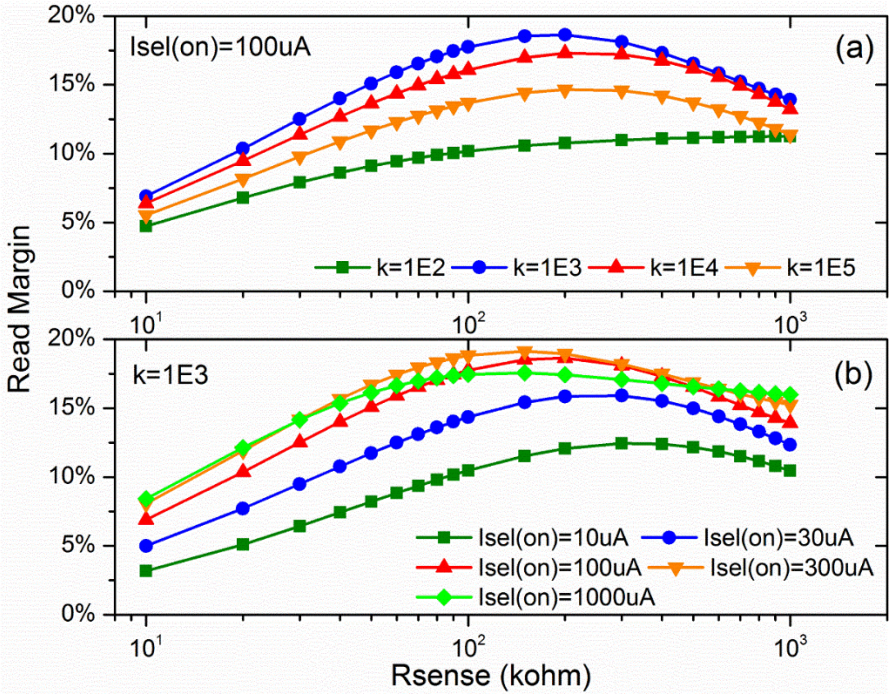


Figure 2.10. Read margin's dependence on sense resistor when (a) selector on-state current I_{sel}(on) is fixed (100μA); (b) selector on/off ratio k is fixed (10^3). Reproduced from [28].

According to the simplified model in [19], the maximum read margin is achieved with a sense resistor equal to the geometric mean of R_{on} and R_{off}, i.e. $R_{sense}= (R_{on*}R_{off})^{1/2}$. Nevertheless, it needs to be examined what the optimal sense resistor values is when the selector device and interconnect resistances are included. The impacts of sense resistor on the read margin is shown in Figure 2.10 a-b, while keeping the selector I_{sel}(on) and on/off

ratio fixed. The results indicate that within a wide range of selector parameter space, the optimal value of sense resistor, R_{sense_op}, remains close to the geometric mean of Ron and Roff. The small shift of R_{sense_op} can be understood by the increase of equivalent cell resistance due to the serially-connected selector or electrode resistance. The more conducting the selector is, the smaller the shift is. Since the read margin varies slowly with the sense resistor, possible trade-off adjustments on the sensing circuitry are feasible when other practical parameters need to be optimized, such as power consumption and fan-out.

2.7. Storage Node Parameters

Below we discuss how the storage node parameters affect the read margin in the selector/storage node configuration. In particular, since we only focus on the read operation, the storage node can be modeled as a resistor with resistance value of R_{on} or $R_{off,}$ at the HRS or LRS, respectively.

Figure 2.11 shows the relationship between the storage node resistance values and the read margin. For a given selector, in Figure 2.11a we see that lowering R_{on} does not necessarily lead to improved read margin, even though the R_{off}/R_{on} ratio is improved. This is due to the fact that reducing R_{on} in fact aggravates the voltage divider effect from the serially connected nonlinear select devices and the finite interconnect resistances. On the other hand, an opposite trend is found in Figure 2.11b which shows that increasing R_{off} leads to increased read margin at the same R_{off}/R_{on} ratio. This trend is further verified by examining the read margin as a function of storage node on/off ratio for different R_{on} values. As shown in Figure 2.11c, the read margin generally increases as the R_{off}/R_{on} ratio is

increased, but is always higher for higher R_{on} at the same R_{off}/R_{on} ratio due to the voltage divider effect.

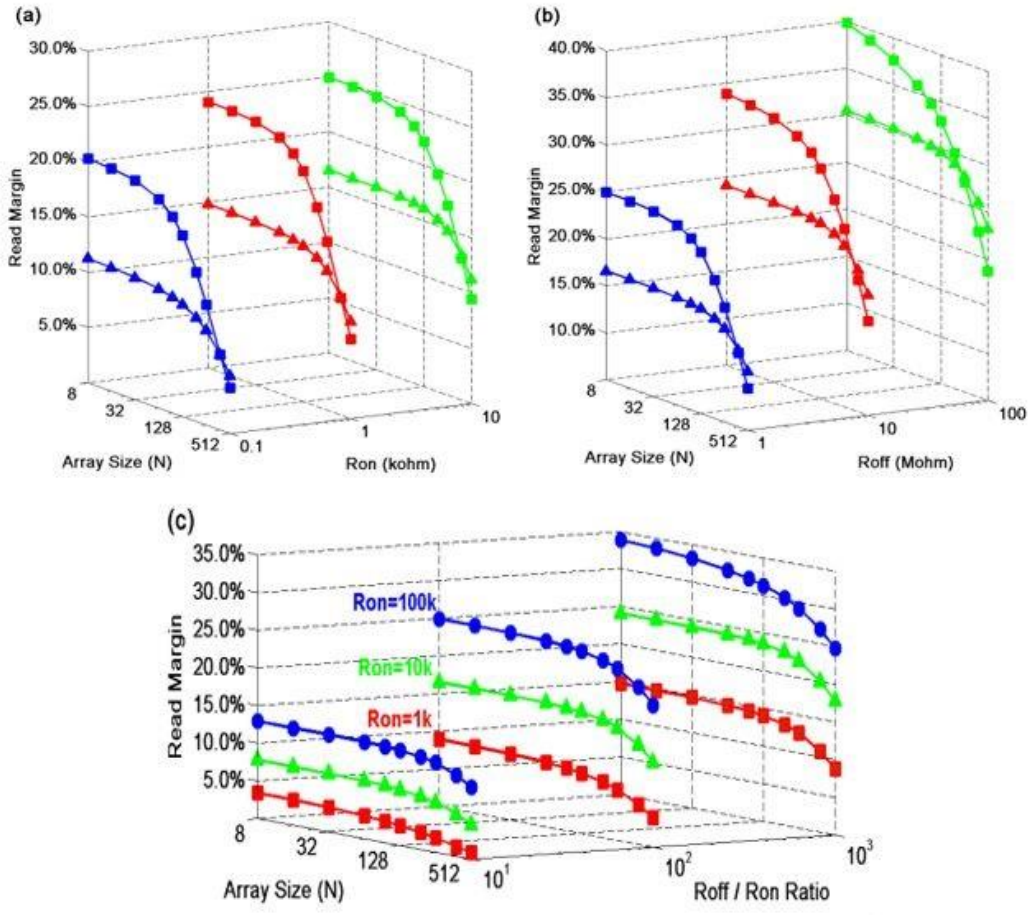


Figure 2.11. Relationship between the read margin and the storage node parameters. (a) with fixed R_{off} = 1M Ω; (b) with fixed R_{on}=10kΩ; (c) with varying R_{off}/R_{on} ratios. In both (a) and (b), upper curves marked by squares correspond to Selector1 ($k=10^4$, $I_{sel}(on)=100\mu A$); lower curves marked by triangles correspond to Selector2 ($k=10^4$, $I_{sel}(on)=1000\mu A$). Selector1 is simulated in (c). The sense resistor in (a)(b)(c) is set to $(R_{on}R_{off})^{1/2}$. Reproduced from [28].

These results suggest that if the device can only provide a limited R_{off}/R_{on} ratio, a desirable solution to mitigating the output signal degradation is to have high values of both

R_{on} and R_{off} instead of low values. However, in practical designs this requirement needs to be balanced with other demands such as speed requirements.

2.8. Conclusions

The read operations of an RRAM crossbar array have been systematically studied for a proposed 1S1R configuration consisting of memory cells with serially-connected selector device and storage node. The read margin was found to closely depend on the nonlinear conduction of the select device. Increasing the crossbar array size demands high values of not only selector on/off ratio (non-linearity) but also high selector on-current especially at low bias, and previously benchmarked selector on-current values seem too low for practical applications. The GN-FT and 1/3 V schemes were found to offer better read margin and lower operation power compared with the conventional GN-GN scheme. If highly nonlinear selectors can be obtained, the FT-FT and 1/2 V schemes are more desirable for ultra-low power application. The optimal sense resistor value was found to be still around the geometric mean of R_{on} and R_{off} but only affects the read margin slowly. The parameters of the storage node device should be optimized by considering the inherent voltage divider effect from the serially connected selector. These results provide a theoretical guidance for the design and optimization of RRAM devices and circuits.

2.9. References

[1] R. Waser and M. Aono, "Nanoionics-based resistive switching memories," *Nat. Mater.*, vol. 6, no. 11, pp. 833–840, 2007.

[2] J. J. Yang, D. B. Strukov, and D. R. Stewart, "Memristive devices for computing," *Nat. Nanotechnol.*, vol. 8, no. 1, pp. 13–24, 2012.

[3] DerChang Kau, S. Tang, I. V. Karpov, R. Dodge, B. Klehn, J. A. Kalb, J. Strand, A. Diaz, N. Leung, J. Wu, Sean Lee, T. Langtry, Kuo-wei Chang, C. Papagianni, Jinwook Lee, J. Hirst, S. Erra, E. Flores, N. Righos, H. Castro, and G. Spadini, "A stackable cross point Phase Change Memory," in *2009 IEEE International Electron Devices Meeting (IEDM)*, pp. 27.1.1–27.1.4, 2009.

[4] X. A. Tran, W. Zhu, W. J. Liu, Y. C. Yeo, B. Y. Nguyen, and H. Y. Yu, "A self-rectifying AlO_y bipolar RRAM with sub-50-μA set/reset current for cross-bar architecture," *IEEE Electron Device Lett.*, vol. 33, no. 10, pp. 1402–1404, 2012.

[5] E. Linn, R. Rosezin, C. Kügeler, and R. Waser, "Complementary resistive switches for passive nanocrossbar memories," *Nat. Mater.*, vol. 9, no. 5, pp. 403–406, 2010.

[6] M.-J. Lee, S. I. Kim, C. B. Lee, H. Yin, S.-E. Ahn, B. S. Kang, K. H. Kim, J. C. Park, C. J. Kim, I. Song, S. W. Kim, G. Stefanovich, J. H. Lee, S. J. Chung, Y. H. Kim, and Y. Park, "Low-Temperature-Grown Transition Metal Oxide Based Storage Materials and Oxide Transistors for High-Density Non-volatile Memory," *Adv. Funct. Mater.*, vol. 19, no. 10, pp. 1587–1593, 2009.

[7] K.-H. Kim, S. Hyun Jo, S. Gaba, and W. Lu, "Nanoscale resistive memory with intrinsic diode characteristics and long endurance," *Appl. Phys. Lett.*, vol. 96, no. 5, p. 053106, 2010.

[8] J. J. Huang, Y. M. Tseng, C. W. Hsu, and T. H. Hou, "Bipolar nonlinear Ni/TiO_2/Ni selector for 1S1R crossbar array applications," *IEEE Electron Device Lett.*, vol. 32, no. 10, pp. 1427–1429, 2011.

[9] A. Kawahara, R. Azuma, Y. Ikeda, K. Kawai, Y. Katoh, Y. Hayakawa, K. Tsuji, S. Yoneda, A. Himeno, K. Shimakawa, T. Takagi, T. Mikawa, and K. Aono, "An 8 Mb Multi-Layered Cross-Point ReRAM Macro With 443 MB/s Write Throughput," *IEEE J. Solid-State Circuits*, vol. 48, no. 1, pp. 178–185, 2013.

[10] W. Lee, J. Park, J. Shin, J. Woo, S. Kim, G. Choi, S. Jung, S. Park, D. Lee, E. Cha, H. D. Lee, S. G. Kim, S. Chung, and H. Hwang, "Varistor-type bidirectional switch (JMAX>10^7A/cm^2, selectivity~10^4) for 3D bipolar resistive memory arrays," in *2012 Symposium on VLSI Technology (VLSIT)*, vol. 1, no. 2010, pp. 37–38, 2012.

[11] M. J. Lee, Y. Park, D. S. Suh, E. H. Lee, S. Seo, D. C. Kim, R. Jung, B. S. Kang, S. E. Ahn, C. B. Lee, D. H. Seo, Y. K. Cha, I. K. Yoo, J. S. Kim, and B. H. Park, "Two series oxide resistors applicable to high speed and high density nonvolatile memory," *Adv. Mater.*, vol. 19, no. 22, pp. 3919–3923, 2007.

[12] X. Liu, S. M. Sadaf, M. Son, J. Shin, J. Park, J. Lee, S. Park, and H. Hwang, "Diode-less bilayer oxide (WO_x–NbO_x) device for cross-point resistive memory applications," *Nanotechnology*, vol. 22, no. 47, p. 475702, 2011.

[13] E. Cha, J. Woo, D. Lee, S. Lee, J. Song, Y. Koo, J. Lee, C. G. Park, M. Y. Yang, K. Kamiya, K. Shiraishi, B. Magyari-Kope, Y. Nishi, and H. Hwang, "Nanoscale (~10nm) 3D vertical ReRAM and NbO_2 threshold selector with TiN electrode," in *2013 IEEE International Electron Devices Meeting*, pp. 10.5.1–10.5.4, 2013.

[14] S. C. Puthentheradam, D. K. Schroder, and M. N. Kozicki, "Inherent diode isolation in programmable metallization cell resistive memory elements," *Appl. Phys. A*, vol. 102, no. 4, pp. 817–826, 2011.

[15] G.W. Burr, K. Virwani, R. S. Shenoy, A. Padilla, M. BrightSky, E. A. Joseph, M. Lofaro, A. J. Kellock, R. S. King, K. Nguyen, A. N. Bowers, M. Jurich, C. T. Rettner, B. Jackson, D. S. Bethune, R. M. Shelby, T. Topuria, N. Arellano, P. M. Rice, B. N. Ku, "Large-scale (512kbit) integration of Multilayer-ready Access-Devices based on Mixed-Ionic-Electronic-Conduction (MIEC) at 100% yield," in *2012 Symp. VLSI Technol. (VLSIT)*, pp. 41–42, 2012.

[16] V. S. S. Srinivasan, S. Chopra, P. Karkare, P. Bafna, S. Lashkare, P. Kumbhare, Y. Kim, S. Srinivasan, S. Kuppurao, S. Lodha, and U. Ganguly, "Punchthrough-Diode-Based Bipolar RRAM Selector by Si Epitaxy," *IEEE Electron Device Lett.*, vol. 33, no. 10, pp. 1396–1398, 2012.

[17] L. Zhang, A. Redolfi, C. Adelmann, S. Clima, I. P. Radu, Y.-Y. Chen, D. J. Wouters, G. Groeseneken, M. Jurczak, and B. Govoreanu, "Ultrathin Metal/Amorphous-Silicon/Metal Diode for Bipolar RRAM Selector Applications," *IEEE Electron Device Lett.*, vol. 35, no. 2, pp. 199–201, 2014.

[18] A. Flocke and T. G. Noll, "Fundamental analysis of resistive nano-crossbars for the use in hybrid Nano/CMOS-memory," in *ESSCIRC 2007 - 33rd European Solid-State Circuits Conference*, pp. 328–331, 2007.

[19] C. J. Amsinck, N. H. Di Spigna, D. P. Nackashi, and P. D. Franzon, "Scaling constraints in nanoelectronic random-access memories," *Nanotechnology*, vol. 16, no. 10, pp. 2251–2260, 2005.

[20] M. M. Ziegler and M. R. Stan, "CMOS/nano Co-design for crossbar-based molecular electronic systems," *IEEE Trans. Nanotechnol.*, vol. 2, no. 4, pp. 217–230, 2003.

[21] Cong Xu, Xiangyu Dong, N. P. Jouppi, and Yuan Xie, "Design implications of memristor-based RRAM cross-point structures," in *2011 Design, Automation & Test in Europe*, pp. 1–6, 2011.

[22] J. Liang and H.-S. P. Wong, "Cross-Point Memory Array Without Cell Selectors - Device Characteristics and Data Storage Pattern Dependencies," *IEEE Trans. Electron Devices*, vol. 57, no. 10, pp. 2531–2538, 2010.

[23] A. Chen, “A Comprehensive Crossbar Array Model With Solutions for Line Resistance and Nonlinear Device Characteristics,” *IEEE Trans. Electron Devices*, vol. 60, no. 4, pp. 1318–1326, 2013.

[24] 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, 2013.

[25] L. Zhang, S. Cosemans, D. J. Wouters, G. Groeseneken, M. Jurczak, and B. Govoreanu, “Selector design considerations and requirements for 1 SIR RRAM crossbar array,” in *2014 IEEE 6th International Memory Workshop (IMW)*, pp. 1–4, 2014.

[26] “International Technology Roadmap for Semiconductors (ITRS),” 2013.

[27] W. Kwon, J. Jeon, L. Hutin, and T.-J. K. Liu, “Electromechanical Diode Cell for Cross-Point Nonvolatile Memory Arrays,” *IEEE Electron Device Lett.*, vol. 33, no. 2, pp. 131–133, 2012.

[28] J. Zhou, K.-H. Kim, and W. Lu, “Crossbar RRAM Arrays: Selector Device Requirements During Read Operation,” *IEEE Trans. Electron Devices*, vol. 61, no. 5, pp. 1369–1376, 2014.

Chapter 3

Conduction Mechanism and Array Application of Tantalum Oxide Selector

3.1. Introduction

As mentioned in the previous chapter, even though resistive random access memory (RRAM) has numerous advantages, including simple structure, fast switching speed, low power consumption and excellent scalability[1]–[3], the sneak current issue needs to be solved first before RRAM devices can be implemented within a large-scale crossbar array for high-density memory applications [4], [5]. To suppress the sneak current, an additional non-linear selector element, such as diode [6], threshold switching device [7], [8], mixed ionic-electronic conductor device [9], tunneling device [4], [10], and punch-through diode [11] needs to be integrated with the memory cell. Among various selector devices, multilayer-oxide (TaO_x or TiO_x) based selectors have received wide attention owning to their high selectivity and high current density [4], [10], [12], [13]. By integrating the multilayer-oxide structure with a resistive switching layer without an intermediate electrode, “selector-less” or “self-rectifying” RRAM devices can be created [14], [15]. The integrated device can retain both the resistive switching behavior of the switching layer and the high nonlinear characteristics of the selector layer, and thus can be directly

implemented in cross-point memory arrays and mitigate the sneak current issue. Moreover, the selector-less RRAM device will simplify the fabrication process of array integration and is compatible with three-dimensional (3D) stacking[14]. However, a clear physical explanation of the conduction mechanism of multilayer-oxide based selectors is still lacking.

In this chapter, we first investigate the mechanism of a TaO_x-based selector device which exhibits high non-linearity (~10^4) and excellent uniformity. The selector consists of a Pd/TaO_x/Ta/Pd structure. Thermionic emission is observed when the device is positively biased (with respect to the bottom electrode) and tunnel emission is observed at reverse bias. The current-voltage characteristics at different temperatures were simulated by the Simmons trapezoidal energy barrier model and found to be consistent with the measured results. Based on the non-linear characteristic of the Pd/TaO_x/Ta/Pd device, we further demonstrate a selector-less RRAM device using a Pd/TaO_x/HfO_2/Pd structure, which exhibited high non-linearity (~5×10^3) in low resistance state (LRS) and reproducible resistive switching behaviors. To test the effectiveness of the integrated selector-less device, the read margin of the RRAM crossbar array was evaluated under different read schemes (V/2, V/3, Ground). Our calculations suggested a maximum array size up to 1M bits without significant degradation on the read margin, indicating the prospect of implementing this selector-less RRAM into high-density memory applications.

3.2. Device Fabrication and Measurement Setup

The Pd/TaO_x/Ta/Pd selector device studied in this chapter was fabricated following the steps below: A 50-nm-thickness Pd bottom electrode was deposited on a SiO2/Si

substrate by photo-lithography, e-beam evaporation and liftoff process, followed by the sputtering of a 12-nm thick Ta layer. The Ta layer was subsequently oxidized at 300 °C for 30 min under oxygen ambient in a furnace to form the TaO_x layer. Finally, a Pd top electrode was patterned and deposited using the same method as the bottom electrode, thus forming a crossbar structure with a device size of 2 μm × 2 μm. For the $Pd/TaO_x/HfO_2/Pd$ selector-less RRAM, HfO_2 layer was deposited by an Oxford atomic layer deposition (ALD) system with tetrakis-hafnium (TEMAH) precursor at 250 °C. The thickness of the TaO_x film was characterized by transmission electron spectroscopy (TEM). The electrical characteristics of selector and RRAM devices were measured by a Keithely 4200 semiconductor parameter analyzer. During the electrical testing, the bias voltage was always applied on the top electrode with the bottom electrode grounded. Temperature-dependent characteristics were measured in a Desert Cryogenics TTP4 probe station system in vacuum.

3.3. Conduction Mechanism Study of Tantalum-Oxide Selector Devices

Figure 3.1a and 3.1b show the SEM image and high resolution cross-sectional TEM image of a fabricated $Pd/TaO_x/Ta/Pd$ selector device, respectively. From the cross-sectional TEM image, it can be seen that there are two distinct layers between the top and the bottom Pd electrodes: an upper bright layer with a 7.8 nm thickness and a lower dark layer with an 8 nm thickness. The upper brighter layer was identified as oxidized TaO_x while the lower layer with a metallic contrast was identified as remaining un-oxidized Ta. To verify this assignment, the chemical compositions of the two layers were analyzed inside the TEM apparatus by energy-dispersive X-ray spectroscopy (EDS), as shown in

Figure 3.1c. An oxygen peak was clearly observed in the upper layer, while the lower 8 nm thick layer shows essentially no O composition, verifying its origin as metallic Ta. Since a total of 12 nm thick Ta film was deposited initially, the thickness of the upper layer, 7.8 nm, corresponds to a volume expansion factor of 1.95 if assuming 4 nm of Ta is converted into TaO_x via Ta oxidation. This value is lower than the volume expansion factor of 2.3 when Ta is fully oxidized into Ta_2O_5[16], and is consistent with a sub-oxide TaO_x formation during oxidation. As a result, a trapezoidal energy barrier is expected for conduction through the TaO_x layer, with different barrier heights of $q\phi_1$, $q\phi_2$ at the bottom and the top interfaces, respectively, due to the work function difference of the bottom Ta and top Pd electrodes, as shown in the inset.

The Pd/TaO_x/Ta/Pd selector device shows pronounced, reproducible non-linear *I-V* characteristics, as shown in Figure 3.1d, where results from 100 DC cycles are overlaid and plotted in semi-log scale. The non-linearity (*NL*), which is a crucial factor for evaluating selector, is defined as[17]:

$$NL = \frac{I_{V_D}}{I_{1/2 V_D}}$$

where V_D is the read voltage. In our device, the *NL* between 1.7 V and 0.85 V is over 10^4, suggesting this device can effectively suppress the leakage current at low bias. Significantly, measurements from 100 successive DC cycles show negligible variations during cycling, indicating excellent cycle-to-cycle uniformity.

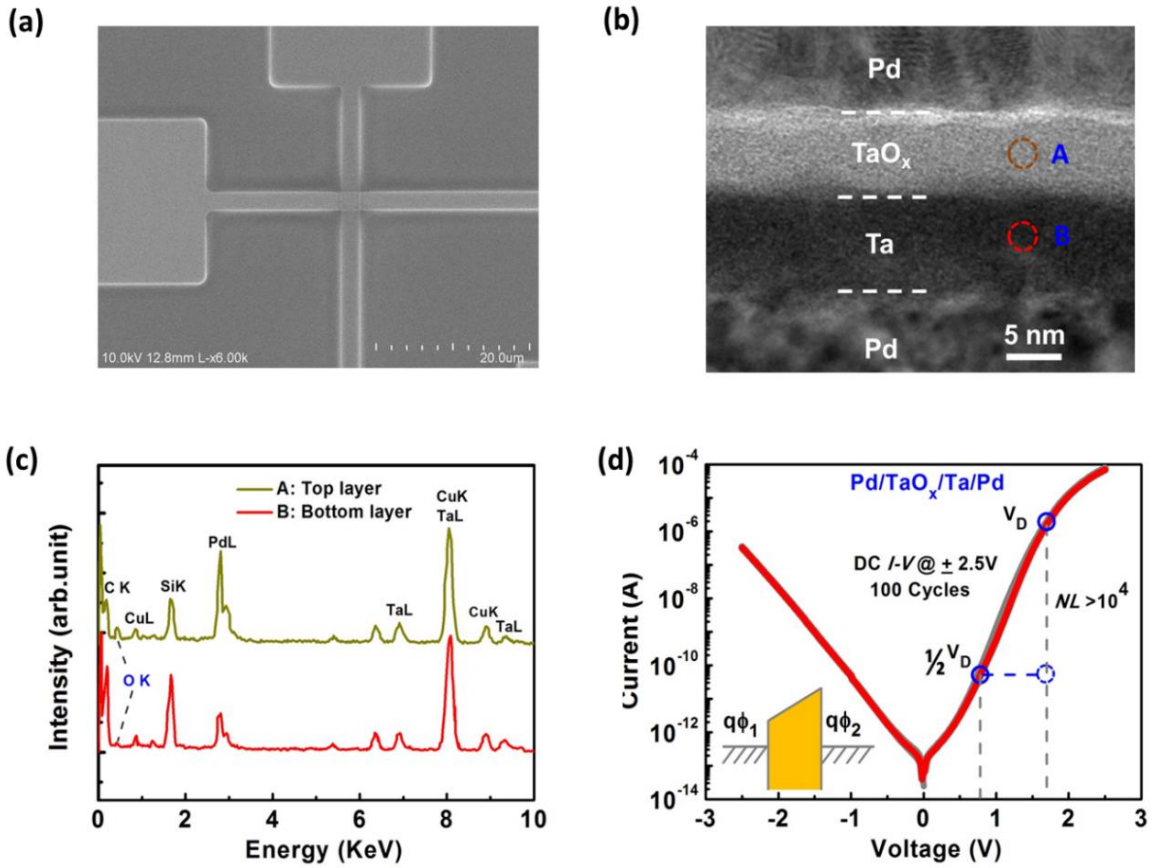


Figure 3.1. (a) SEM image of the as-fabricated Pd/TaO_x/Ta/Pd devices in a crossbar structure. Scale bar: 20 μm. (b) High resolution cross-sectional TEM image of a Pd/TaO_x/Ta/Pd device showing the different layers. The brown and red circles A and B mark the locations for the EDS analysis. (c) EDS spectrum of the TEM specimen tested in locations A and B. (d) Typical *I-V* characteristics of a Pd/TaO_x/Ta/Pd selector device. Results from 100 consecutive measurements are overlaid together. Very good reproducibility and a high *NL* (10^4 between 1.7 V and 0.85 V) can be obtained. The inset shows the schematic of a trapezoidal energy barrier for electron conduction through the TaO_x layer with energy barrier $q\phi_1$ and $q\phi_2$ at the bottom (TaO_x/Ta) and the top (Pd/TaO_x) interfaces, respectively. Reproduced from [26].

The device also exhibits a clear asymmetry with a higher current density and steeper slope at positive bias compared to those at negative bias (Fig 3.1d). To clarify the conduction mechanism that leads to the asymmetric and highly non-linear *I-V* characteristics of the Pd/TaO_x/Ta/Pd selector device, the device was measured in a wide

temperature range from 150 K to 300 K. As shown in Fig 3.2c, the measured current I_1 under positive bias is strongly temperature-dependent; while the measured current I_2 under negative bias shows little temperature dependence within the temperature range studied. The distinct temperature dependent behaviors between I_1 and I_2 suggest that different conduction mechanisms are dominating at positive and negative biases, and may explain the observed asymmetric and non-linear transport behaviors in the Pd/TaO_x/Ta/Pd selector device.

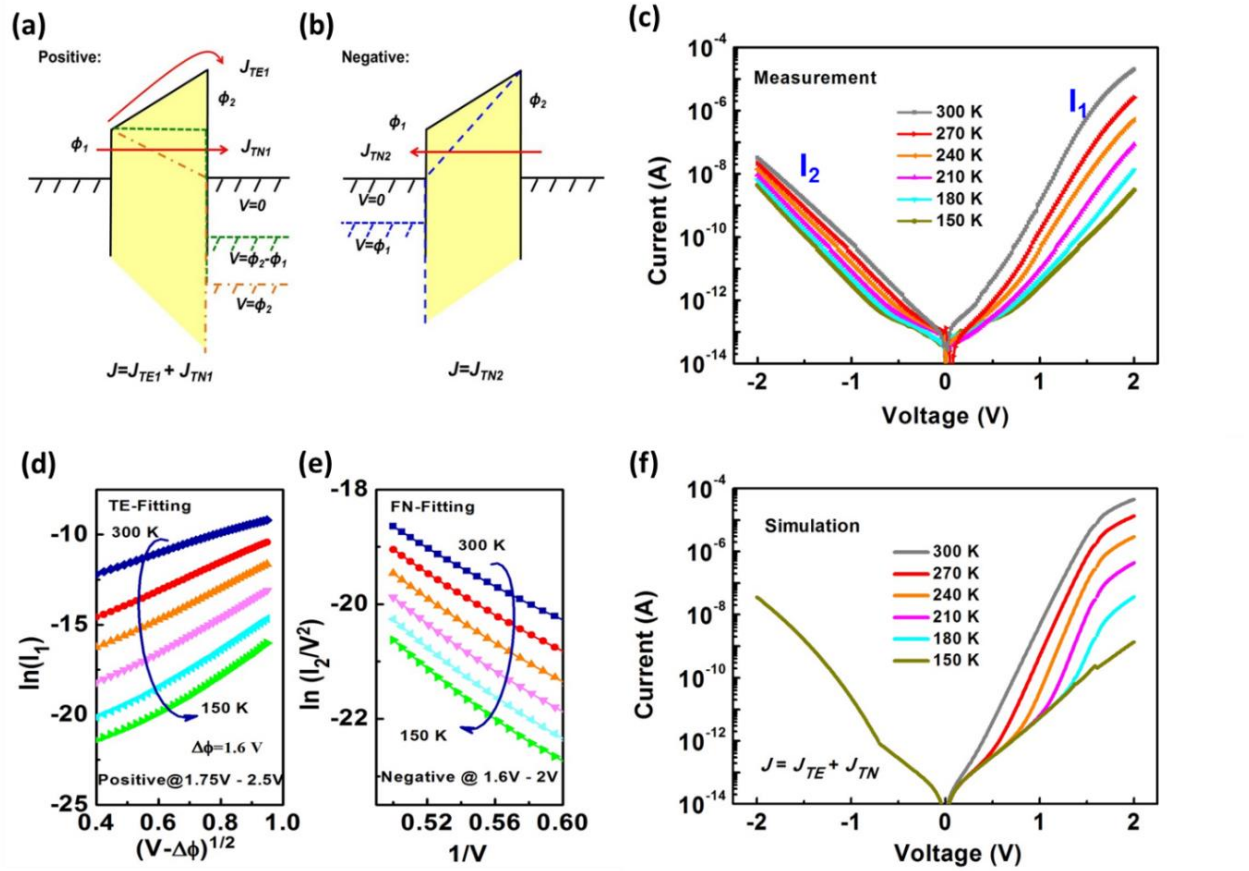


Figure 3.2. Band diagram of the Pd/TaO_x/Ta/Pd selector device under positive (a) and negative bias (b) conditions. (c) Temperatures dependence of device *I-V* (150 K to 300 K). (d) ln(I_1) vs. $(V\text{-}\Delta\phi)^{1/2}$ measured at different temperatures for the voltage range from 1.75 V to 2.5V (e) ln(I_2/V^2) vs. 1/*V* measured at different temperatures for the voltage range from -1.6 V to -2 V. (f) Simulated *I-V* curves at different temperatures. Reproduced from [26].

We note similar asymmetric *I-V* characteristics have been observed in thin insulating Langmuir films, and can be explained by thermionic and tunneling emission with a trapezoidal barrier model[18]. Following a similar approach, the band diagram of the Pd/TaO_x/Ta/Pd junction under positive and negative bias can be schematically illustrated in Fig 3.2a and 3.2b. Here, J_{TE1} (J_{TE2}), J_{TN1} (J_{TN2}), J_1 (J_2) are the net thermionic emission current density, tunneling current density and total current density of current flowing from the bottom (top) electrode to the top (bottom) electrode, respectively. Due to the lower barrier height of $q\phi_1$ at the bottom interface, in the positive bias region both thermionic emission and tunneling emission contribute to the electron transport process. In particular, thermionic emission dominates the device current at high temperature or large positive bias. In the negative region, thermionic emission can be ignored due to the high energy barrier $q\phi_2$, and the electron transport process is dominantly controlled by tunneling. Based on the Simmons' model[19], [20], the thermionic emission at positive bias can be expressed as:

$$J_{TE1} = A^{*}T^{2}\exp\left(\frac{q(V-\phi_2)}{\eta kT}\right)\exp\left(\frac{\left\{14.4\left[7+Ks(q\Delta\phi-qV)\right]\right\}^{\frac{1}{2}}}{KskT}\right)\ (0 \le V \le \Delta\phi = \phi_2 - \phi_1) \quad (1)$$

$$J_{TE1} = A^{*}T^{2}\exp\left(\frac{-q\phi_1}{\eta kT}\right)\exp\left(\frac{\left[14.4Ks(qV-q\Delta\phi)\right]^{\frac{1}{2}}}{KskT}\right)\ (V \ge \Delta\phi) \quad (2)$$

The tunneling emission at positive and negative bias can be calculated respectively as:

$$J_{TN1} = J_0\left\{\bar{q\phi_1}\exp(-A(\bar{q\phi_1})^{\frac{1}{2}}) - (\bar{q\phi_1}+qV)\exp\left[-A(\bar{q\phi_1}+qV)^{\frac{1}{2}}\right]\right\}$$

$$(0 < V \le \Delta\phi, \Delta\phi < V \le \phi_2, V > \phi_2) \quad (3)$$

$$J_{TN2} = J_0 \left\{ q\bar{\phi}_2 \exp(-A(q\bar{\phi}_2)^{\frac{1}{2}}) - (q\bar{\phi}_2 + qV)\exp\left[-A(q\bar{\phi}_2 + qV)^{\frac{1}{2}} \right] \right\} \quad (0 < V \le \phi_1, V > \phi_1) \quad (4)$$

Here, A* is the effective Richardson constant, q is the charge of electron, K is the dielectric constant of TaO_x, s is the thickness of TaO_x film, V is the applied voltage, η is the idea ideality factor[21]; $\bar{\phi}_1, \bar{\phi}_2$ are mean barrier heights as defined in [18].

Table 3.1: Simulation parameters for *I-V* curves of the Pd/TaO_x/Ta/Pd selector. Reproduced from [26].

Parameter	**Pd work function**[22] **(eV)**	**Ta work function**[22] **(eV)**	**TaO_x affinity**[23] **(eV)**	**$q\phi_1$ (eV)**	**$q\phi_2$ (eV)**	**$e\Delta\phi$ (eV)**	**TaO_x thickness (nm)**	**Area (μm^2)**	**η**
Value	5.6	4	3.3	0.7	2.3	1.6	7.8	2×2	2.4

Figure 3.2f shows a group of simulated *I-V* curves based on Eq. 1-4 at different temperatures. The Pd and Ta work function values (5.6 eV and 4 eV, respectively) and TaO_x electron affinity (3.3 eV) were obtained from literature [22], [23] $_1$ *and* $q\phi_2$ of 0.7 eV and 2.3 eV at the bottom and the top interface, respectively. The TaO_x film thickness and the device area were obtained from device geometry measurements. The only fitting parameter in the simulation is the idea ideality factor η, which was chosen to be 2.4 to yield the best simulation results. From Figure 3.2c and 3.2f, it can be seen that our model with only one fitting parameter agrees well both qualitatively and quantitatively with the experimental results in a broad temperature and bias range. Additionally, Eq. (2) predicts a linear dependence of ln(I_1) versus $(V\text{-}\Delta\phi)^{1/2}$ for V >> ϕ_1=0.7 V. Indeed, plotting the ln(I_1) - $(V\text{-}\Delta\phi)^{1/2}$ curves at different temperatures in the high voltage regime from 1.75 V to 2.5

V reveals a good linear dependence, as shown in Figure 3.2d, indicating the thermionic emission as the dominating mechanism at high positive bias. Additionally, Simmons' theory [33] tunneling emission will approach the general Fowler-Nordheim formula in the high bias region. Indeed, plotting $\ln(I/V^2)$ and $1/V$ at high negative bias (-1.6~-2V, Figure 3.2e) revealed a linear $\ln(I/V^2)$ - $1/V$ dependence at different temperatures, consistent with the hypothesis that tunneling emission is the dominant current mechanism at the negative bias region due to the asymmetric energy barrier.

The excellent agreement of the model with experimental results, along with the distinct temperature dependence of the device, reveals the physical mechanism behind the observed non-linear and asymmetric *I-V* characteristics of the TaO_x selector device. It also serves to provide a guidance for the continued design and optimization of selector devices in the future, where highly non-linear and asymmetric *I-V* behaviors are desirable [17], [24] .

3.4. Multilayer Oxide Based Selector-less RRAM Devices

The simple stack structure and fabrication process, along with the reliable non-linear *I-V* characteristics makes the TaO_x-based selector well suited for RRAM applications. To this end, we investigated the feasibility of integrating the TaO_x selective layer with an oxide-based switching layer to form selector-less RRAM devices. HfO_2, which exhibits reliable resistive switching behavior with a wide tunable resistance range [25], was chosen as the resistive switching layer to be integrated with the TaO_x selective layer.

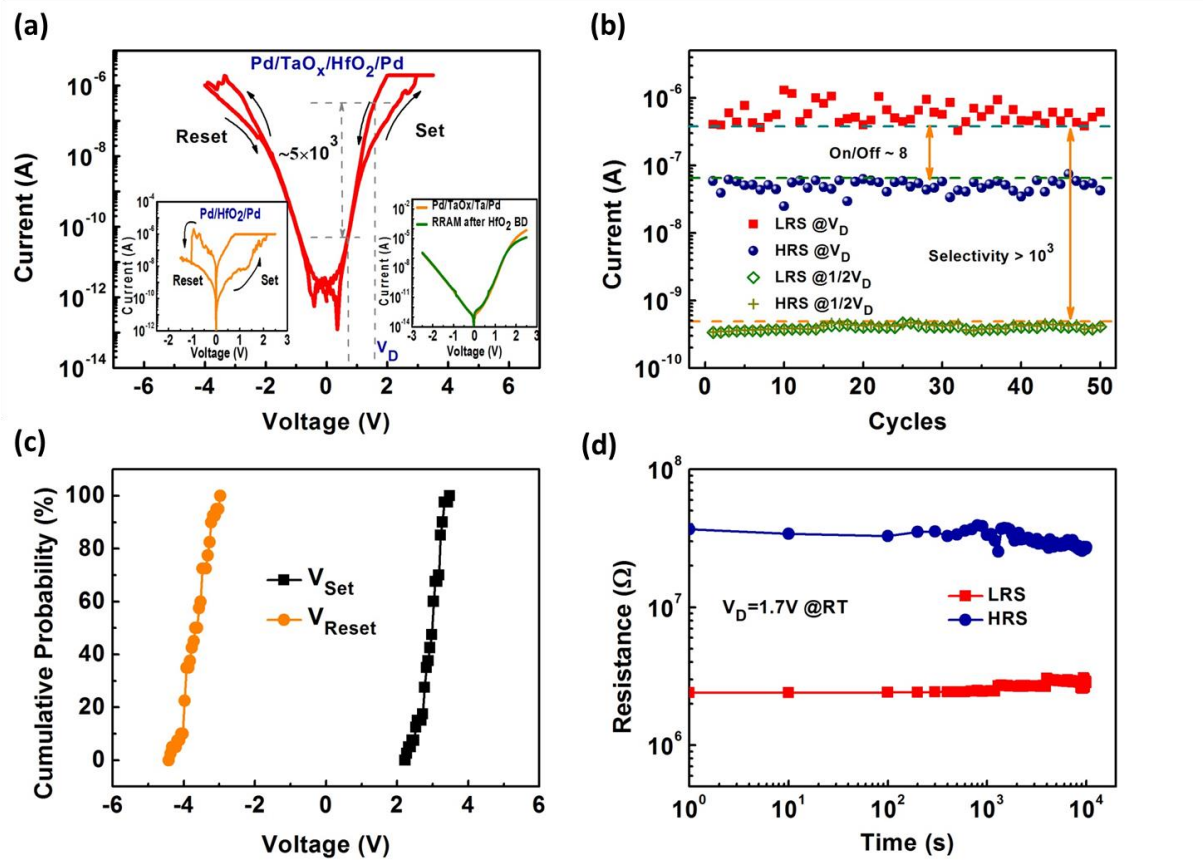


Figure 3.3. (a) Bipolar resistive switching of the selector-less Pd/TaO_x/HfO_2/Pd device. A current compliance of 2 μA was used during set. Left inset: *I-V* of a Pd/HfO_2/Pd control device. Right inset: *I-V* of the Pd/TaO_x/HfO_2/Pd device (green) after breakdown of the HfO_2 layer and *I-V* of the Pd/TaO_x/Ta/Pd selector device (yellow). (b) Endurance of the selector-less device measured at V_D=1.7 V. (c) Cumulative distributions of V_{Set} and V_{Reset}. (d) Room temperature (RT) retention of HRS and LRS. The HRS and LRS were tested every 100s with a read pulse (10 ms, 1.7 V). Reproduced from [26].

Figure 3.3a shows the bipolar resistive switching behaviors of an integrated Pd/TaO_x/HfO_2/Pd device. With increasing positive voltage (0→3.5 V) applied on the top electrode, an abrupt increase in current occurred at the set voltage (~ 2.9 V) and the device switched from the high resistance state (HRS) to the LRS. During negative voltage sweep (0→-4 V), a sudden drop in current appeared at the reset voltage (~ -3.3 V), when the device switched from the LRS to HRS. Significantly, the Pd/TaO_x/HfO_2/Pd device also

shows a pronounced non-linear behavior in the on-state (LRS state), with a 5×10^3-fold reduction in LRS current observed at $½V_D$ compared with current at V_D. For comparison, devices based on the Pd/HfO_2/Pd structure without the TaO_x layer shows well-defined resistive switching behaviors but only weak non-linearity in LRS (as shown in the left inset of Figure 3.3a). Additionally, after intentional breakdown (BD) of the HfO_2 resistive switching layer at high bias voltage, the *I-V* curve of the Pd/TaO_x/HfO_2/Pd device closely matches that of the Pd/TaO_x/Ta/Pd selector device (as shown in the right inset of Figure 3.3a), indicating that the device non-linearity is originated from the TaO_x selective layer. These results from the prototype Pd/TaO_x/HfO_2/Pd device suggest that the TaO_x-based selective element could be used to effectively mitigate the sneak leakage in crossbar RRAM arrays in a selector-less device structure.

Other important parameters for RRAM operations, including endurance, operation voltage and retention, were evaluated in the selector-less Pd/TaO_x/HfO_2/Pd device. Figure 3.3b shows results obtained from 50 successive DC cycles. The LRS current is reduced in the integrated Pd/TaO_x/HfO_2/Pd device compared with that in the standalone Pd/HfO_2/Pd device, but a minimum on/off ratio of 8 can still be obtained during cycling. Additionally, the current values of HRS and LRS at $½V_D$ are essentially the same and stay unchanged during cycling, due to the ability of the TaO_x selective layer to effectively suppress current at low bias, maintaining a high selectivity ($>10^3$) during cycling. The cumulative distributions of V_{Set} and V_{Reset} are illustrated in Figure 3.3c showing that tight distributions can still be obtained in the integrated device. Figure 3.3d shows the data retention characteristics of HRS and LRS at room temperature (RT). No significant degradation was observed for both resistance states.

3.5. Crossbar Array Simulation for TaO_x/HfO_2 RRAM Devices

To demonstrate the feasibility of the selector-less $Pd/TaO_x/HfO_2/Pd$ RRAM device in passive crossbar RRAM array applications, circuit level simulations were performed using a HSPICE model to estimate the read margin in different read schemes for different array sizes. In this model, the HfO_2 resistive switching layer is modeled to be in series with the TaO_x selective layer. Additionally, only the dynamic behaviors of the selector element are considered in the model, since the resistive switching element (the HfO_2 layer) is not considered to be disturbed during read. As a result, the resistance of the HfO_2 resistive switching layer was modeled as a resistor with either R_{on} ($4\times10^5\,\Omega$) or R_{off} ($1\times10^7\,\Omega$), and the *I-V* characteristics of the TaO_x selective layer was extracted from the numerical fitting of the $Pd/TaO_x/Ta/Pd$ selector device. The simulated *I-V* curve of the $Pd/TaO_x/HfO_2/Pd$ device is shown in Figure 3.4a, which agrees with the measured data reasonably well. Figure 3.4b shows the calculated read margin versus array size under two different read voltages (V_D=2.1 V and 1.8 V). The worse-case scenario was assumed in the simulation, with the target cell located at the farthest corner from the voltage source and all unselected cells in LRS [17]. Three different read schemes (V/2, V/3 and Ground, with the unselected cells connected to $V_D/2$, $V_D/3$ and ground, respectively[17]) were evaluated to reveal the optimal read operation approach. Here, the ground scheme refers to the scheme where the selected WL is biased to V_D and all other WL/BLs are grounded. The V/2 (V/3) scheme biases the unselected WLs to 1/2 (1/3) V_{read}, the unselected BLs at 1/2 (2/3) V_D, and makes the selected BL grounded.

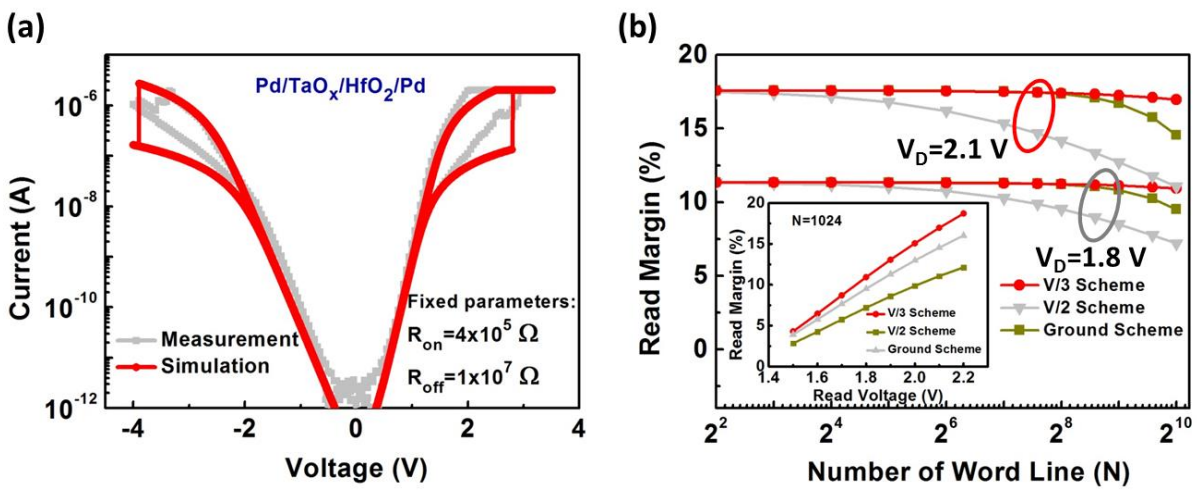


Figure 3.4. (a) Simulated *I-V* curve of the Pd/TaO_x/HfO_2/Pd device using fixed values of R_{on}=4×10^5 Ω and R_{off} =1×10^7 Ω. (b) Calculated read margin under different read voltages (V_D=2.1 V and 1.8 V) for three read schemes: V/2, V/3 and Ground. The read margin generally improves with the increase in read voltage, as shown in the inset. Reproduced from [26].

As shown in Figure 3.4b, the V/3 and Ground read schemes are more preferable to the proposed device structure, and read margin degradation is less than 4% under V/3 scheme for array size up to 1M bits. These results suggest that the selector-less Pd/TaO_x/HfO_2/Pd device can be implemented in high-density crossbar arrays and can effectively mitigate the sneak-current problem. In general, better read margin can be obtained at higher read voltage, as shown in the inset of Figure 3.4b. This effect can be explained by the non-linear behavior of the selector element such that the increased read voltages reduces the voltage divider effect from the selector layer and improves the read current [17].

3.6. Conclusions

In summary, we investigated the conduction mechanism of a Pd/TaO_x/Ta/Pd selector device which shows high non-linearity and asymmetric *I*-*V*. The trapezoidal energy barrier model qualitatively and quantitatively explains the observed electrical transport process at different temperatures, and reveals that thermionic emission is dominant at the positive bias and tunnel emission is dominant at the negative bias. A selector-less RRAM device based on the Pd/TaO_x/HfO_2/Pd structure was further demonstrated by integrating the TaO_x selective layer with the HfO_2 resistive switching layer. The integrated RRAM device shows a high selectivity (5×10^3) in LRS with the strong ability to significantly suppress sneak current at low bias. The read margin of crossbar memory array based on the selector-less RRAM cell was evaluated for different read schemes. The simulation results showed that the Pd/TaO_x/HfO_2/Pd device can be implemented in high-density arrays up to 1 Mbit with minimal degradation of the read margin. These results not only provide insight in understanding the origin of non-linearity in selector devices and selector-less RRAM devices, but also will guide the design and optimization of high-density crossbar RRAM applications.

3.7. References

[1] I. Valov, R. Waser, J. R. Jameson, and M. N. Kozicki, "Electrochemical metallization memories--fundamentals, applications, prospects.," *Nanotechnology*, vol. 22, no. 25, p. 254003, 2011.

[2] J. J. Yang, M. D. Pickett, X. Li, D. a a Ohlberg, D. R. Stewart, and R. S. Williams, "Memristive switching mechanism for metal/oxide/metal nanodevices," *Nat. Nanotechnol.*, vol. 3, no. 7, pp. 429–433, 2008.

[3] Y. Yang and W. Lu, "Nanoscale resistive switching devices: mechanisms and modeling," *Nanoscale*, vol. 5, no. 21, p. 10076, 2013.

[4] W. Lee, J. Park, S. Kim, J. Woo, J. Shin, G. Choi, S. Park, D. Lee, E. Cha, B. H. Lee, and H. Hwang, "High Current Density and Nonlinearity Combination of Selection Device Based on $TaO_x/TiO_2/TaO_x$ Structure for One Selector–One Resistor Arrays," *ACS Nano*, vol. 6, no. 9, pp. 8166–8172, Sep. 2012.

[5] 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, no. 1, pp. 389–395, Jan. 2012.

[6] M.-J. Lee, Y. Park, B.-S. Kang, S.-E. Ahn, C. Lee, K. Kim, W. Xianyu, G. Stefanovich, J.-H. Lee, S.-J. Chung, Y.-H. Kim, C.-S. Lee, J.-B. Park, I.-G. Baek, and I.-K. Yoo, "2-stack 1D-1R Cross-point Structure with Oxide Diodes as Switch Elements for High Density Resistance RAM Applications," in *2007 IEEE International Electron Devices Meeting*, pp. 771–774, 2007.

[7] M. J. Lee, Y. Park, D. S. Suh, E. H. Lee, S. Seo, D. C. Kim, R. Jung, B. S. Kang, S. E. Ahn, C. B. Lee, D. H. Seo, Y. K. Cha, I. K. Yoo, J. S. Kim, and B. H. Park, "Two series oxide resistors applicable to high speed and high density nonvolatile memory," *Adv. Mater.*, vol. 19, no. 22, pp. 3919–3923, 2007.

[8] X. Liu, S. M. Sadaf, M. Son, J. Shin, J. Park, J. Lee, S. Park, and H. Hwang, "Diode-less bilayer oxide (WO_x–NbO_x) device for cross-point resistive memory applications," *Nanotechnology*, vol. 22, no. 47, p. 475702, 2011.

[9] R. S. Shenoy, K. Gopalakrishnan, B. Jackson, K. Virwani, G. W. Burr, C. T. Rettner, a. Padilla, D. S. Bethune, R. M. Shelby, a. J. Kellock, M. Breitwisch, E. a. Joseph, R. Dasaka, R. S. King, K. Nguyen, a. N. Bowers, M. Jurich, a. M. Friz, T. Topuria, P. M. Rice, and B. N. Kurdi, "Endurance and scaling trends of novel access-devices for multi-layer crosspoint-memory based on mixed-ionic-electronic-conduction (MIEC) materials," *2011 Symp. VLSI Technol.*, vol. 4, pp. 94–95, 2011.

[10] J. Woo, W. Lee, S. Park, S. Kim, D. Lee, G. Choi, E. Cha, J. Lee, W. Jung, C. Park, and H. Hwang, "Multi-layer Tunnel Barrier ($Ta_2O_5/TaO_x/TiO_2$) Engineering for Bipolar RRAM Selector Applications," *2013 Symp. VLSI Circuits*, pp. 168–169, 2013.

[11] V. S. S. Srinivasan, S. Chopra, P. Karkare, P. Bafna, S. Lashkare, P. Kumbhare, Y. Kim, S. Srinivasan, S. Kuppurao, S. Lodha, and U. Ganguly, "Punchthrough-Diode-

Based Bipolar RRAM Selector by Si Epitaxy," *IEEE Electron Device Lett.*, vol. 33, no. 10, pp. 1396–1398, 2012.

[12] J. Shin, I. Kim, K. P. Biju, M. Jo, J. Park, J. Lee, S. Jung, W. Lee, S. Kim, S. Park, and H. Hwang, "TiO_2-based metal-insulator-metal selection device for bipolar resistive random access memory cross-point application," *J. Appl. Phys.*, vol. 109, no. 3, 2011.

[13] J. J. Huang, C. W. Kuo, W. C. Chang, and T. H. Hou, "Transition of stable rectification to resistive-switching in Ti/TiO_2/Pt oxide diode," *Appl. Phys. Lett.*, vol. 96, no. 26, pp. 2008–2011, 2010.

[14] C.-W. Hsu, Y.-F. Wang, C.-C. Wan, I.-T. Wang, C.-T. Chou, W.-L. Lai, Y.-J. Lee, and T.-H. Hou, "Homogeneous barrier modulation of TaO_x/TiO_2 bilayers for ultra-high endurance three-dimensional storage-class memory," *Nanotechnology*, vol. 25, no. 16, p. 165202, 2014.

[15] J. Woo, S. Kim, W. Lee, D. Lee, S. Park, G. Choi, E. Cha, and H. Hwang, "Thermally activated non-linearity of device in resistance-switching memory for cross-point array applications," *Appl. Phys. Lett.*, vol. 102, no. 12, p. 122115, 2013.

[16] S. W. Park, "Effects of oxidation conditions on the properties of tantalum oxide films on silicon substrates," vol. 207, pp. 258–264, 1992.

[17] J. Zhou, K.-H. Kim, and W. Lu, "Crossbar RRAM Arrays: Selector Device Requirements During Read Operation," *IEEE Trans. Electron Devices*, vol. 61, no. 5, pp. 1369–1376, 2014.

[18] S. Horiuchi and K. Naito, "Electric Conduction through Thin Insulating Langmuir Film," *J. Electrochem. Soc.*, vol. 115, no. 6, pp. 634–637, 1968.

[19] J. G. Simmons, "Generalized Formula for the Electric Tunnel Effect between Similar Electrodes Separated by a Thin Insulating Film," *J. Appl. Phys.*, vol. 34, no. 6, p. 1793, 1963.

[20] J. G. Simmons, "Potential Barriers and Emission-Limited Current Flow Between Closely Spaced Parallel Metal Electrodes," *J. Appl. Phys.*, vol. 35, no. 8, p. 2472, 1964.

[21] S. K. Cheung and N. W. Cheung, "Extraction of Schottky diode parameters from forward current-voltage characteristics," *Appl. Phys. Lett.*, vol. 49, no. 2, pp. 85–87, 1986.

[22] H. B. Michaelson, "The work function of the elements and its periodicity," *J. Appl. Phys.*, vol. 48, no. 11, p. 4729, 1977.

[23] J. Robertson and C. W. Chen, "Schottky barrier heights of tantalum oxide, barium strontium titanate, lead titanate, and strontium bismuth tantalate," *Appl. Phys. Lett.*, vol. 74, no. 8, p. 1168, 1999.

[24] S. Kim, J. Zhou, and W. D. Lu, “Crossbar RRAM Arrays: Selector Device Requirements During Write Operation,” *IEEE Trans. Electron Devices*, vol. 61, no. 8, pp. 2820–2826, 2014.

[25] B. Govoreanu, G. S. Kar, Y.-Y. Chen, V. Paraschiv, S. Kubicek, A. Fantini, I. P. Radu, L. Goux, S. Clima, R. Degraeve, N. Jossart, O. Richard, T. Vandeweyer, K. Seo, P. Hendrickx, G. Pourtois, H. Bender, L. Altimime, D. J. Wouters, J. A. Kittl, and M. Jurczak, “10×10nm2 Hf/HfO_x crossbar resistive RAM with excellent performance, reliability and low-energy operation,” in *2011 International Electron Devices Meeting*, pp. 31.6.1–31.6.4, 2011.

[26] M. Wang, J. Zhou, Y. Yang, S. Gaba, M. Liu and W. D. Lu, “Conduction mechanism of a TaO x-based selector and its application in crossbar memory arrays,” *Nanoscale*, vol. 7, no. 11, pp. 4964–4970, 2015.

Chapter 4

Low Programming-Current Resistive Switching Memory

4.1. Introduction and Motivation

To approach commercial nonvolatile memory applications, RRAM devices with low operating current are highly preferable. Scaling the RRAM programming current will reduce the total power consumption and relax the requirement on current-driving capacity of the select transistor (1T1R configuration) or the selector device (in 1S1R configuration). In the meantime, with lower operation current, less voltage will drop on the word/bit lines in a crossbar array, thus mitigating the parasitic series-resistance effect from the electrodes and improving the read margin as well as the write disturbance[1]–[3].

Recently low-power RRAMs with sub-μA operating current have been demonstrated in different material systems [4]–[9]. Typically the low programming current leads to the formation of very weak conductive filaments in the switching layer [10]–[12]. Given the limited diameter, however, these filaments may be more susceptible to noise and retention degradation [13]–[15].

In this chapter, we first demonstrate reliable sub-nA operations using a copper based RRAM structure with a polycrystalline silicon in-cell resistor. To further support the high

density integration and alleviate the sneak leakage issue in crossbar RRAM arrays, a novel $Cu/Al_2O_3/aSi/Ta$ RRAM structure with self-rectifying characteristics is further developed. The $Cu/Al_2O_3/aSi/Ta$ RRAM cell exhibits low operating current (~nA), high on/off ratios (>100x) and pronounced nonlinearity. The use of low programming-current RRAM elements avoids the current-driving capability bottleneck of selectors, while the integrated rectifying layer improves the RRAM operation reliability.

4.2. Copper Based RRAM with Polycrystalline Silicon In-cell Resistor

4.2.1. Device Structure and Fabrication

Experiments were conducted on two-terminal cross-point devices with lateral size of 2μm × 2μm. 50nm boron-doped polysilicon (1000 ohm/square, ~$1e^{19}$ cm^{-3} doping level) was deposited using low pressure chemical vapor deposition (LPCVD) at 580°C on a Si/SiO_2 wafer. The polysilicon bottom electrodes (BEs) were patterned by photolithography and reactive ion etching (RIE) process. Immediately after a short 1:20 HF dip to remove the native oxide, 10nm Al_2O_3 was deposited using trimethylaluminum (TMA) and water precursors in an Oxford Opal atomic layer deposition system at 150 °C. Copper top electrodes (TEs) were then patterned using photolithography and liftoff processes. A stack of 800Å Cu/ 400Å Au was used to passivate the Cu electrode. All measurements were conducted using a Keithley 4200 semiconductor characterization system (SCS). For all electrical testing, the external bias was applied to the Cu TE while the polysilicon BE was grounded.

4.2.2. Results and Discussions

As shown in Figure 4.1, the device can be programmed and erased with sub-nA current compliance (0.5nA) provided by the 4200-SCS. The sharp increase (decrease) in the current corresponds to the formation (rupture) of the filament during set (reset) process. Particularly, the polysilicon BE serves as an in-cell resistor to effectively limit potential voltage overshoot during filament growth to prevent the formation of thick filaments (Figure 4.2). The effect of the polysilicon BE in-cell resistor was confirmed by fabricating devices with metal (e.g. W) BEs. All devices with W BEs shorted out after the forming operation even when very low current compliance (0.5nA) was used. This is consistent with previous reports suggesting reduction in voltage/current overshoot during forming / SET process is required to prevent over-programming of the device [16].

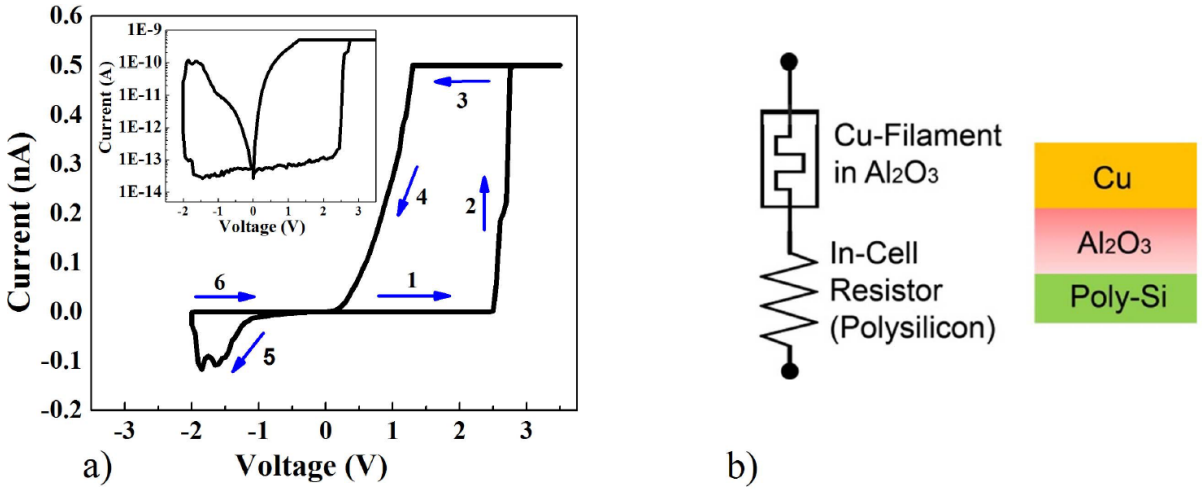


Figure 4.1. (a) Linear and log (inset) scale I–V curve showing sub-1nA current operation; (b) Device structure schematic: The polysilicon BE effectively acts as an in-cell resistor and prevents overshoot during writing. Reproduced from [21].

With the low current compliance and the in-cell resistor to prevent voltage overshoot, we expect improved control of the filament growth process. Specifically, instead of the formation of a very weak filament in previously studied low-current RRAM devices, we target an incomplete filament which has a solid base but leaves a gap between the filament tip and the BE, schematically shown in the inset of Figure 4.2. In this case, low current is obtained since the filament does not completely bridge the two electrodes; while good retention can still be maintained as the filament does not have a very weak tip that leads to retention loss. The concept of controlling filament growth has been recently demonstrated by us and verified through in-situ TEM studies [10], [11].

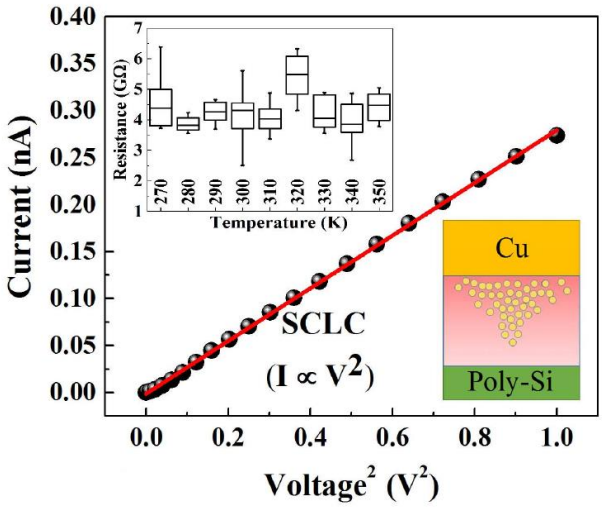


Figure 4.2. Linear fitting for LRS current vs. V^2 indicating SCLC as the conduction mechanism. Since the copper filament does not completely bridge the two electrodes (lower inset), very low operating current can be obtained. Upper inset shows that the LRS resistance is apparently independent of temperature, consistent with the SCLC conduction mechanism. Reproduced from [21].

The hypothesis of the incomplete filament formation was supported by analyzing the measured I–V curve in the low-resistance state (LRS). The LRS I–V can be well fitted with

a SCLC (space-charge-limited-conduction) model (Figure 4.2), consistent with electron transport through a thin ALD Al_2O_3 layer [17]. The space charge limited current through a thin material of thickness d and dielectric constant ε is given by

$$J = \frac{9\varepsilon\mu V^2}{8d^3},$$

where μ is the electron mobility in the specific medium, and V is the applied voltage. By assuming a gap d between the filament tip and the BE to be ~2–3nm [18], μ (electron mobility in Al_2O_3) = $7e^{-9}$ m^2/V-s [17], ε_r = 4 [17], the effective electrode area responsible for LRS conduction can be calculated to be 8–26nm^2, suggesting the presence of a dominant filament with an effective tip area of 8–26nm^2. The estimated filament and gap size are consistent with the observed filament shape/characteristics from experiments targeted at visualizing the actual filament [10] and support the concept of having a partially formed filament in the LRS to maintain low programming current and retention.

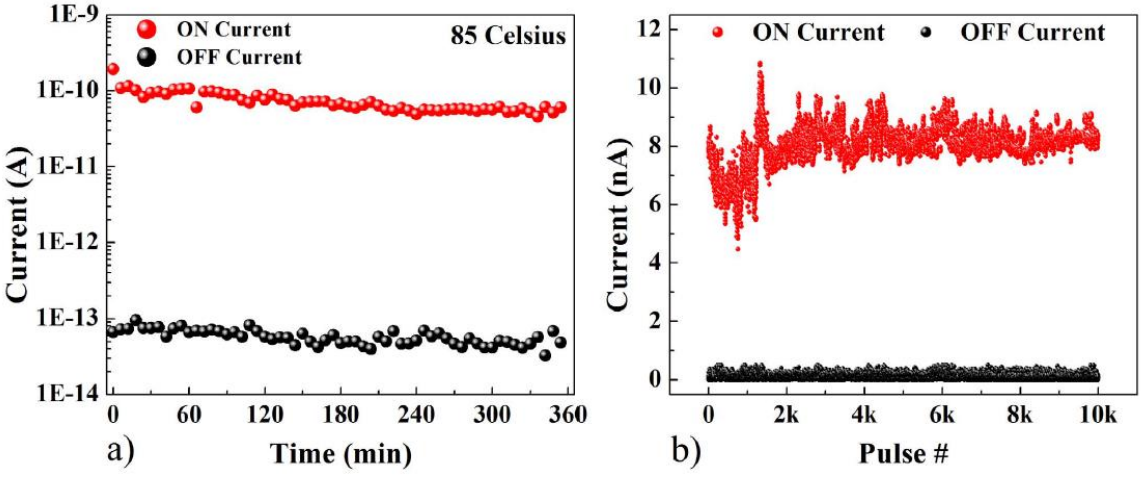


Figure 4.3. (a) Elevated temperature retention test. Read pulse (1V/10ms) was repeated every 6 minutes. Large read window is maintained after 6 hours at 85 °C. (b) 10000 cycle pulse data indicating robust endurance. Write pulse: 5V/5ms, erase pulse: −2.5V/4ms, read pulse: 1V, 10ms. Reproduced from [21].

Despite the ultra-low programming current, the devices show good retention behavior (Figure 4.3a), since the current is not limited by the narrow filament width but rather by the gap between the filament tip and the bottom electrode. An incomplete but robust filament results in stable read current even at 85 °C. Excellent device endurance of 10000 pulse cycles can be also obtained with the help of a 1Gohm series resistance (Figure 4.3b).

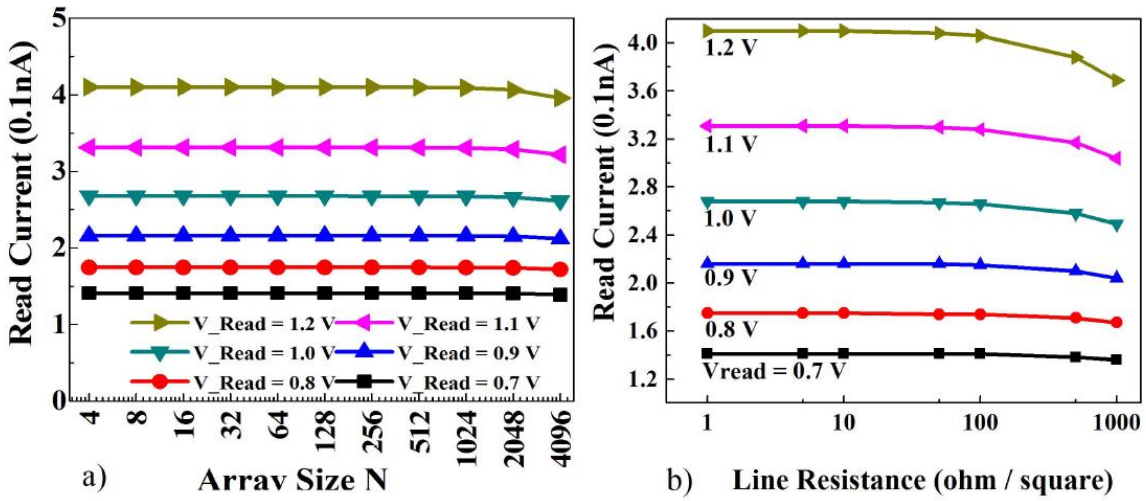


Figure 4.4. (a) Read margin for square arrays with N rows and N columns. Line resistance of 100 ohm/sq. is assumed. Grounding scheme is utilized for array simulation - unselected word/bit-lines are held at ground potential. The selected word-line is biased at V_{read} while the selected bit-line is grounded. Worst case scenario is assumed with the target cell located at the farthest corner and all unselected cells in low resistance state. (b) Read current as a function of word / bit line resistance for N = 512. Reproduced from [21].

Based on a grounding scheme, HSPICE numerical simulations were conducted to verify the read operation of crossbar array integrating this low-current RRAM cell. As shown in Figure 4.4a, the read current loss is negligible in a 1M-bit array and is less than 10% within a 16M-bit array. Besides, the low current negates the detrimental effects of line resistance of the word/bit lines, particularly for the grounding scheme during read. Normally, increased line resistance causes increased voltage drops on the selected word/bit

line, and also raises the potential on unselected word lines which can lead to inaccurate read current even for the grounding configuration. Due to the very low current, the voltage drops on the electrodes are minimized and no significant read current degradation is observed for large arrays (Figure 4.4b). Additionally, the total power dissipation can be reduced. For example, the worst-case total power dissipation during read using the grounding scheme for a 1 Mb array was calculated to be $2.7e^{-7}$W at 1V read voltage, lower than the power dissipation for reading a single cell with a read current of 1μA.

4.3. Copper Based RRAM with Self-Rectifying Characteristics

4.3.1. Device Concept

In the last section, a sub-nA copper based RRAM with polycrystalline silicon in-cell resistor has been demonstrated. Reliable retention and excellent endurance are proven on the individual cell level. However, several inherent drawbacks hinder the practical application of this cell structure: 1) the LPCVD step for polysilicon bottom electrode, which requires high temperature, cannot adapt to the mainstream CMOS technology and the back-end-of-line (BEOL) integration, thus eliminating RRAM's advantage of 3D stacking; 2) compared to metal electrodes (e.g. Cu), the polysilicon electrodes is much more resistive and will result in severe RC delays on read/write operations; 3) it is challenging to precisely control the growth of incomplete conductive filament in the switching layer; 4) the LRS fail to provide sufficient nonlinearity and will suffer from the sneak leakage issue, in particular when the crossbar array is very large.

A straightforward solution to (1) and (2) is replacing polysilicon with metal materials that are CMOS compatible. However, as pointed out in [16], changing bottom electrodes alone may sacrifice the low operating current and lead to reliability issues from high transient current during forming/set processes. As a RRAM device switches to the LRS state, the accompanying current spike from the discharging of the parasitic capacitance tends to over-program the memory and cause the over-growth of the filament and a very low resistance state. From a microscopic perspective, the variability of incomplete conductive filaments in (3) is also highly associated with this discharging current, since the migration of Cu ions is driven by the device current which increases significantly as the gap distance reduces. Intuitively, a robust barrier layer inserted between the switching material and the metal electrode can help resolve the problems by reliably regulating the transient current, effectively confining the conductive filament within the switching layer and clearly defining the gap between filaments and bottom electrodes.

It's worth noting that the inserted barrier layer may also potentially act as a rectifying/selector element and provide nonlinear I-V characteristics for the RRAM device in the meantime. A major bottleneck for the selector development is the high current-density requirement to minimize the voltage divider effect with the serially connected memory element, and allow sufficient read/write margin for the array operation, as discussed in Chapter 2 [1], [2]. The mismatch between the RRAM on-resistance and the selector on-resistance lowers the on/off memory window and reduces the effectiveness of the non-linear selector in 1S1R memories[3]. In principle, by integrating selectors with low-current RRAMs, the requirement for high selector current may be alleviated while the advantages of high nonlinearity and low power consumption can be maintained.

4.3.2. Amorphous Silicon Based Selector Devices

Metal/amorphous-silicon/metal (MSM) structures have been reported and demonstrated for selector applications in the 1S1R configuration[19], [20] . Using an ultrathin amorphous silicon layer acting as the tunneling dielectric, the MSM sandwiched structure has been shown to provide high driving current (exceeding 1MA/cm^2), large nonlinearity and fast switching speed (~ns). With a moderate thermal budget (less than 600 °C), the device performance and reliability can be further improved by thermal annealing treatment and barrier engineering. The relatively simple process is beneficial for the high density integration of this MSM selector.

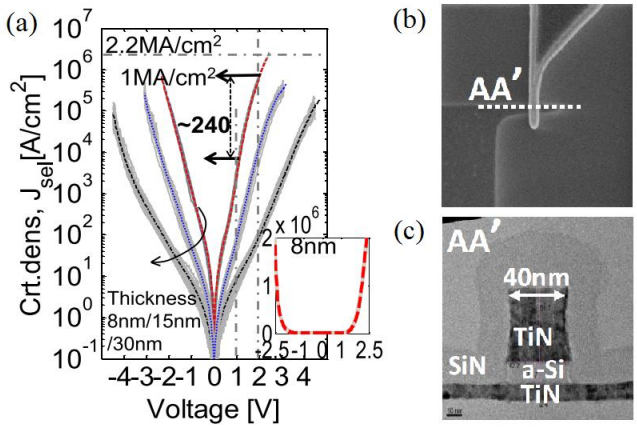


Figure 4.5. (a) I-V plots for the MSM selectors (40nm×40nm) with different silicon thicknesses. (b)(c) Cross-sectional TEM images for MSM selector. Reproduced from [19], [20].

Inspired by the reported selector studies, we expect to implement the amorphous silicon layer into our low-current RRAM as a barrier layer for Cu migration and as a rectifying layer to provide nonlinear IV characteristics. Based on these device conceptions,

a novel Cu/Al_2O_3/aSi/Ta RRAM structure (Figure 4.7) is proposed based on the previous Cu/Al_2O_3/polysilicon system.

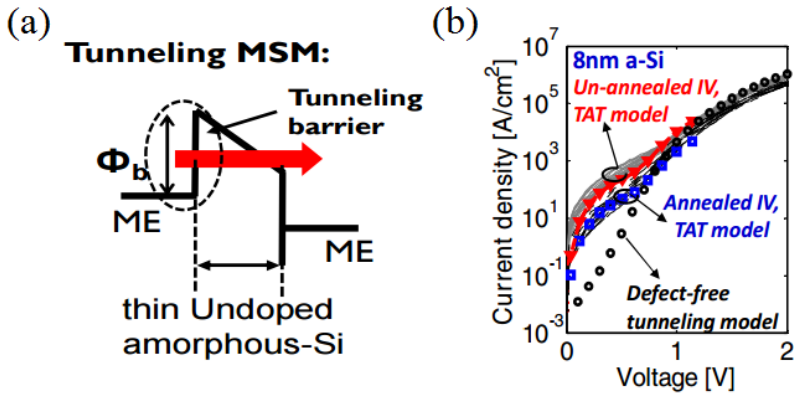


Figure 4.6. (a) Band diagram of the MSM selector (under bias), with ultrathin undoped amorphous silicon, which behaves as a low-bandgap tunnel dielectric. (b) Trap assisted tunneling (TAT) and defect-free tunneling conduction models explain the I-V difference between annealed and un-annealed samples. Reproduced from [19], [20].

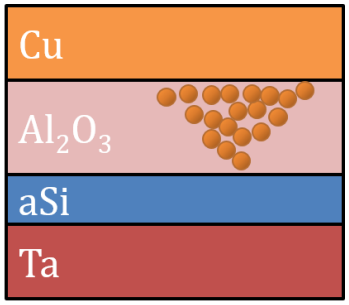


Figure 4.7. Conceptual device schematic of a Cu/Al_2O_3/aSi/Ta RRAM cell at LRS. Cu conductive filaments are expected to form in the Al_2O_3 and stop growth at the aSi/Al_2O_3 interface.

4.3.3. Device Structure and Fabrication

The Cu/Al_2O_3/aSi/Ta cells were fabricated in a crossbar configuration. Starting from a Si/SiO_2 substrate, the bottom electrodes (BE) were defined by photolithography and

50nm sputtered tantalum (Ta). An aSi film was then deposited with a thickness in a range from 4.5nm to 9nm by room temperature DC sputtering using a p-type silicon target (with a boron doping level of ~10^{19} cm^{-3}). Immediately after the aSi sputtering, 6nm Al_2O_3 was grown by atomic layer deposition (ALD) at 150 °C, the peak temperature of the entire process. The Cu top electrodes (TE) were then patterned by photolithography and liftoff, and followed by an Au passivation layer deposition. The device size was 2μm×2μm. The device SEM image is shown in Figure 4.8. Besides, control samples with MSM stacks similar to [19] but different metal electrodes were also fabricated.

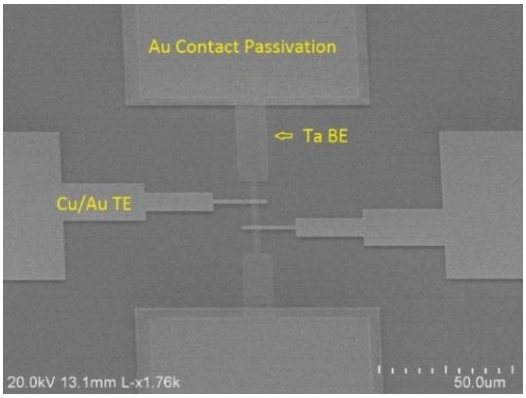


Figure 4.8. Scanning electron microscope image of two Cu/Al_2O_3/aSi/Ta RRAM devices sharing a Ta bottom electrode.

4.3.4. Low-current Operation and Self-Rectifying Characteristics

A typical I-V curve of a Cu/Al_2O_3/aSi/Ta device with a 6nm thick aSi layer is shown in Figure 4.9. The voltage bias is always applied on the TE; while the BE is grounded. After an initial DC forming process with a compliance current (I_{cc}) of 10nA at 5V, the device shows repeatable bipolar switching characteristics with very low operating current (~nA) even without current compliance. It can be set from the high resistance state (HRS)

to the low resistance state (LRS) at ~3.6V and reset back to HRS at ~-2V. The sharp increase/decrease in the device current can be understood by the formation/rupture of Cu filament in the Al_2O_3 switching layer, respectively[21]. The inset of Figure 4.9 shows the I-V curve plotted in linear scale and resistance variations in cycling tests, highlighting the LRS nonlinearity and the on/off ratio of the cell. The half-bias nonlinearity $NL_{1/2}$ = $I(V_{read})/I(1/2V_{read})$ is ~13, and the 1/3-bias nonlinearity $NL_{1/3}$ is ~34 when the device is read at 2V. The on/off window is ~130x at 2V with a low OFF current of ~1pA. Additionally, the device can be erased with a reset current as low as 30pA.

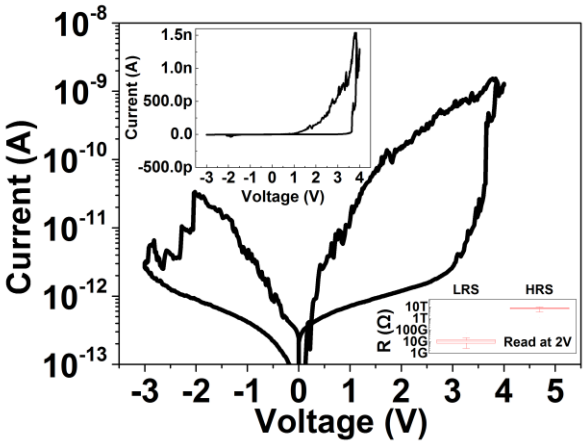


Figure 4.9. Typical *I-V* curve of a $Cu/Al_2O_3/aSi/Ta$ device. Upper inset: data plotted in linear scale. Lower inset: resistance variations of HRS/LRS over 50 DC cycles. Reproduced from [26].

The low operating current demonstrated in this paper can be well explained by the controlled growth of filaments [10] where a small gap between the filament and the bottom electrode limits the LRS current. However, as discussed in Section 4.3.1, the polysilicon bottom electrode that helps control the filament growth in [21] will introduce numerous problems in large-scale crossbar applications, such as parasitic resistance and incompatible

processes. Here by inserting a thin amorphous silicon layer as a barrier layer that stops filament growth, the gap in the conduction path can now be defined definitely while the series resistance of electrodes is minimized by replacing polysilicon with a metal (e.g. Ta). Additionally, the CBRAM cell structure proposed here allows low current and LRS nonlinearity without an extra middle electrode. This will help alleviate the sneak leakage in crossbar array and simply the integration steps.

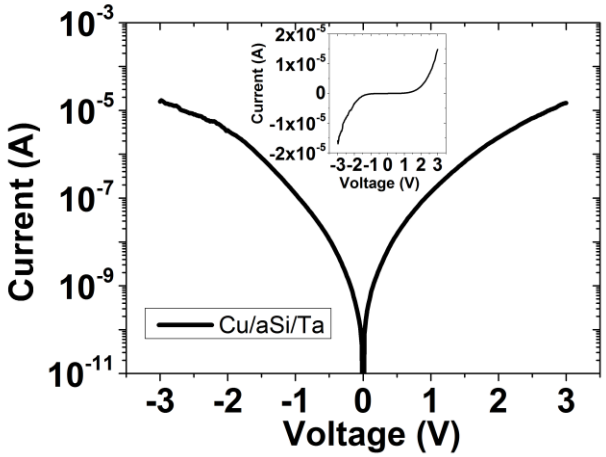


Figure 4.10. Typical I-V curve of a Cu/aSi/Ta device in log and linear scale (inset). Reproduced from [26].

To further validate the role of aSi barrier layer, control samples without the Al_2O_3 switching layer were also measured. Figure 4.10 shows the I-V curve of a Cu/aSi(6nm)/Ta control device. It exhibits the expected nonlinear conduction ($NL_{1/2}$~17 at 2V) very similar to that observed in the LRS of the Cu/Al_2O_3/aSi/Ta device. It is also worth noting that the aSi barrier/rectifying layer does not break down at 4V and at current levels more than 10000 times higher than the programming current of the memory device. These observations suggest that the formation/rupture of the filament is confined within the Al_2O_3

layer, and the aSi layer provides the nonlinear conduction effect of the LRS. Figure 4.11 shows that the nonlinearity and current level strongly depend on the electrode/aSi interface. For example, titanium (Ti) electrode yields an improved $NL_{1/2}$ of 65 and can drive 70 times higher current at 2V compared to the Cu/aSi/Ta device. Therefore, further optimizations on the NL and the current density (if needed) are feasible by tuning the rectifying layer material thickness, the effective Schottky barrier height at the electrode/aSi interface through the choice of the BE metal, and the electrode/aSi interface, following an approach discussed in [20]. The choice of the rectifying layer also depends on the breakdown electric field and its ability to suppress ion migration [11].

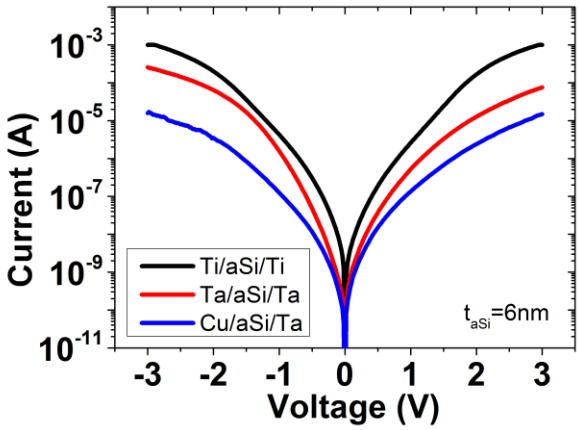


Figure 4.11. The electrode effect on the aSi rectifying layer performance. Reproduced from [26].

The LRS conduction mechanism of the integrated $Cu/Al_2O_3/aSi/Ta$ cell was further investigated. Figure 4.12(a) compares the log(I)–V plots of the $Cu/Al_2O_3/aSi/Ta$ cell in LRS and the Cu/aSi/Ta device. Both curves follow the same trend over the entire bias range. By simply scaling the control device current (by a factor of 62500), the log(I)–V curve

almost perfectly overlaps with the memory device curve in LRS state. Therefore, it is reasonable to assume that: (1) the LRS behavior of the $Cu/Al_2O_3/aSi/Ta$ cell is dominated by the rectifying layer; (2) the reduced absolute current level in the $Cu/Al_2O_3/aSi/Ta$ cell is likely a result from the effective cross-sectional area of the Cu filament/aSi contact in the cell. Based on these assumptions, from the current ratio of the two cases and the known electrode size of the Cu/aSi/Ta device, the diameter of the Cu filament in the integrated $Cu/Al_2O_3/aSi/Ta$ device is estimated to be ~9nm, which agreed well with the filament sizes obtained from direct measurements [10], [12]. As shown in Figure 4.12(b), the current shows negligible temperature dependence, consistent with a tunneling current model across the thin film [22]. Specifically, the I-V at high bias region can be well modeled by F-N tunneling (inset). From the fitting of the $\ln(J/E^2)$ - 1/E plot, the slope

$$K = -\frac{8\pi(2qm^*)^{1/2}\emptyset_B^{3/2}}{3h}$$

can be obtained, where m* ($0.1m_0$ [22]) is the tunneling effective mass, h is the Planck constant, and the barrier height Φ_B can be extracted to be 0.58eV. Optimizing the tunnel barrier height and the metal/aSi interface quality can lead to higher nonlinearity and current driving capability by choosing metal materials with proper work function and reactivity with aSi [20].

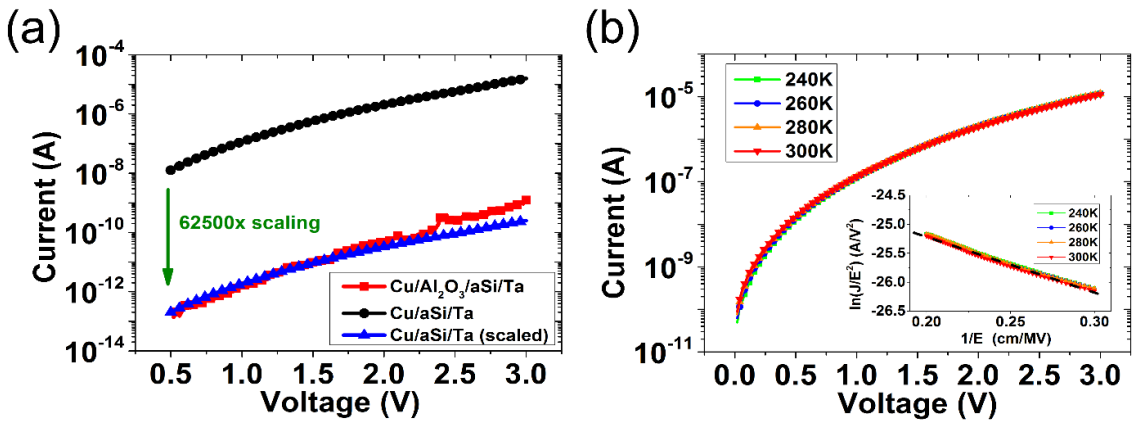


Figure 4.12. (a) Log(I)-V plots of the Cu/aSi/Ta control device (before/after scaling) and the Cu/Al_2O_3/aSi/Ta memory device in LRS. (b) Temperature dependence of the Cu/aSi/Ta device. The I-V of control sample can be modeled by F-N tunneling at high bias (inset). Reproduced from [26].

4.3.5. Retention and Endurance

The ultralow current operation does not come at a cost of performance degradation in the proposed structure. The Cu/Al2O3/aSi/Ta cell in this study shows stable retention in LRS and HRS for over 10^4 seconds at 100°C, as shown in Figure 4.13a. Over 500 SET/RESET pulse cycles can be obtained without the application of external current compliance (Figure 4.13b). The switching speed during the pulse test is limited by the RC delay in the system due to the very low programing/read current and the relatively large device size. We expect the switching speed to be dramatically improved by scaling down the device size hence reducing the device capacitance, along with the minimization of parasitic capacitances. Nevertheless, very fast switching (~ns level) may still be difficult to achieve for such device with extremely low operating current (~nA). On the other hand, these devices may be suitable for applications where low power is critical and relatively low speed (at the individual device level) can be tolerated, such as in neuromorphic circuits.

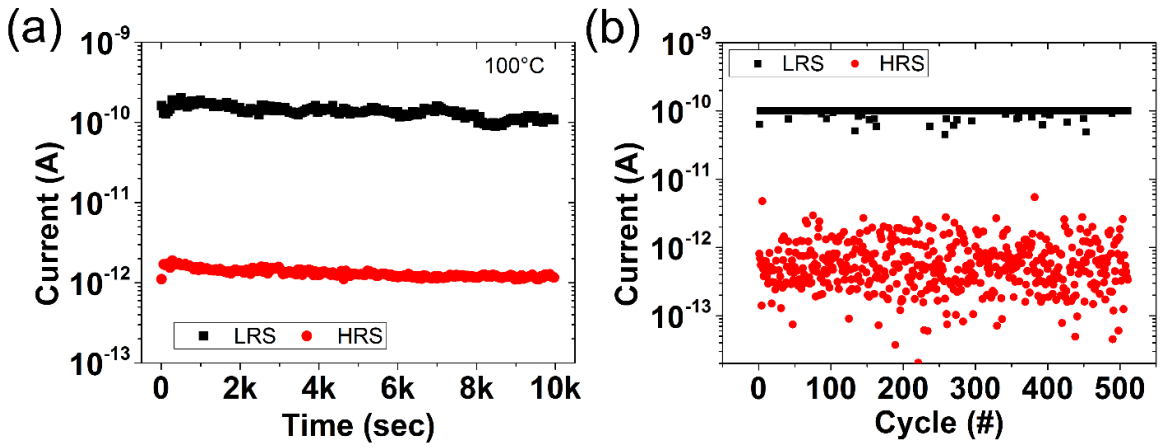


Figure 4.13. (a) HRS/LRS retention showing over 10000 seconds at 100 °C. (b) Pulse endurance test without current compliance. 3.5V/-3V pulses (10ms width) were used for SET/RESET in the endurance test. Read was performed with 2V pulses (10ms width). Reproduced from [26].

4.3.6. Multilevel Capability

Apart from the characteristic dimension (CD) scaling guided by the Moore's Law, multilevel operation will further reduce the cost per bit. By optimizing the program/erase procedures, such as adjusting current compliance [23] and pulse amplitude/width [24], multilevel cell storage can be demonstrated in a single memory device, thus allowing enhanced storage density without extra fabrication complexity.

In this work, multibit capacity can be obtained by the precise control on the filament growth. With different compliance current applied during the SET process ranging from 100pA to 10nA, the RRAM device can be programmed into distinct low resistance states (Levels 1/2/3 in Figure 4.14a). Stable device current can be still read for each state over the time, even Level 1 corresponds to an on-current of ~10pA. A well-defined read margin

exceeding 10 can be reliably achieved between four different states in the same device, as shown in Figure 4.14b.

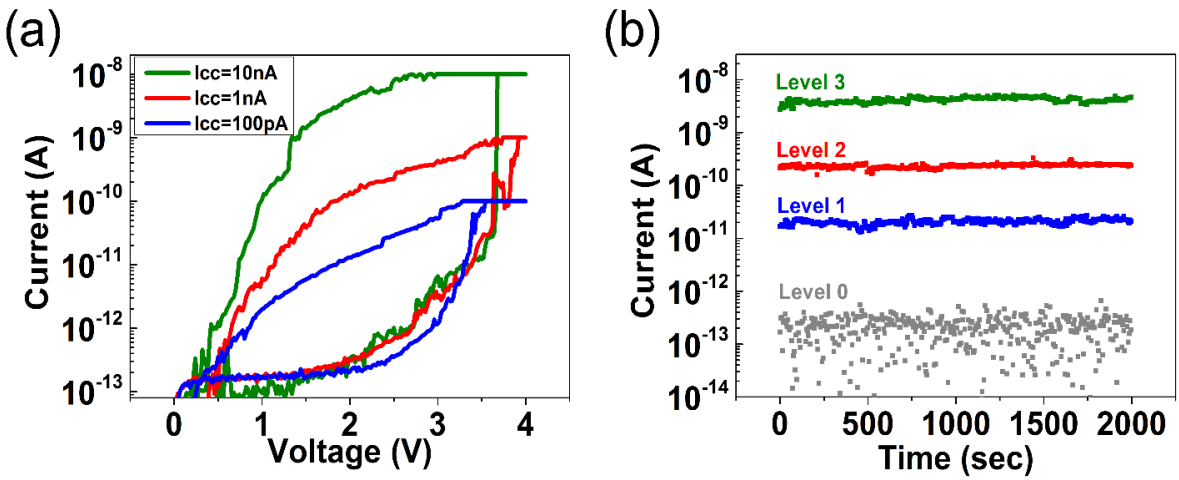


Figure 4.14. (a) Different compliance currents I_{cc} are employed during the SET process. (b) Multibit data retention measured from single device at room temperature. Reproduced from [26].

4.3.7. Device Variability and Optimization

Figure 4.15a shows the device-to-device variation of the forming voltage (red circles). Relatively narrow distribution of the formation process can be obtained. On the other hand, the cycle-to-cycle variations is non-negligible even for the same cell (black squares). The SET voltage was measured from the DC sweeps and defined as the voltage where the device current increases above a threshold value of 0.1nA. These different behaviors can be consistently explained by the facts that the forming voltage is determined by the initial film stack which can be very uniform between different cells; while the programming voltage is determined by the remaining filament shape after the previous reset which can exhibit larger variations than the initial virgin state configurations [25]. Figure 4.15b shows

the set/forming voltage as a function of the aSi thickness. Overall the devices with a thick rectifying layer require higher forming voltage and may be more susceptible to the voltage overshoot and programming failure issues, while lowering the rectifying layer thickness reduces the forming voltage but may lead to reliability issues such as film breakdown and copper injection under extremely high electric fields at reduced film thickness. The parameters of the switching/rectifying layer stack have to be carefully engineered, taking into account the inherent voltage divider effect in the series configuration.

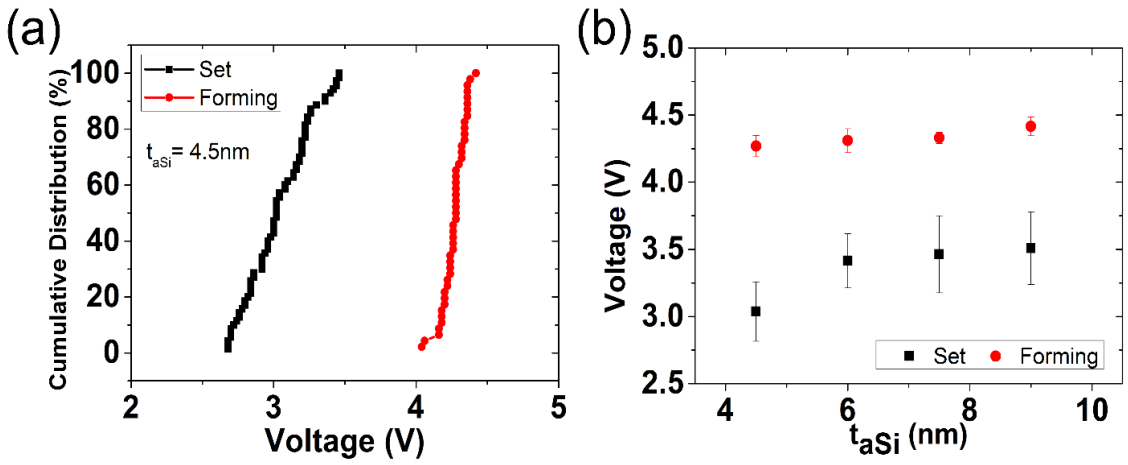


Figure 4.15. (a) cycle-to-cycle variation of V_{set} and device-to-device variation of V_{form} for cells with 4.5nm thick aSi. (b) V_{set} and V_{form} of the memory cells as a function of the aSi layer thickness. Reproduced from [26].

4.4. Conclusions

Sub-nA operation of RRAM devices has been demonstrated for the first time on a Cu based CBRAM with polysilicon in-cell resistor. A novel Cu/Al_2O_3/aSi/Ta RRAM cell with self-rectifying characteristics was further developed to address the series resistance and high-thermal budget issues in the polysilicon electrode devices. By combining a low

current RRAM with an amorphous silicon barrier/rectifying layer, the device exhibits low current (~nA), high on/off ratio (>100x) and pronounced nonlinearity. Stable retention, endurance and multilevel operation can be demonstrated by the precise filament growth control inherently offered by the device structure. Further optimizations on the device performances, e.g. higher nonlinearity and less parameter variations, can potentially lead to the application of such devices in future low-power large-scale crossbar memory arrays.

4.5. References

[1] J. Zhou, K.-H. Kim, and Wei Lu, "Crossbar RRAM Arrays: Selector Device Requirements During Read Operation," *IEEE Trans. Electron Devices*, vol. 61, no. 5, pp. 1369–1376, 2014.

[2] Sungho Kim, Jiantao Zhou, and W. D. Lu, "Crossbar RRAM Arrays: Selector Device Requirements During Write Operation," *IEEE Trans. Electron Devices*, vol. 61, no. 8, pp. 2820–2826, 2014.

[3] L. Zhang, S. Cosemans, D. J. Wouters, G. Groeseneken, M. Jurczak, and B. Govoreanu, "On the Optimal ON/OFF Resistance Ratio for Resistive Switching Element in One-Selector One-Resistor Crosspoint Arrays," *IEEE Electron Device Lett.*, vol. 36, no. 6, pp. 570–572, 2015.

[4] C. Schindler, M. Weides, M. N. Kozicki, and R. Waser, "Ultra-low current resistive memory based on Cu-SiO2," in *2008 IEEE Silicon Nanoelectronics Workshop*, pp. 1–2, 2008.

[5] Y. Wu, B. Lee, and H.-S. P. Wong, "Ultra-low power Al_2O_3-based RRAM with 1μA reset current," in *Proceedings of 2010 International Symposium on VLSI Technology, System and Application*, pp. 136–137, 2010.

[6] W. Kim, S. Il Park, Z. Zhang, Y. Yang-Liauw, D. Sekar, H.-S. P. Wong, and S. S. Wong, "Forming-free nitrogen-doped AlO_X RRAM with sub-μA programming current," *2011 Symp. VLSI Technol.*, vol. 31, pp. 22–23, 2011.

[7] K.-H. Kim, S. Hyun Jo, S. Gaba, and W. Lu, "Nanoscale resistive memory with intrinsic diode characteristics and long endurance," *Appl. Phys. Lett.*, vol. 96, no. 5, p. 053106, 2010.

[8] S.-G. Park, M. K. Yang, H. Ju, D.-J. Seong, J. M. Lee, E. Kim, S. Jung, L. Zhang, Y. C. Shin, I.-G. Baek, J. Choi, H.-K. Kang, and C. Chung, "A non-linear ReRAM cell with sub-1μA ultralow operating current for high density vertical resistive memory (VRRAM)," in *2012 International Electron Devices Meeting*, pp. 20.8.1–20.8.4, 2012.

[9] Y. S. Chen, H. Y. Lee, P. S. Chen, W. S. Chen, K. H. Tsai, P. Y. Gu, T. Y. Wu, C. H. Tsai, S. Z. Rahaman, Y. De Lin, F. Chen, M. J. Tsai, and T. K. Ku, "Novel defects-trapping TaO_X/HfO_X RRAM with reliable self-compliance, high nonlinearity, and ultra-low current," *IEEE Electron Device Lett.*, vol. 35, no. 2, pp. 202–204, 2014.

[10] Y. Yang, P. Gao, S. Gaba, T. Chang, X. Pan, and W. Lu, "Observation of conducting filament growth in nanoscale resistive memories," *Nat. Commun.*, vol. 3, p. 732, 2012.

[11] Y. Yang, P. Gao, L. Li, X. Pan, S. Tappertzhofen, S. Choi, R. Waser, I. Valov, and W. D. Lu, "Electrochemical dynamics of nanoscale metallic inclusions in dielectrics," *Nat. Commun.*, vol. 5, no. May, p. 4232, 2014.

[12] U. Celano, L. Goux, A. Belmonte, K. Opsomer, A. Franquet, A. Schulze, C. Detavernier, O. Richard, H. Bender, M. Jurczak, and W. Vandervorst, "Three-Dimensional Observation of the Conductive Filament in Nanoscaled Resistive Memory Devices," *Nano Lett.*, vol. 14, no. 5, pp. 2401–2406, 2014.

[13] S. Choi, Y. Yang, and W. Lu, "Random telegraph noise and resistance switching analysis of oxide based resistive memory," *Nanoscale*, vol. 6, no. 1, pp. 400–404, 2014.

[14] S. Ambrogio, S. Balatti, A. Cubeta, A. Calderoni, N. Ramaswamy, and D. Ielmini, "Understanding switching variability and random telegraph noise in resistive RAM," in *2013 IEEE International Electron Devices Meeting*, pp. 31.5.1–31.5.4, 2013.

[15] Y. Y. Chen, M. Komura, R. Degraeve, B. Govoreanu, L. Goux, A. Fantini, N. Raghavan, S. Clima, L. Zhang, A. Belmonte, A. Redolfi, G. S. Kar, G. Groeseneken, D. J. Wouters, and M. Jurczak, "Improvement of data retention in HfO_2/Hf 1T1R RRAM cell under low operating current," in *2013 IEEE International Electron Devices Meeting*, pp. 10.1.1–10.1.4, 2013.

[16] S. Gaba, S. Choi, P. Sheridan, T. Chang, Y. Yang, and W. Lu, "Improvement of RRAM Device Performance Through On-Chip Resistors," in *Procs. MRS*, vol. 1430, pp. 177–182, 2012.

[17] H. Spahr, J. Reinker, T. Bülow, D. Nanova, H.-H. Johannes, and W. Kowalsky, "Regimes of leakage current in ALD-processed Al_2O_3 thin-film layers," *J. Phys. D. Appl. Phys.*, vol. 46, no. 15, p. 155302, 2013.

[18] S. H. Jo, K. H. Kim, and W. Lu, "High-density crossbar arrays based on aSi memristive system," *Nano Lett.*, vol. 9, no. 2, pp. 870–874, 2009.

[19] L. Zhang, A. Redolfi, C. Adelmann, S. Clima, I. P. Radu, Y.-Y. Chen, D. J. Wouters, G. Groeseneken, M. Jurczak, and B. Govoreanu, "Ultrathin Metal/Amorphous-Silicon/Metal Diode for Bipolar RRAM Selector Applications," *IEEE Electron Device Lett.*, vol. 35, no. 2, pp. 199–201, 2014.

[20] L. Zhang, B. Govoreanu, A. Redolfi, D. Crotti, H. Hody, V. Paraschiv, S. Cosemans, C. Adelmann, T. Witters, S. Clima, Y. Y. Chen, P. Hendrickx, D. J. Wouters, G. Groeseneken, and M. Jurczak, "High-drive current (>1MA/cm^2) and highly nonlinear (>10^3) TiN/amorphous-Silicon/TiN scalable bidirectional selector with excellent reliability and its variability impact on the 1S1R array performance," in *Technical Digest - International Electron Devices Meeting, IEDM*, pp. 6.8.1–6.8.4, 2015.

[21] S. Gaba, F. Cai, J. Zhou, and W. D. Lu, "Ultralow Sub-1-nA Operating Current Resistive Memory With Intrinsic Non-Linear Characteristics," *IEEE Electron Device Lett.*, vol. 35, no. 12, pp. 1239–1241, 2014.

[22] J. M. Shannon and K. J. B. M. Nieuwesteeg, "Tunneling effective mass in hydrogenated amorphous silicon," *Appl. Phys. Lett.*, vol. 62, no. 15, p. 1815, 1993.

[23] S.-S. Sheu, P.-C. Chiang, W.-P. Lin, H.-Y. Lee, P.-S. Chen, Y.-S. Chen, T.-Y. Wu, F. T. Chen, K.-L. Su, M.-J. Kao, K.-H. Cheng, and M.-J. Tsai, "A 5ns fast write multi-level non-volatile 1 K bits RRAM memory with advance write scheme," in *2009 Symposium on VLSI Circuits*, pp. 82–83, 2009.

[24] F. Alibart, L. Gao, B. D. Hoskins, and D. B. Strukov, "High precision tuning of state for memristive devices by adaptable variation-tolerant algorithm," *Nanotechnology*, vol. 23, no. 7, p. 075201, 2012.

[25] U. Celano, G. Giammaria, L. Goux, A. Belmonte, M. Jurczak, and W. Vandervorst, "Nanoscopic structural rearrangements of the Cu-filament in conductive-bridge memories," *Nanoscale*, vol. 8, no. 29, pp. 13915–13923, 2016.

[26] J. Zhou, F. Cai, Q. Wang, B. Chen, S. Gaba, and W. D. Lu, "Very Low-Programming-Current RRAM With Self-Rectifying Characteristics," *IEEE Electron Device Lett.*, vol. 37, no. 4, pp. 404–407, 2016.

Chapter 5

Reconfigurable Devices Based on Complex Oxide Heterojunctions

5.1. Introduction

In previous chapters, selector and low-power RRAM devices have been investigated by experimental and simulation studies. In this chapter, we aim to utilize the ionic motions in resistive switching processes in logic devices, and demonstrate reconfigurable transistors for potential neuromorphic computing applications. Specifically, several recent studies with similar device concepts of a programmable resistive transistor have been reported, using the hybrid integration where the RRAM device was stacked on the gate terminal of bulk silicon MOSFET [1] or horizontal nanowire transistor [2]. The nonvolatile modulation of transistor transconductance was essentially discrete and demonstrated by the voltage divider effect between the fixed dielectric capacitor and the variable capacitor of resistive switching medium. Two metal gates, namely the resistive-switch gate and the control gate, have to be processed in one transistor device.

Different from these proposed structures, our device structure is expected to directly control the carrier concentration in a gradual and analog fashion with only one gate [3]. The device is based on the two-dimensional electron gas (2DEG) formed at the complex oxide heterojunction between the lanthanum aluminate ($LaAlO_3$, or LAO) and strontium

titanate ($SrTiO_3$, or STO) interface. Figure 5.1 shows the band diagram of this oxide heterostructure.

Over the last decade, the $LaAlO_3/SrTiO_3$ interface has attracted intensive attention due to its exotic physical properties. Even though $LaAlO_3$ thin film and $SrTiO_3$ substrate are individually insulators, a 2DEG was surprisingly discovered at their interface [4]. The $LaAlO_3/SrTiO_3$ heterojunction exhibits many other interesting phenomena that do not normally co-exist in bulk materials, including superconductivity [5], magnetism [6], [7], enhanced Rashba spin-orbital coupling [8] and persistent photoconductivity [9], [10]. These interesting properties can potentially lead to not only important advances in the condense matters theory, but also promising device applications.

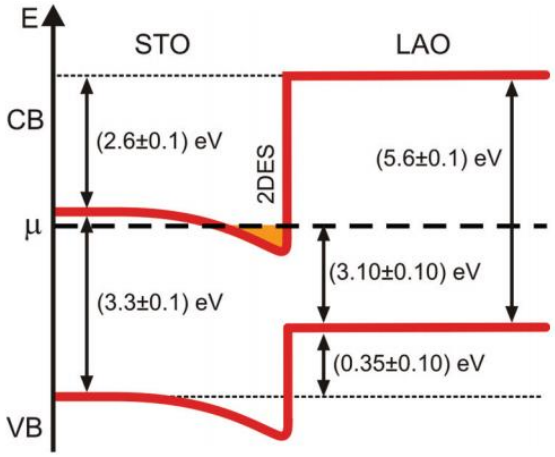


Figure 5.1. Band diagram of the $LaAlO_3/SrTiO_3$ heterostructure. Reproduced from [11].

The microscopic origin to these phenomena still remains highly debated. In particular, the origin of 2DEG is still not clear. As shown in Figure 5.2, intrinsic electronic reconstruction ("polar catastrophe") [12], [13], oxygen vacancy [14]–[16] and cation mixing [17]–[19] have been suggested as the likely mechanisms, although none of these theories could perfectly agree with the experimental observations. More recently, Yu et al.

proposed a polarity-induced defect mechanism, combining the polar discontinuity with surface oxygen vacancy and interfacial anti-site cation mixing [20]. There seems to be a consensus that these different mechanisms may not be mutually exclusive, and instead jointly contribute to the formation of 2DEG at the $LaAlO_3/SrTiO_3$ heterojunction.

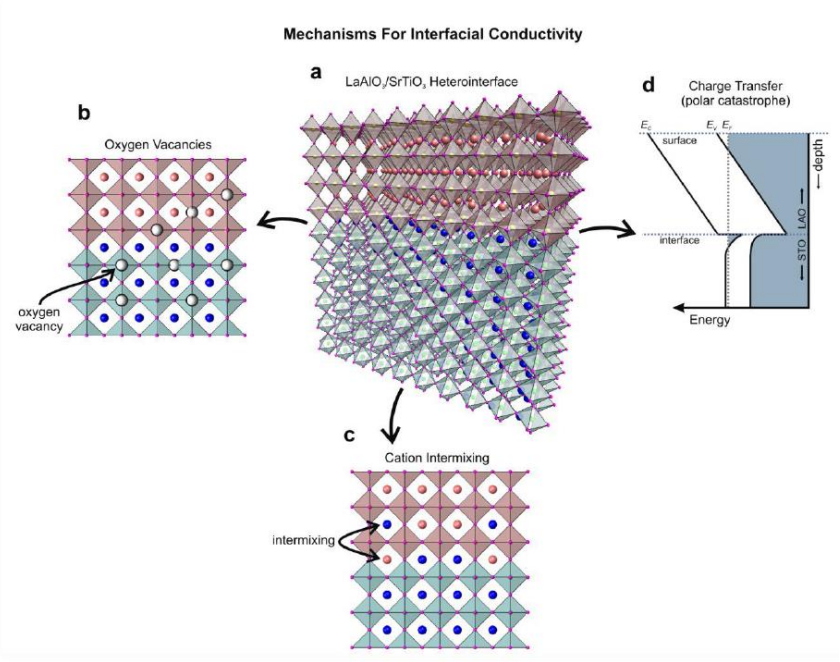


Figure 5.2. Possible mechanisms for 2DEG formation in oxide heterostructure. (a) the $LaAlO_3/SrTiO_3$ interface. (b) Oxygen vacancy donate charge to 2DEG. (c) Cation intermixing (La/Sr) dopes the interface. (d) Built-in electric field due to charge transfer. Reproduced from [21]

The electric field effect and the memory effect of the $LaAlO_3/SrTiO_3$ interface conductivity have been intensively investigated [12], [15], [22]–[26]. In most studies, the interface conductivity was modulated by either a back-gate with a very high gate voltage, or a C-AFM tip. Both methods were not applicable to the practical device applications. In the following sections, we demonstrate a reconfigurable top-gated transistor based on the $LaAlO_3/SrTiO_3$ 2DEG system. Sample preparation and device fabrication processes will be introduced in Section 5.2. Basic transistor operations and repeatable switching behaviors

of the proposed device structure are carefully characterized by electrical measurement and discussed in Section 5.3. The microscopic origins of interface conductance switching are investigated and modeled in Section 5.4.

5.2. Sample Preparation and Device Fabrication

The $LaAlO_3$ films used in this study were epitaxially deposited on TiO_2-terminated (001) $SrTiO_3$ substrates by pulsed laser deposition using a 248nm KrF excimer (Coherent Inc.). The samples were provided to us by our collaborators at Case Western Reserve University (Prof. Alp Sehirlioglu). Before growth, the substrates were etched with a chemical solution of ammonium fluoride and hydrofluoric acid at pH=6 to obtain a TiO_2-terminated surface and then pre-annealed at 950 °C for one hour in an oxygen-rich atmosphere. The film growth was conducted in an oxygen pressure of 10^{-4} torr at 750 °C, at a repetition rate of 2Hz. Two samples with $LaAlO_3$ thickness of 5 and 15 u.c. were prepared.

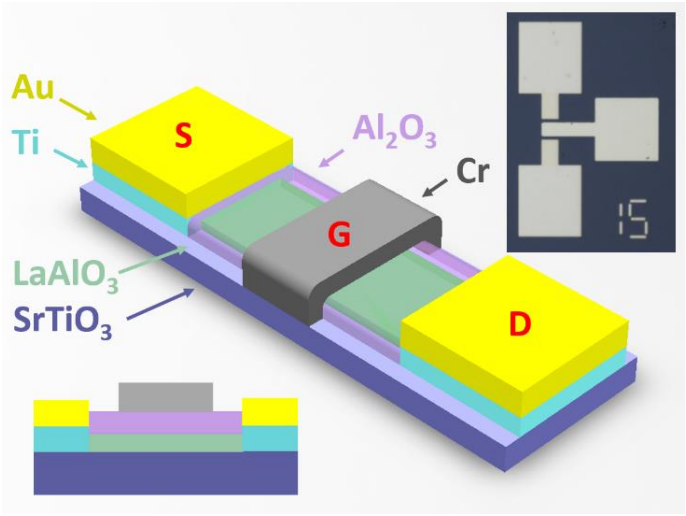


Figure 5.3. Schematic view of the device structure: the left inset illustrates the cross-section along the channel direction; the right inset is the microscopic image of a device (W/L=15μm/15μm).

After receiving the LAO/STO samples, the transistor device was processed at UM in a gate-last manner. First, photolithography was performed to define the conductive channel of the transistor device. Using photoresist (SPR220-3.0) as hard mask, a reactive ion etching process with argon and fluoride chemistry ($Ar/SF_6/C_4F_8$, Lam 9400) was carried out to etch the $LaAlO_3$ film outside the channel region and isolate different transistor devices. Ti/Au (100/1500Å) electrodes were then patterned and deposited by e-beam evaporation, forming the source/drain Ohmic contacts. 10nm Al_2O_3 was deposited by atomic layer deposition at 150 °C as the gate dielectric, using the standard tri-methyl-aluminum (TMA) and water precursors in an Oxford OpAL system. Cr top gate (1000Å) and Au passivation (1000Å) were processed by DC sputtering (Kurt. J. Lesker Lab-18) and patterned using photolithography and lift processes. Figure 5.3 shows the schematic and the optic image of a completed device.

5.3. Electrical Measurement

The current-voltage (I-V) characteristics of device were measured using a Keithley 4200 semiconductor analyzer. Capacitance-voltage tests were performed using an Andeen-Hagerling 2700A capacitance bridge and a Tektronix AFG 3101 arbitrary signal generator. All the electrical measurement was conducted in dark and under vacuum.

5.3.1. Virgin State and Forming Process

As shown in Figure 5.4, the as-fabricated device with 5 u.c. $LaAlO_3$ (W/L=30μm/15μm) exhibits an n-type FET behavior; while the drain current does not saturate at room temperature until the gate voltage is more negative than -3V. At -3V V_{gs}, the saturation

current density is calculated to be 8.3μA/μm at 5V $V_{d,}$, comparable to the reported works [23], [26].

We note that at room temperature the drain current does not deplete even at -5V $V_{g.}$ When increasing the gate voltage sweep range to -10V, it was found that the device showed a very distinct transport behavior during the negative sweep (0V to -10V) vs. during the back-sweep (-10V to 0V). During the negative gate sweep, the drain current decreases sharply at V_g~-6V, accompanied by a significant increase in gate (leakage) current. The gate current reaches a peak value at V_g= -7.8V when the drain current fully turns off. During the subsequent back-sweep, the transistor device shows a dramatically reduced drain current and a large hysteresis can be observed.

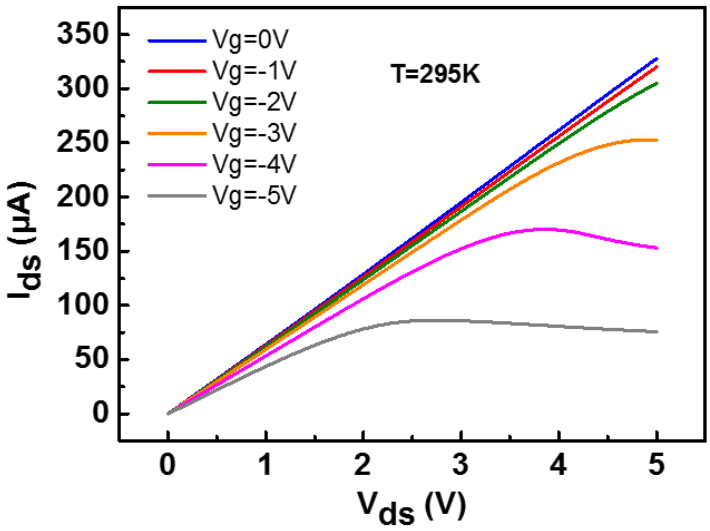


Figure 5.4. Output curves of a virgin $LaAlO_3/SrTiO_3$ transistor (W/L=30μm/15μm)

After this one-time negative DC sweep, termed as the "forming process", the device current will increase slowly and eventually stabilize, typically after a few hours (depending

on the forming voltage and temperature). However, the drain current does not recover to its (high value) virgin state even after applying positive gate sweeps again, suggesting an irreversible process occurring within the oxide stacks during the forming process.

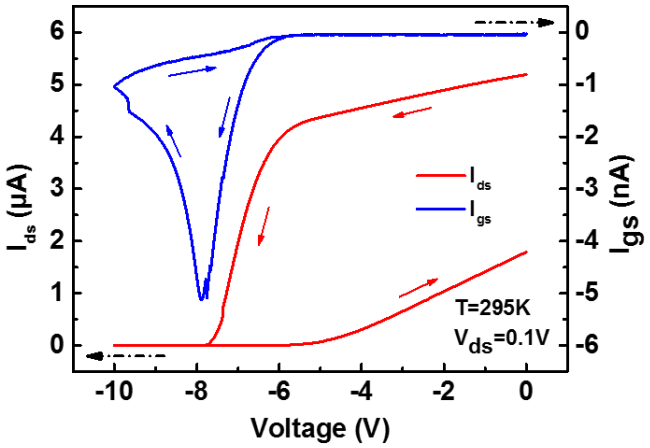


Figure 5.5. Drain current (red) and gate current (blue) measured during the forming process of the $LaAlO_3/SrTiO_3$ transistor

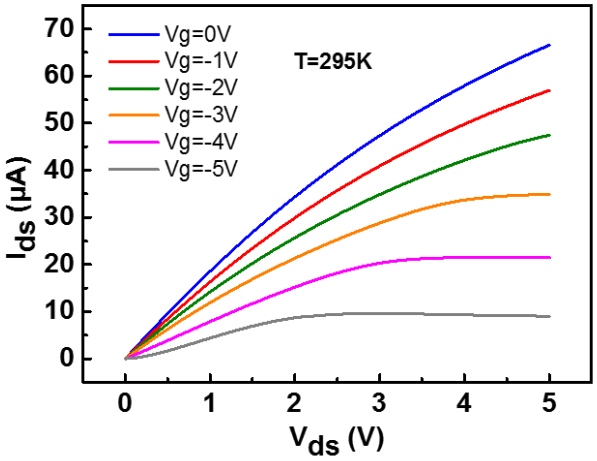


Figure 5.6. Stabilized output curves of the transistor measured after the forming process. Compared to Figure 5.4, the current was ~5 times lower.

Based on the transconductance extracted from the transistor transfer curves and the directly measured gate capacitance, the mobility and carrier concentration of 2DEG system

at the equilibrium state are obtained as the functions of temperature in Figure 5.7, consistent with the Hall measurement results in past studies [23], [24].

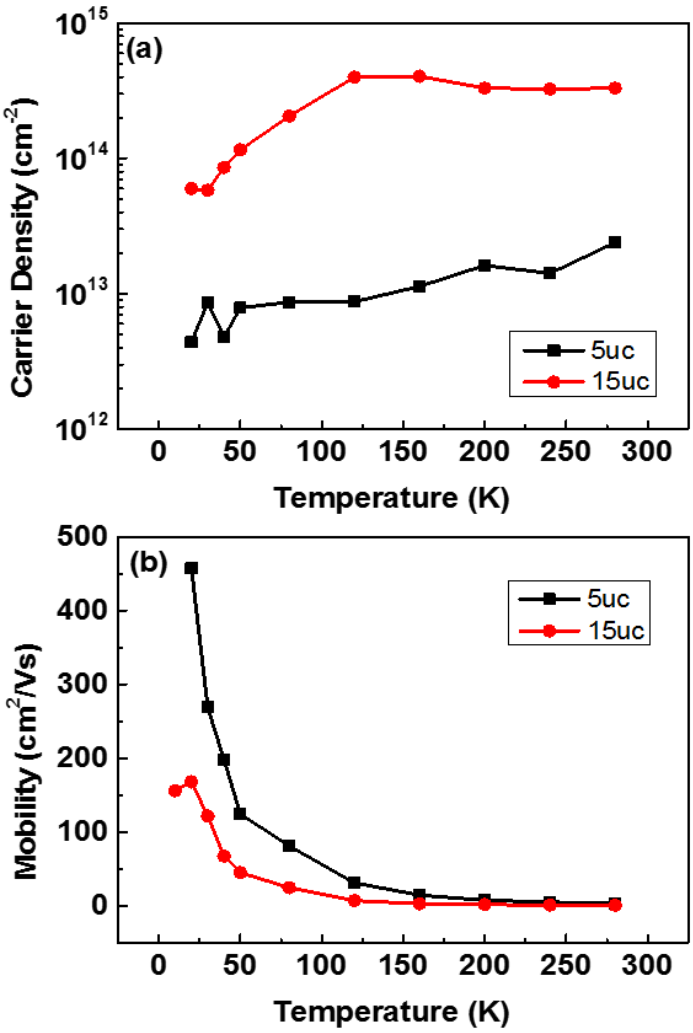


Figure 5.7. Extracted 2DEG (a) carrier concentration and (b) mobility at different temperatures.

5.3.2. Reconfigurable Modulation

The device after the forming process shows stable transistor behavior with a reduced (negative) threshold voltage compared with the virgin state. Additionally, the threshold voltage and the transistor behavior can be further tuned between an ON state (with overall

a more negative threshold voltage) and an OFF state (with overall a less negative threshold voltage). Starting from an equilibrium state after the device has stabilized, this all-oxide transistor can be reconfigured to the ON (OFF) state with high (low) channel conductivity by applying an external positive (negative) bias stress. Figure 5.8 (a) and (b) show the device output characteristics (I_{ds}-V_{ds}) of the ON (OFF) state, obtained after 300 seconds of 5V (-5V) stress programming (erasing), respectively. The current read at 0V V_g exhibits a 5-fold modulation between the ON and OFF state (45.6μA vs. 9.7μA @5V V_{ds}).

Figure 5.9 illustrates the device transfer characteristics (I_{ds}-V_{gs}) in the ON and OFF states, showing a clearly shift of the threshold voltage and transconductance. The current window ($I_{ds(ON)}/I_{ds(OFF)}$) between the two states is also plotted, and over two orders of magnitude $I_{ds(ON)}/I_{ds(OFF)}$ can be obtained at a negative V_{gs} read voltage (e.g. at V_{gs} = 4V). After the programming and erasing bias stress, the threshold voltage of the transistor V_{th} shifts between the ON state (-5V) and the OFF state (-2.7V).

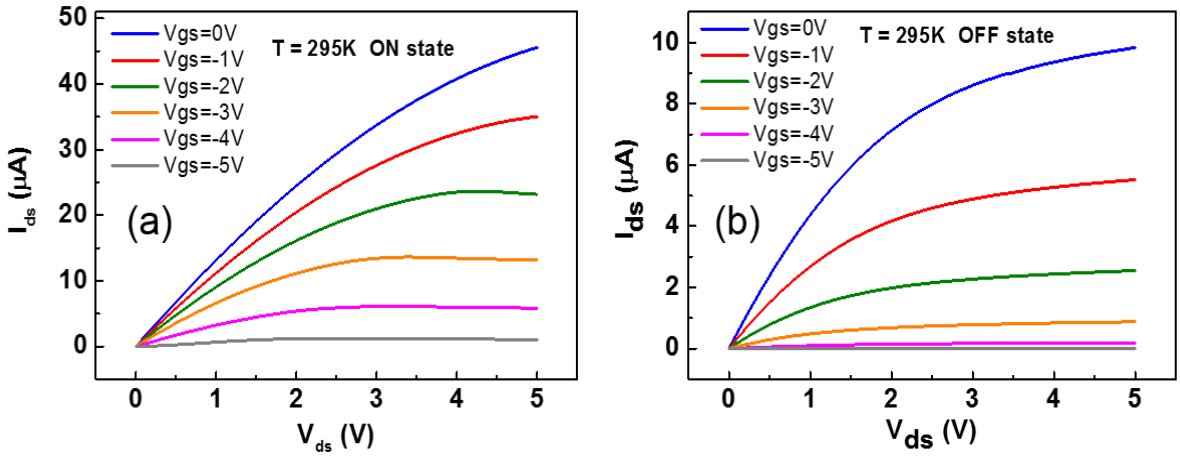


Figure 5.8. Output curves of (a) the ON state programmed by 300sec 5V stress; and (b) the OFF state erased by 300sec -5V stress. (W/L=30μm/25μm)

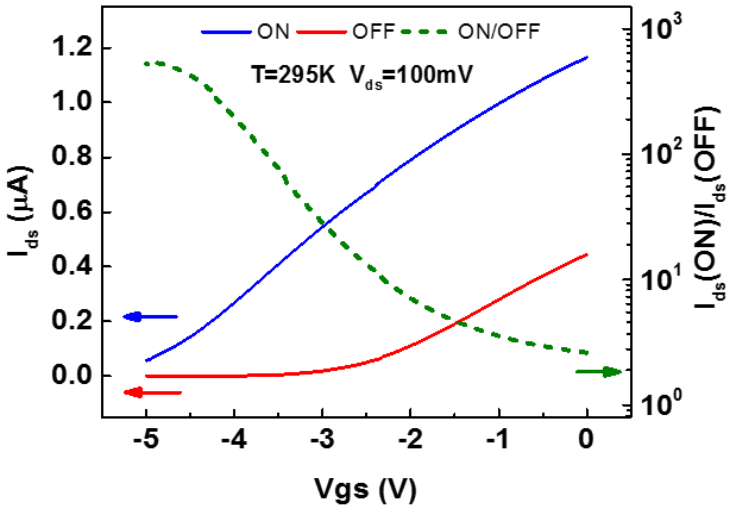


Figure 5.9. Transfer curves of the ON and OFF states in the $LaAlO_3/SrTiO_3$ transistor. The dash line shows the current ratio.

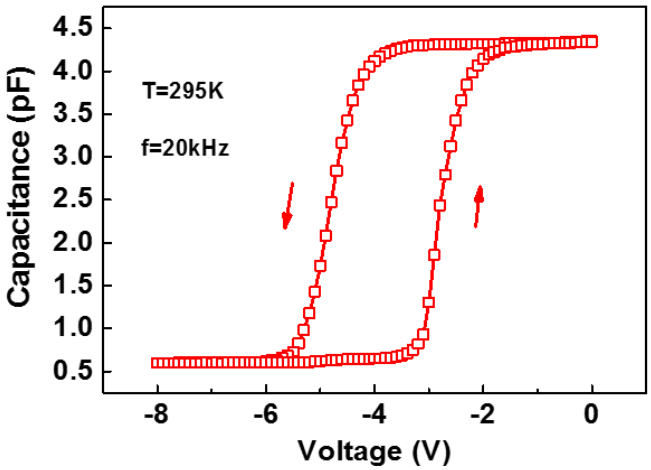


Figure 5.10. Capacitance-voltage (C-V) curve for the gate capacitance. The drain and the source were grounded during the measurement. 30mV$_{ac}$ (rms) amplitude and a DC sweep rate of 100mV/sec are used.

The capacitance-voltage curve of the LAO/STO device is shown in Figure 5.10. A constant capacitance (4.3pF) was measured above V_{gs}=-4V; while the capacitance

significantly reduces at V_{gs}=-5.5V, indicating the depletion of the 2DEG channel. The gate capacitance, directly measured from the C-V measurement, was further used to calculate the carrier mobility/concentration in the 2DEG. Interestingly, during gate voltage sweep (0V to -8V back to 0V), the C-V curve shows a counter-clockwise hysteresis loop with ~2V shift in V_{th}, consistently with the observed V_{th} in transport measurements (Figure 5.8/5.9) of the same device. We note the modulation of the threshold voltage is non-volatile. As shown in Figure 5.11, the device stays in the ON (OFF) state after the program (erase) stress has been removed. Despite a gradual decay, the ON and OFF states maintain a large read-out window (~100%) over at least 10^4 seconds.

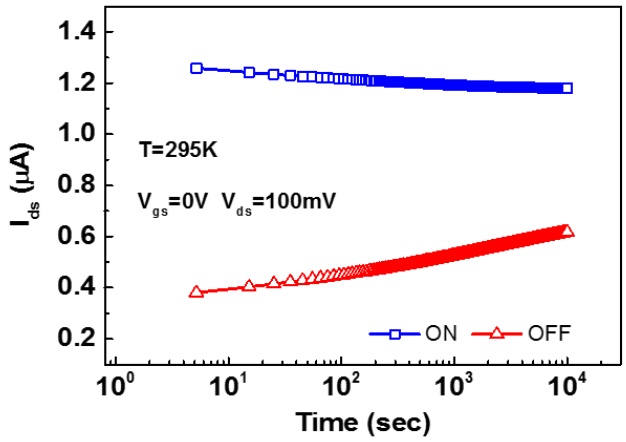


Figure 5.11. Data retention of the ON state and the OFF state.

From a classic transistor model [27], the shift of V_{th} can be explained as the movement of mobile charges across the oxide layers driven by the electric field during the programming/erasing stresses. By assuming a parallel plate model and using parameters as $\varepsilon(LaAlO_3)$=25 [28], $\varepsilon(Al_2O_3)$=9 [29], the areal charge density to cause such shift in threshold voltage is approximately 10^{13} cm^{-2}, if all the charges are assumed to move back

and forth between the LAO/STO interface and the Al_2O_3 surface. We note this estimated amount of mobile charge is of the same order of magnitude as the estimated 2DEG carrier concentration.

5.3.3. Cycling Operations

Program/erase (P/E) cycling tests were also conducted to verify the device operation. Immediately after applying positive/negative program/erase pulse, the transfer curve was measured to check the device status. As shown in Figure 5.12, the device exhibited repeatable ON/OFF states during each cycle, including the drain current and the threshold voltage. The targeted conduction states and the V_{th} shift can be accurately controlled by adjusting the width/amplitude of P/E pulses on the fly, which potentially enables the "memistor" functionality for neuromorphic circuit applications[3], [30].

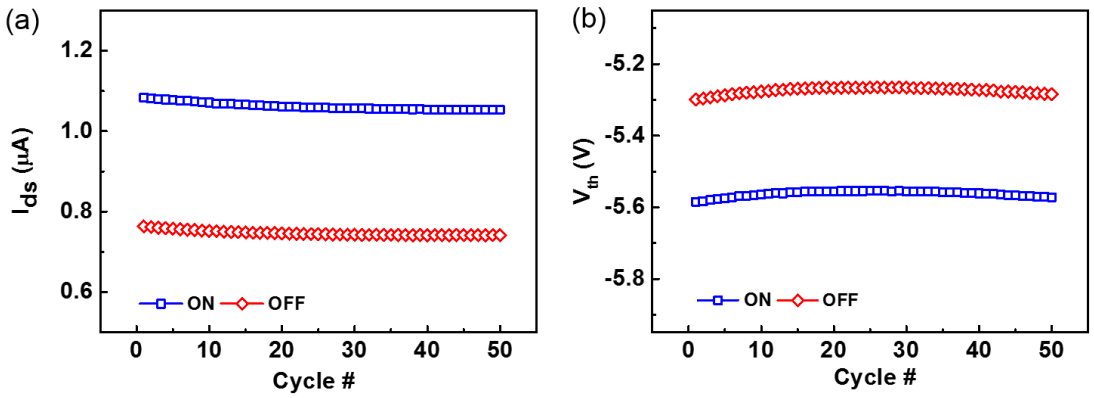


Figure 5.12. (a) Drain current and (b) threshold voltage measured over 50 P/E cycles. 60sec 5V positive pulse was used for each programming, and 75sec -5V negative pulse was used to erase the $LaAlO_3/SrTiO_3$ device. The 50 cycles showed repeatable ON/OFF states.

5.3.4. Switching Dynamics: Stress Test

To shed light on the observed switching behavior, stress tests that enable in-situ monitoring of the time-dependent, cumulative effects of the applied electric field on the 2DEG conductivity are carried out. Figure 5.13a shows the normalized change of the drain current (measured at ±5V V_g) during a 5V (-5V) V_g stress over a 300 second period, starting from the equilibrium state. During the positive V_g stress, 26% increase in I_{ds} is observed at 280K, whereas ~80% I_{ds} reduction is observed for negative V_g. Figure 5.13b shows the temperature dependence of the current modulation during a positive gate bias. It is evident that the current (i.e. conductance) increase is faster at higher temperatures. For instance, after 300 second stress, I_{ds} increases by 44%, 55% and 84% at 295K, 310K and 325K respectively. These results indicate that the device switching process is thermally assisted.

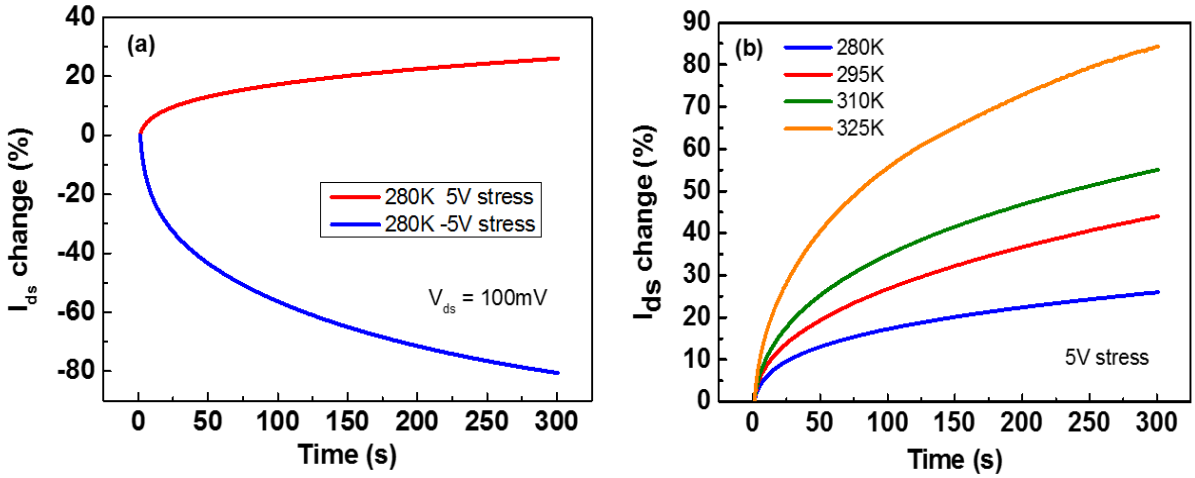


Figure 5.13. (a) Normalized modulation of drain current with 5V/-5V stress at 280K (b) temperature dependence of current modulation with 5V stress.

To perform semi-quantitative analysis on the switching dynamics, we extract the stress time needed to achieve a specific current modulation (e.g. 20% increase) and plot the

required stress time vs. temperature and the stress voltage. As shown in Figure 5.14a, the relation between the required stress time t_w and the absolute temperature T can be fitted with an Arrhenius plot with $ln(t_w)$ proportional to $1/T$. Similarly, the required stress time t_w is found to reduce exponentially with the stress voltage (Figure 5.14b). As a result, the stress time required to reach a certain conductivity state can be empirically described as

$$t_w = t_0 \cdot exp\left(\frac{E_A - l \cdot E}{k_B T}\right) \qquad (1)$$

where E_A is an effective activation energy, k_B is the Boltzmann constant, E is the electric field. t_0 is the inverse of the attempt frequency, and l represents a characteristic length. Eq. (1) includes effects from the electric field (E) as well as thermal effects represented by the local temperature (T). From the experimental data in Figure 5.14, E_A and l were extracted to be 0.67eV and 3Å, respectively.

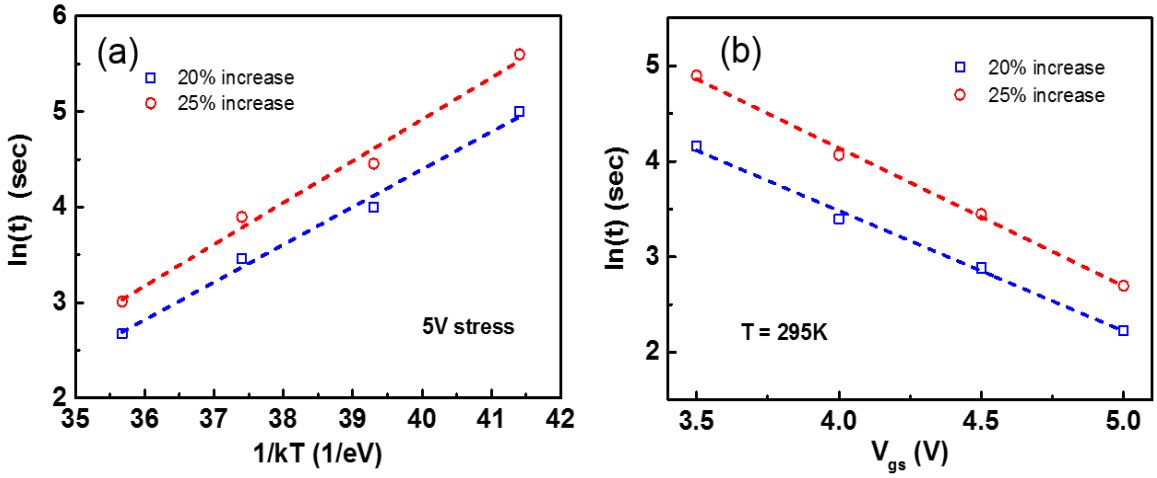


Figure 5.14. (a) Arrhenius plot of the required stress time as a function of temperature under 5V stress. (b) Required stress time as a function of the applied positive bias at a fixed temperature (295K). In (a) and (b), results of switching time to achieve 20% and 25% current increase are presented.

5.3.5. Switching Dynamics: Pulse Test

The data in Figure 5.13 and Figure 5.14 represent the switching dynamics of the LAO/STO reconfigurable device on a relatively long time scale (hundreds of seconds). To reveal transient behavior of the device that may provide better understand the microscopic origin of the conductance switching, a train of short pulses (<=100ms) were applied on the gate electrode and the drain current response was monitored. In Figure 5.15, the normalized change of I_{ds} ($\Delta I_{ds}/I_{ds(0)}$, $I_{ds(0)}$ is the initial drain current measured before any pulse is applied) is plotted after each pulse in the pulse train, at three different temperatures. With a fixed drain voltage V_{ds} of 100mV, the pulse train includes 100 consecutive positive pulses (4V, 100ms, Pulse# 1-100), followed by 100 consecutive negative pulses (-4V, 100ms, Pulse# 101-200). Each pulse is followed by a small read pulse (0.1V, 50μs) which allows reading the drain current without changing the device state. At 280K, I_{ds} increases gradually after the application of each positive pulse and decreases gradually after the application of each negative pulse, consistent with the observations shown in Figure 5.13. When the temperature is cooled down to 220K, however, the response of the current change becomes much smaller. Specifically, the first positive pulse leads to a small reduction in the drain current, while the subsequent positive pulses gradually increase the drain current. The opposite trend is observed for the negative pulses. Interestingly, reducing the temperature to 160K leads to a completely opposite behavior compared to the behavior at high temperatures (e.g. 280K), where a positive (negative) pulse leads to a decrease (increase) of I_{ds} instead. The reversal of the I_{ds} modulation direction shown in Figure 5.15 suggests that competing mechanisms may co-exist and become dominate at different temperature regimes.

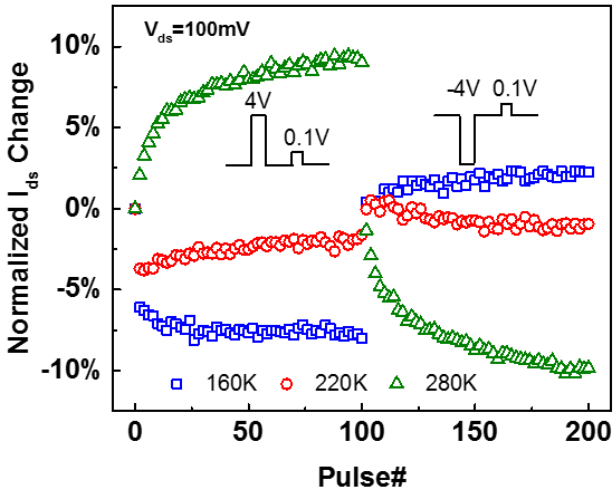


Figure 5.15. Drain current I_{ds} change measured after repeated positive (4V, #1-100) and negative (-4V, #101-200) gate pulses at different temperatures. The width of 4V/-4V pulses is 100ms. I_{ds} is read with fixed V_{gs} (100mV) and V_{ds} (100mV).

The occurrence of competing mechanisms can be better illustrated in Figure 5.16, where the accumulated change of I_{ds} over all 100 positive pulses is plotted as a function of temperature, for different pulse voltage amplitudes. This figure clearly shows a transition from a reduction of conductance (negative changes in I_{ds}) under positive gate voltage stress at low temperatures to an enhancement of conductance (positive change in I_{ds}) under the same stress conditions at high temperatures. The same trend is observed for stresses with different pulse amplitudes, with a higher amplitude leading to a larger modulation, both for conductance reduction at low temperatures and conductance enhancement at high temperatures. At the transition temperature regime (220~240K) very weak modulation was observed from the pulses, suggesting that effects from the different mechanisms are mutually cancelled.

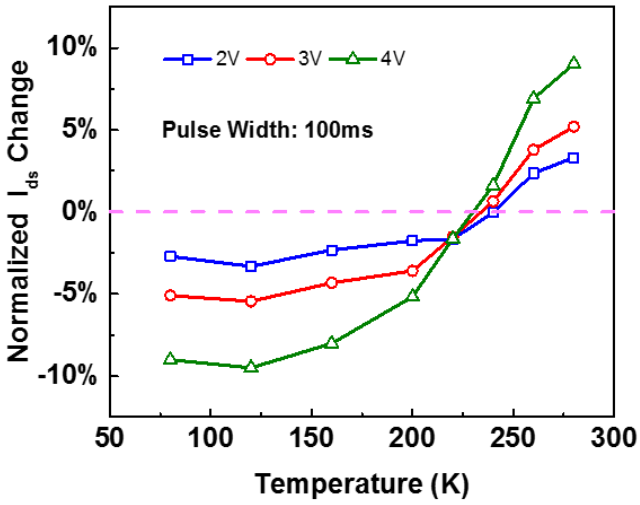


Figure 5.16. The accumulated I_{ds} change measured after 100 positive pulses (100ms) as a function of temperature and pulse amplitude. Fixed V_{gs} (100mV) and V_{ds} (100mV) are used for reading I_{ds}.

The influence of the pulse width on the conductance modulation near the transition temperature (220K) was investigated in detail (Figure 5.17). With different pulse widths (1/10/100ms), the first positive pulse always causes a reduced I_{ds}. However, a remarkable distinction can be found during the subsequent positive pulses. When short pulses of 1ms are applied, I_{ds} keeps decreasing and eventually becomes saturated. For longer pulses of 10ms, the LAO/STO device generally maintains the same drain current. At even longer pulses of 100ms, a rebound in I_{ds} is observed. The exactly reverse trend can be observed in Figure 5.17 when negative pulses (#101-200) are applied. Based on these findings, the conductance modulation observed in these reconfigurable devices can again be attributed to two competing processes. Taking positive pulses as an example, one of the processes is comparatively fast and dominates at low temperatures. It leads to a reduction of the

interface 2DEG conductance; while the other process is slow and dominants at high temperatures, and leads to an enhancement of the interface 2DEG conductance.

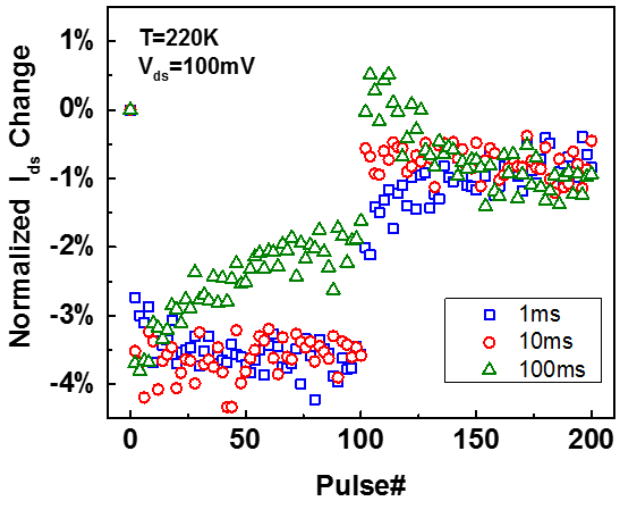


Figure 5.17. Pulse-width dependency of I_{ds} change at 220K. The amplitude of positive/negative pulse is 4V/-4V. I_{ds} is read by applying fixed V_{gs} (100mV) and V_{ds} (100mV).

5.4. Microscopic Origins and Simulation Study

Field-induced conductance switching behaviors have been reported in the LAO/STO system with a back-gate[12], [25] or a C-AFM tip[15], [31], [32]. In the previous studies, the memory effect on the interface conductivity was generally believed to originate from AFM-tip-induced water-ion injection[15], [31]–[33], field driven migration of oxygen vacancies (V_Os)[34] or the presence of charge trapping states in the STO layer[35], [36]. Based on these earlier findings and the observed transition of the conductance modulation effects, the current modulation observed in our LAO/STO device may be explained by the competing effects of V_O migration and electron trapping in the LAO layer.

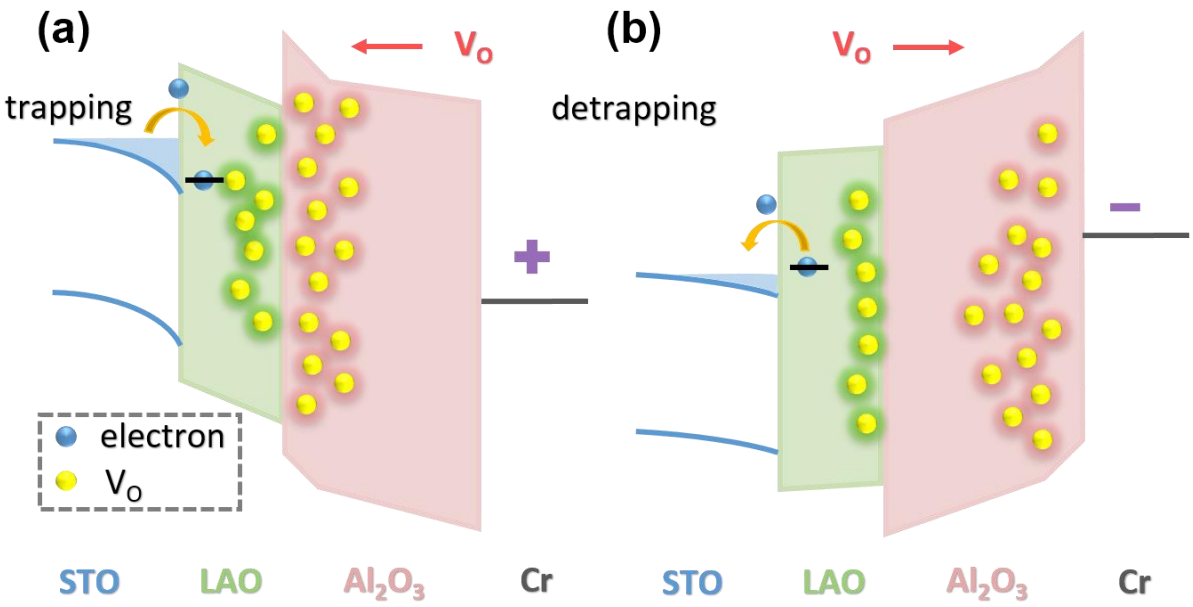


Figure 5.18. Band diagrams of oxide heterointerfaces in the proposed device under positive (a) and negative V_g (b). The band effect of positively charged V_Os are included.

As pointed out by Yu et al.[20], the built-in polar field across the LAO film triggers the spontaneous formation of V_Os in LAO (especially at the LAO surface). The electrons originated from the (positively) ionized V_O donors contribute to the formation of the 2DEG at the LAO/STO interface and compensate the polarization field, although part of the electrons may also be trapped by acceptor defects (including anti-site cations like Sr-on-La, and cation vacancies like La vacancies and Al vacancies) around the interface. Within this picture, we speculate that the external top gate bias applied on the LAO/STO device can lead to both ionic and electronic migration processes in the oxide heterostructure. Here the majority of the gate voltage is assumed to drop in the LAO film and in the Al_2O_3 dielectric layer, based on simple electrostatic analysis. On one side, the applied gate voltage can modify the spatial distribution of V_Os in these two oxide layers. The re-distribution of

the positively charged V_Os between the LAO layer and the Al_2O_3 layer will lead to changes in the device threshold voltage and subsequently the surface potential of the STO layer, resulting in modulated 2DEG conductivity. In this case, a positive gate voltage will push the V_Os towards the LAO/STO interface and make the threshold voltage more negative, i.e. enhanced conductance in this n-type transistor device. On the other side, the electric field can shift the fermi level and change the occupation state of the acceptor defects in the LAO film. When the electrons are captured (released) by these acceptors, the free carrier concentration of 2DEG will decrease (increase). In this case, a positive gate voltage will lead to the electron trapping in deep traps near the 2DEG and cause a decrease of the 2DEG conductance. Figure 5.18 schematically shows the two competing processes: electron trapping and V_O migration that lead to opposite changes in the threshold voltage and 2DEG conductivity of the device.

This proposed microscopic picture can consistently explain our experimental observations. For a virgin device, a large amount of oxygen vacancies exist around the LAO/Al_2O_3 interface and give rise to the initial high conductivity[37]. After the one-time forming process with high negative gate bias, part of the V_Os are extracted into the Al_2O_3 dielectric. At a positive bias, the V_O migration from the gate electrode to LAO induces a negative shift of V_{th} and the ON state with high conductivity, and vice versa (Figure 5.8). Considering the 2DEG carrier concentration (>10^{13} cm^{-2}, Figure 5.7), it is reasonable to ascribe the mobile charge responsible for ΔV_{th} (10^{13} cm^{-2}) to the oxygen vacancies after considering possible acceptor trapping. The V_O migration can be also supported indirectly by the switching dynamics of the device. The temperature (field) dependence of switching time shown in Figure 5.14 is very similar to the case of ionic transport, where increased

temperature helps ions overcome the hopping barrier while the external electric field lowers the hopping barrier and exponentially accelerates the ion migration process. Indeed, our extracted activation energy E_A and the characteristic length *l* are comparable to the reported V_O activation energy in amorphous Al_2O_3 (0.85eV) [39] and the lattice parameter of α-Al_2O_3 (4.875Å)[40].

Compared to the slow ionic migration of V_O, the electron trapping process is much faster. Consequently, at low temperatures (<220K), the electron trapping is more dominant, since V_Os lack necessary energy to overcome the hopping barrier for ionic movement. During electron trapping, the deep level traps become filled (unfilled) and causes I_{ds} reduction (gain) when positive (negative) pulses are applied. Once the V_O migration is activated at higher temperature (>240K), the electron trapping process is masked by the stronger ionic effect.

The effect of V_o migration on the transistor behavior is further modeled by numerical simulation. This simulation specifically focuses on the slow evolution of V_O re-distribution at room temperature (280K). The oxygen vacancy distributions of ON and OFF states are schematically illustrated in Figure 5.19 (a) and (b). The Poisson's equation and the ion drift-diffusion models are used to describe the charge redistribution and the V_O migration processes in a 1D mesh (Figure 5.19c). Figure 5.20 shows the entire simulation flow.

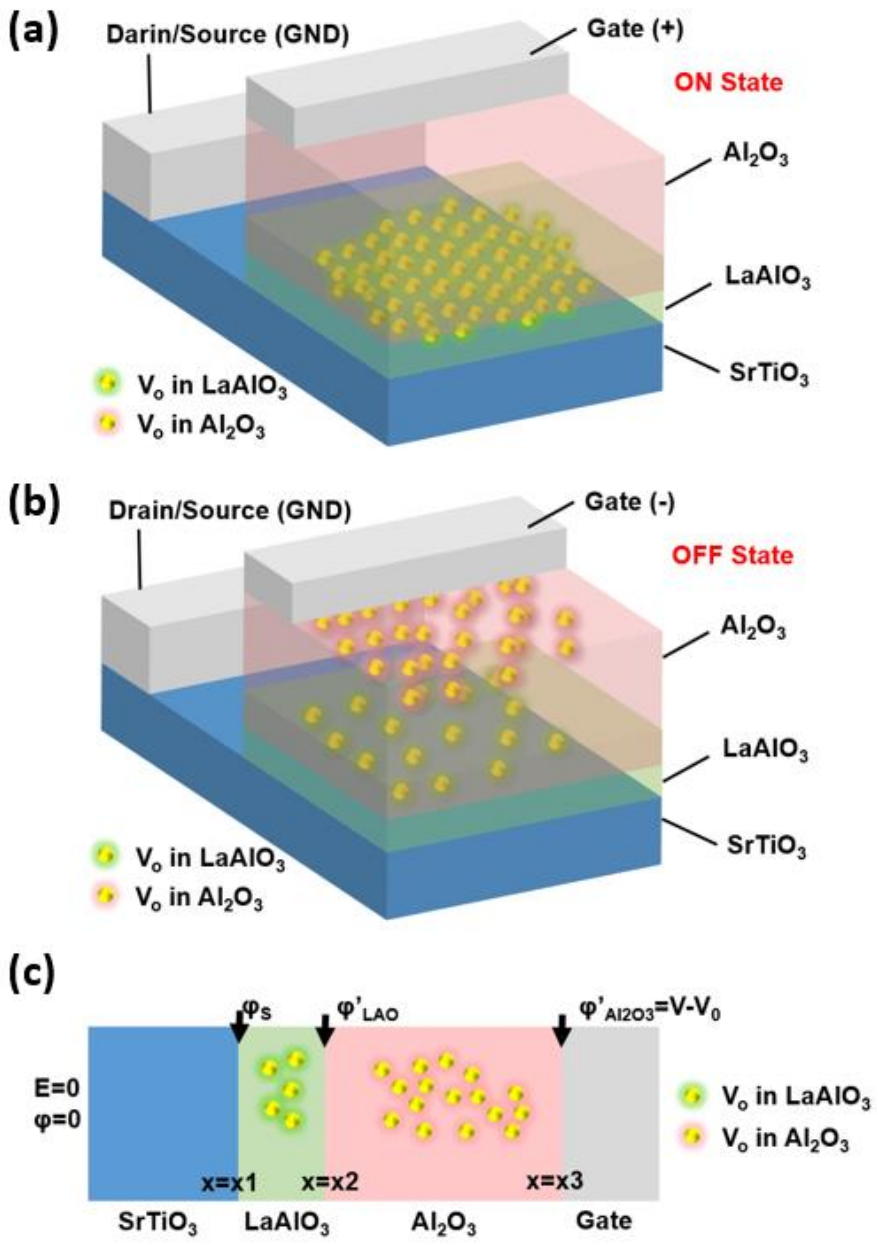


Figure 5.19. Schematic illustrations of (a) the ON state and (b) the OFF state in the $LaAlO_3/SrTiO_3$ transistor. (c) 1D mesh setup for the numerical simulation across the gate direction.

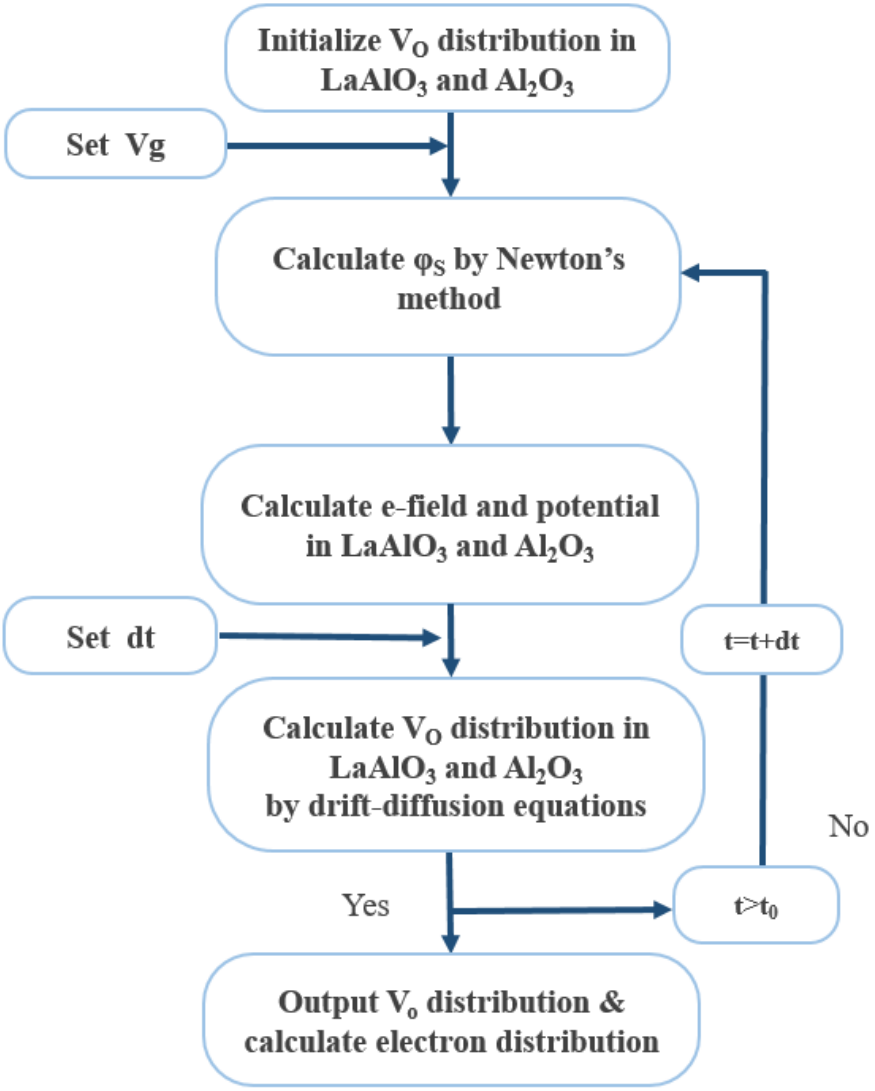


Figure 5.20. Workflow of the MATLAB based simulations.

In this model, prior to the application of any program/erase stress, an initialization step including forming and stabilization is performed to redistribute the oxygen vacancies originally located at the LAO surface. Upon a negative stress (-5V), the oxygen vacancies are attracted closer to the Al_2O_3 surface and reduces the electron concentration in the LAO/STO interface (Figure 5.21a). Once the negative erase bias is released, the oxygen vacancies gradually move away from the gate electrode due to spontaneous diffusion,

leading to a gradual recovery of the electron concentration (Figure 5.21b). With the application of a positive stress, oxygen vacancies are driven towards the LAO layer (Figure 5.21c) and leading to an enhancement of the 2DEG density and device conductance. A slow decay in n_e is also observed when the stress bias is removed and the oxygen vacancies revert to the equilibrium distribution (Figure 5.21d).

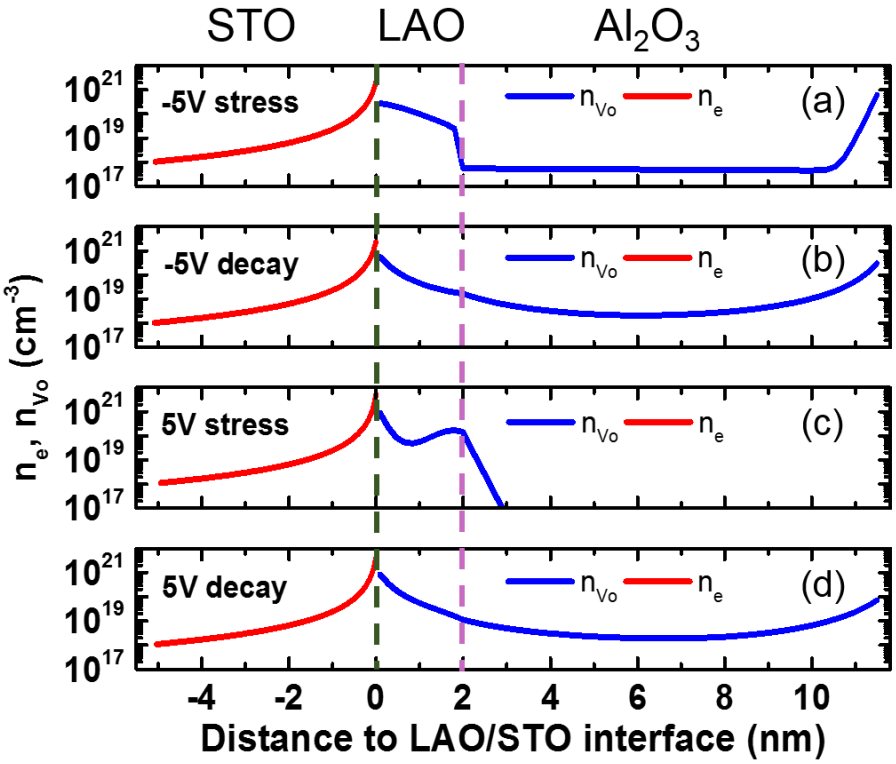


Figure 5.21. Simulated density of oxygen vacancies n_{Vo} (blue) in LAO/Al_2O_3 and electrons n_e (red) in STO at different moments: (a) after 300seconds' -5V stress; (b) 10000 seconds later after removing -5V stress; (c) after 300seconds' 5V stress; (b) 10000 seconds later after removing -5V stress.

Suppose that the 2DEG mobility remains roughly constant during the V_O migration, the current modulation observed of our LAO/STO device is proportional to the change of carrier density n_e. In this case, the normalized change of device current is simulated, shown

in Figure 5.22, along with the experimental observations. Good match with experiments can be observed (Figure 5.22).

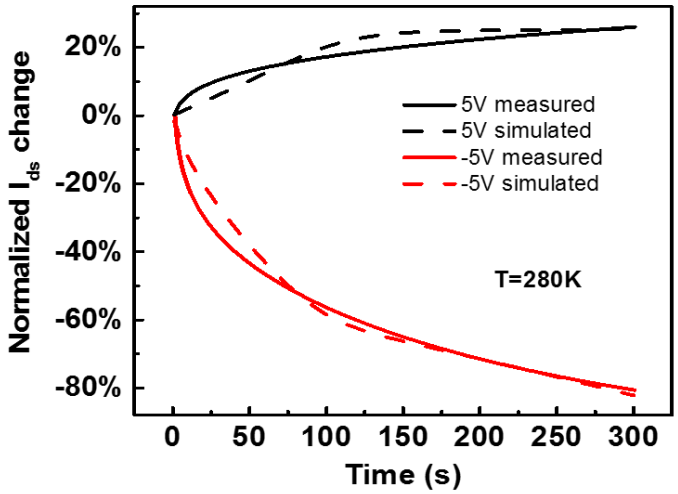


Figure 5.22. Measured (solid) and simulated (dash) I_{ds} change as a function of stress time at 280K.

5.5. Conclusions

In this study, we demonstrate a reconfigurable top-gate transistor device based on the LAO/STO heterojunction. Through electrical measurements and numerical simulations, the field induced switching of device conductivity is found to stem from the competing effects of ionic migration and electron trapping. Looking into the future, we expect that the proposed device can be continuously optimized in terms of performance and scalability and implemented for novel computing systems.

5.6. Reference

[1] Q. Huang, R. Huang, Y. Pan, S. Tan, and Y. Wang, "Resistive-Gate Field-Effect Transistor: A Novel Steep-Slope Device Based on a Metal-Insulator-Metal-Oxide Gate Stack," *IEEE Electron Device Lett.*, vol. 35, no. 8, pp. 877–879, 2014.

[2] W. Shim, J. Yao, and C. M. Lieber, "Programmable Resistive-Switch Nanowire Transistor Logic Circuits," *Nano Lett.*, vol. 14, no. 9, pp. 5430–5436, 2014.

[3] Y. Yang, B. Chen, and W. D. Lu, "Memristive Physically Evolving Networks Enabling the Emulation of Heterosynaptic Plasticity," *Adv. Mater.*, vol. 27, no. 47, pp. 7720–7727, 2015.

[4] A. Ohtomo and H. Y. Hwang, "A high-mobility electron gas at the LAO/STO heterointerface.," *Nature*, vol. 427, no. 6973, pp. 423–426, 2004.

[5] N. Reyren, S. Thiel, A. D. Caviglia, L. F. Kourkoutis, G. Hammerl, C. Richter, C. W. Schneider, T. Kopp, A.-S. Ruetschi, D. Jaccard, M. Gabay, D. A. Muller, J.-M. Triscone, and J. Mannhart, "Superconducting Interfaces Between Insulating Oxides," *Science*, vol. 317, no. 5842, pp. 1196–1199, 2007.

[6] A. Brinkman, M. Huijben, M. van Zalk, J. Huijben, U. Zeitler, J. C. Maan, W. G. van der Wiel, G. Rijnders, D. H. a Blank, and H. Hilgenkamp, "Magnetic effects at the interface between non-magnetic oxides.," *Nat. Mater.*, vol. 6, no. 7, pp. 493–496, 2007.

[7] J. A. Bert, B. Kalisky, C. Bell, M. Kim, Y. Hikita, H. Y. Hwang, and K. a. Moler, "Direct imaging of the coexistence of ferromagnetism and superconductivity at the $LaAlO_3/SrTiO_3$ interface," *Nat. Phys.*, vol. 7, no. 10, pp. 767–771, 2011.

[8] A. D. Caviglia, M. Gabay, S. Gariglio, N. Reyren, C. Cancellieri, and J.-M. Triscone, "Tunable Rashba Spin-Orbit Interaction at Oxide Interfaces," *Phys. Rev. Lett.*, vol. 104, no. 12, p. 126803, 2010.

[9] A. Tebano, E. Fabbri, D. Pergolesi, G. Balestrino, and E. Traversa, "Room-Temperature Giant Persistent Photoconductivity in $SrTiO_3/LaAlO_3$ Heterostructures," *ACS Nano*, vol. 6, no. 2, pp. 1278–1283, 2012.

[10] M. C. Tarun, F. A. Selim, and M. D. McCluskey, "Persistent Photoconductivity in Strontium Titanate," *Phys. Rev. Lett.*, vol. 111, no. 18, p. 187403, Oct. 2013.

[11] G. Berner, A. Müller, F. Pfaff, J. Walde, C. Richter, J. Mannhart, S. Thiess, A. Gloskovskii, W. Drube, M. Sing, and R. Claessen, "Band alignment in $LaAlO_3/SrTiO_3$ oxide heterostructures inferred from hard x-ray photoelectron spectroscopy," *Phys. Rev. B*, vol. 88, no. 11, p. 115111, 2013.

[12] S. Thiel, "Tunable Quasi-Two-Dimensional Electron Gases in Oxide Heterostructures," *Science*, vol. 313, no. 5795, pp. 1942–1945, 2006.

[13] N. Nakagawa, H. Y. Hwang, and D. A. Muller, "Why some interfaces cannot be sharp," *Nat. Mater.*, vol. 5, no. 3, pp. 204–209, 2006.

[14] W. Siemons, G. Koster, H. Yamamoto, W. A. Harrison, G. Lucovsky, T. H. Geballe, D. H. A. Blank, and M. R. Beasley, "Origin of charge density at $LaAlO_3$ on $SrTiO_3$ heterointerfaces: possibility of intrinsic doping.," *Phys. Rev. Lett.*, vol. 98, no. 19, p. 196802, 2007.

[15] C. Cen, S. Thiel, G. Hammerl, C. W. Schneider, K. E. Andersen, C. S. Hellberg, J. Mannhart, and J. Levy, "Nanoscale control of an interfacial metal–insulator transition at room temperature," *Nat. Mater.*, vol. 7, no. 4, pp. 298–302, 2008.

[16] N. C. Bristowe, P. B. Littlewood, and E. Artacho, "Surface defects and conduction in polar oxide heterostructures," *Phys. Rev. B*, vol. 83, no. 20, p. 205405, 2011.

[17] A. S. Kalabukhov, Y. A. Boikov, I. T. Serenkov, V. I. Sakharov, V. N. Popok, R. Gunnarsson, J. Börjesson, N. Ljustina, E. Olsson, D. Winkler, and T. Claeson, "Cationic disorder and phase segregation in $LaAlO_3/SrTiO_3$ heterointerfaces evidenced by medium-energy ion spectroscopy.," *Phys. Rev. Lett.*, vol. 103, no. 14, p. 146101, 2009.

[18] P. R. Willmott, S. A. Pauli, R. Herger, C. M. Schlepütz, D. Martoccia, B. D. Patterson, B. Delley, R. Clarke, D. Kumah, C. Cionca, and Y. Yacoby, "Structural basis for the conducting interface between $LaAlO_3$ and $SrTiO_3$.," *Phys. Rev. Lett.*, vol. 99, no. 15, p. 155502, 2007.

[19] V. Vonk, J. Huijben, D. Kukuruznyak, A. Stierle, H. Hilgenkamp, A. Brinkman, and S. Harkema, "Polar-discontinuity-retaining A-site intermixing and vacancies at $SrTiO_3/LaAlO_3$ interfaces," *Phys. Rev. B*, vol. 85, no. 4, p. 045401, 2012.

[20] L. Yu and A. Zunger, "A polarity-induced defect mechanism for conductivity and magnetism at polar–nonpolar oxide interfaces," *Nat. Commun.*, vol. 5, p. 5118, 2014.

[21] J. A. Sulpizio, S. Ilani, P. Irvin, and J. Levy, "Nanoscale Phenomena in Oxide Heterostructures," *Annu. Rev. Mater. Res.*, vol. 44, no. 1, pp. 117–149, 2014.

[22] C. W. Bark, P. Sharma, Y. Wang, S. H. Baek, S. Lee, S. Ryu, C. M. Folkman, T. R. Paudel, A. Kumar, S. V. Kalinin, A. Sokolov, E. Y. Tsymbal, M. S. Rzchowski, A. Gruverman, and C. B. Eom, "Switchable Induced Polarization in $LaAlO_3$ /$SrTiO_3$ Heterostructures," *Nano Lett.*, vol. 12, no. 4, pp. 1765–1771, 2012.

[23] M. Hosoda, Y. Hikita, H. Y. Hwang, and C. Bell, "Transistor operation and mobility enhancement in top-gated $LaAlO_3/SrTiO_3$ heterostructures," *Appl. Phys. Lett.*, vol. 103, no. 10, p. 103507, 2013.

[24] C. Bell, S. Harashima, Y. Kozuka, M. Kim, B. G. Kim, Y. Hikita, and H. Y. Hwang, "Dominant mobility modulation by the electric field effect at the $LaAlO_3/SrTiO_3$ interface.," *Phys. Rev. Lett.*, vol. 103, no. 22, p. 226802, 2009.

[25] D. V. Christensen, F. Trier, Y. Z. Chen, A. Smith, J. Nygård, and N. Pryds, "Controlling interfacial states in amorphous/crystalline $LaAlO_3/SrTiO_3$ heterostructures by electric fields," *Appl. Phys. Lett.*, vol. 102, no. 2, p. 021602, 2013.

[26] L. Dong, Y. Q. Liu, M. Xu, Y. Q. Wu, R. Colby, E. a. Stach, R. Droopad, R. G. Gordon, and P. D. Ye, "Atomic-layer-deposited $LaAlO_3/SrTiO_3$ all oxide field-effect transistors," in *2010 International Electron Devices Meeting*, pp. 26.4.1–26.4.4, 2010.

[27] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*. Cambridge university press, 2013.

[28] C. Cazorla and M. Stengel, "First-principles modeling of $Pt/LaAlO_3/SrTiO_3$ capacitors under an external bias potential," *Phys. Rev. B*, vol. 85, no. 7, p. 075426, 2012.

[29] J. Krupka, R. G. Geyer, M. Kuhn, and J. H. Hinken, "Dielectric properties of single crystals of Al_2O_3, $LaAlO_3$, $NdGaO_3$, $SrTiO_3$, and MgO at cryogenic temperatures," *IEEE Trans. Microw. Theory Tech.*, vol. 42, no. 10, pp. 1886–1890, 1994.

[30] 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," *Front. Neurosci.*, vol. 5, no. 26, 2011.

[31] Y. Xie, C. Bell, T. Yajima, Y. Hikita, and H. Y. Hwang, "Charge writing at the $LaAlO_3/SrTiO_3$ surface," *Nano Lett.*, vol. 10, no. 7, pp. 2588–2591, 2010.

[32] Y. Xie, C. Bell, Y. Hikita, and H. Y. Hwang, "Tuning the Electron Gas at an Oxide Heterointerface via Free Surface Charges," *Adv. Mater.*, vol. 23, no. 15, pp. 1744–1747, 2011.

[33] F. Bi, D. F. Bogorin, C. Cen, C. W. Bark, J.-W. Park, C.-B. Eom, and J. Levy, "'Water-cycle' mechanism for writing and erasing nanostructures at the $LaAlO_3/SrTiO_3$ interface," *Appl. Phys. Lett.*, vol. 97, no. 17, p. 173110, 2010.

[34] T. Leisegang, H. Stöcker, A. A. Levin, T. Weissbach, M. Zschornak, E. Gutmann, K. Rickers, S. Gemming, and D. C. Meyer, "Switching Ti valence in $SrTiO_3$ by a dc electric field.," *Phys. Rev. Lett.*, vol. 102, no. 8, p. 087601, 2009.

[35] K. Szot, W. Speier, G. Bihlmayer, and R. Waser, "Switching the electrical resistance of individual dislocations in single-crystalline $SrTiO_3$," *Nat. Mater.*, vol. 5, no. 4, pp. 312–320, 2006.

[36] H. F. Tian, Y. G. Zhao, X. L. Jiang, J. P. Shi, H. J. Zhang, and J. R. Sun, "Resistance switching effect in $LaAlO_3$/Nb-doped $SrTiO_3$ heterostructure," *Appl. Phys. A*, vol. 102, no. 4, pp. 939–942, 2011.

[37] C. Cantoni, J. Gazquez, F. Miletto Granozio, M. P. Oxley, M. Varela, A. R. Lupini, S. J. Pennycook, C. Aruta, U. S. di Uccio, P. Perna, and D. Maccariello, "Electron Transfer and Ionic Displacements at the Origin of the 2D Electron Gas at the LAO/STO Interface: Direct Measurements with Atomic-Column Spatial Resolution," *Adv. Mater.*, vol. 24, no. 29, pp. 3952–3957, 2012.

[38] S. K. Kim, S.-I. Kim, H. Lim, D. S. Jeong, B. Kwon, S.-H. Baek, and J.-S. Kim, "Electric-field-induced Shift in the Threshold Voltage in $LaAlO_3/SrTiO_3$ Heterostructures," *Sci. Rep.*, vol. 5, p. 8023, 2015.

[39] K. Sankaran, L. Goux, S. Clima, M. Mees, J. a. Kittl, M. Jurczak, L. Altimime, G.-M. Rignanese, and G. Pourtois, "Modeling of Copper Diffusion in Amorphous Aluminum Oxide in CBRAM Memory Stack," *ECS Trans.*, vol. 45, no. 3, pp. 317–330, 2012.

[40] M. Lucht, M. Lerche, H.-C. Wille, Y. V Shvyd'ko, H. D. Rüter, E. Gerdau, and P. Becker, "Precise measurement of the lattice parameters of α-Al_2O_3 in the temperature range 4.5–250 K using the Mössbauer wavelength standard," *J. Appl. Crystallogr.*, vol. 36, no. 4, pp. 1075–1081, 2003.

Chapter 6

In-Memory Computing Based on Crossbar RRAM Arrays

6.1. Introduction

The conventional von Neumann architecture is based on separate computing units and memory units[1]. For the past few decades, this design principle has remained effective in maximizing and balancing the performance of computation and storage. On one hand, the processor performance has dramatically benefited from the transistor scaling driven by the Moore's Law. On the other hand, the sophisticated design of memory hierarchy (e.g. cache system) efficiently hides the lag from read/write operations and alleviates the "memory wall" issues[2]. However, as the CMOS technology marches into the 1x nm era, the free ride to exploit computing power from smaller devices is ending. In the meantime, with the exploding data volume as a result of "big data" and Internet of Things, today's computer systems severely suffer from the von Neumann bottleneck where significant energy penalty and wire delay are caused by the frequent data flow and the limited bandwidth between processors and memory[3]. To meet the requirement of data-centric applications, the concepts of near-data and in-memory computing have been proposed[4]–[8]. Instead of moving data towards computing units and writing back, distributing the computation closer to memory can reduce the energy and latency from unnecessary data movement.

The question then arises how the computation can be implemented on the storage side. Some prior works proposed straightforward solutions by incorporating logic components

in the memory dies[6]–[9]. However, the processor performance may be sacrificed by the relatively slow memory technology, and the available memory space allocated to each processor is limited by the chip area. With the rapid progress of through silicon via (TSV) technology[10], an improved version of in-memory computing was demonstrated by vertically stacking DRAM layers on top of logic layer[11].

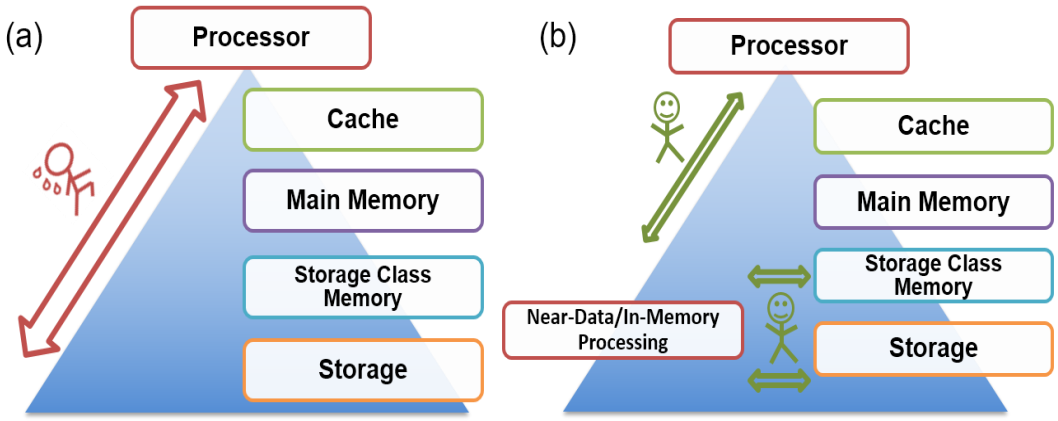


Figure 6.1. (a) The von Neumann Bottleneck results from the frequent data flow across the entire memory hierarchy. (b) In-memory/near-data processing may reduce the data movement and alleviate the related issues concerning energy, bandwidth and delay.

Compared to DRAM, RRAM offers unique advantages, including non-volatility, low power consumption, high density and convenient 3D stacking. More importantly, RRAM incorporates the capabilities of both computation and storage. As an emerging two-terminal device, RRAM is inherently compatible to the crossbar configuration which potentially allows high parallelism. Several recent works have realized the logic functionality using RRAM devices on the individual cell level[12], [13]; while system-level analysis and demonstration are still lacking[14]. Here we expect to explore the feasibility of in-memory computing based on RRAM crossbar arrays.

In this chapter, the basic scheme of proposed computing architecture and its operation

protocol will be firstly discussed in Section 6.2. To verify the architecture concept, a prototype circuit is built to demonstrate NOR logic and a parallel 1-bit fill adder (FA) with RRAM crossbar arrays by experiment and simulation (Section 6.3). A 4-bit multiplier (MULT) is further obtained by preprogramming 2-bit MULT and 2-bit FA. Section 6.4 briefly discusses the performance of this novel architecture.

6.2. Computing Architecture and Design Principle

The proposed computing architecture is schematically illustrated in Figure 6.2. This system consists of RRAM crossbar arrays, microcontroller and related peripheral circuits (e.g. buffers and decoders). The essential point of this design is to parse the complex functions into fundamental operations where all input values will be stored in the systems with the output values computed and recorded simultaneously. As a result, for any previous input, the result of a certain function can be directly read out without conducting new computations.

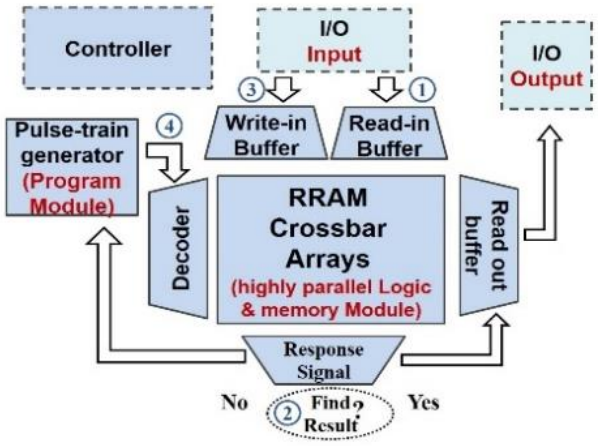


Figure 6.2. Schematic block diagram of the proposed architecture. Locally stored logic and parallel programming and read-out are developed for efficient computation. Reproduced from [16].

To be specific, the RRAM crossbar arrays in this system record all the input and keep updating while new entries are received. Managed by the controller, the results of the target function can be generated by pulse operations and restored within the RRAM crossbar array on the fly. For an input-output pair that has been already stored in the system, the input signals will be fed from I/O into the read-in buffer, and the corresponding output can be found from the array in a parallel fashion and fed to the read-out buffer. If the input-output pair has yet been created, the new function (or input combination) encountered needs to be written into the array. In this case, a series of programming pulses will be generated by the controller to modify the crossbar array content which represents the intermediate and final results of target function. With this process repeated, each stored input-output pair relation can be reused for future computing operations.

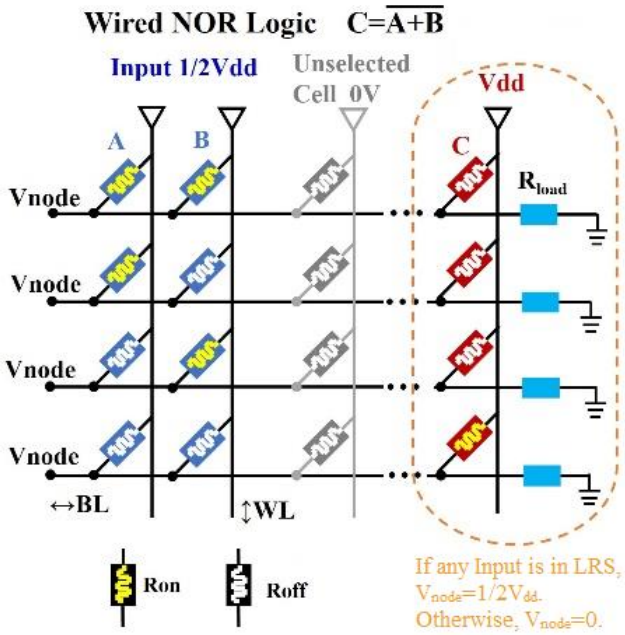


Figure 6.3. Schematic of the wired-NOR logic implementation using the half V_{dd} scheme. Reproduced from [16].

During the array programing, the target function is decomposed into basic NOR logic since it can readily adapt to the RRAM crossbar array and support arbitrary Boolean logic

elements. As shown in Figure 6.3, the basic NOR logic can be programmed in the crossbar by applying V_{dd} and $V_{dd}/2$ on the word-lines (WLs) of output and input cells, respectively. Under such bias scheme, the output cells will be programmed into either HRS (Logic 0) or LRS (Logic 1) depending on the potential of bit-lines (BLs), while the resistance states of input cells remain unchanged.

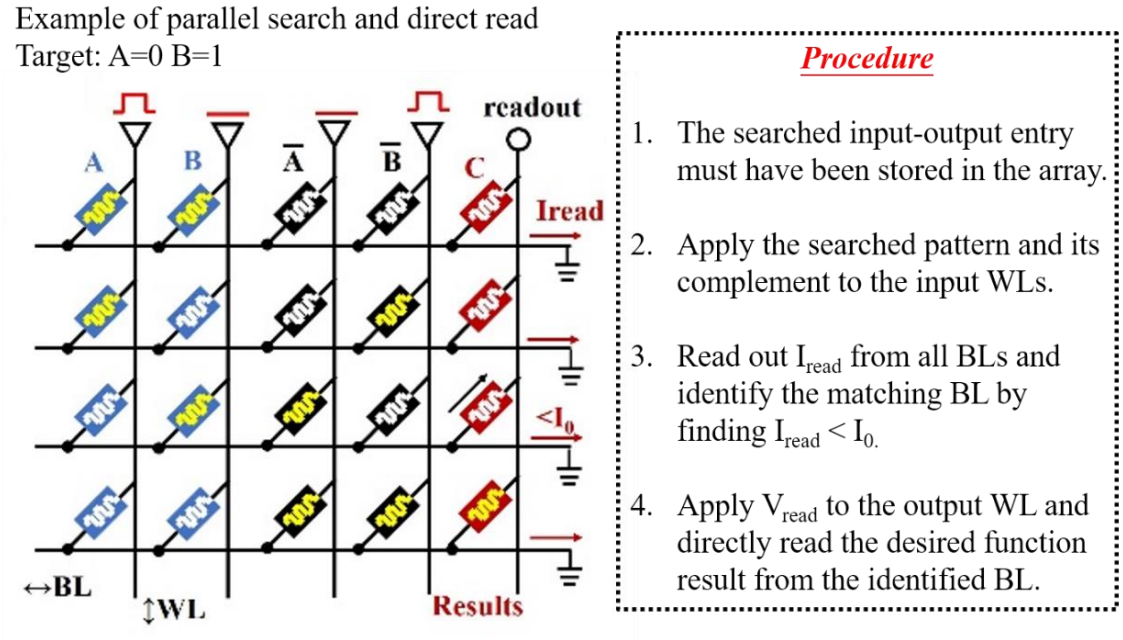


Figure 6.4. Proposed procedures for parallel search of input pattern and direct read-out of function result. Only one BL whose input cell states match the search pattern will produce a current smaller than the predefined threshold current I_0. Reproduced from [16].

The direct parallel search for a programed logic element (an input-output entry for the given function, corresponding to one BL in the crossbar array) can be demonstrated at the constant time cost without involving additional decoders. As illustrated in Figure 6.4, both the input combination and its complement are applied to the corresponding WLs, and the currents through all the BLs will be read by the sense amplifiers (SAs) for the search. Only the BL whose input cell states match the searched values will produce a current significantly smaller than the predefined threshold current I_0 (decided by the array size and

the LRS/HRS of RRAM). Once this low current is found, the searched input/output combination can be located. In the next step, the result of desired function can be directly read out from the identified BL by applying a V_{read} to the output WL.

Generally, by combining stored logic (analogous to local look-up tables) and direct parallel read-out with the non-volatile capabilities of RRAM arrays, the separation between computation and storage is eliminated in the proposed in-memory computing architecture. The true co-located memory and logic in our systems will avoid the von Neumann bottleneck and offer fast operation speed, low power and ultra-high function density.

6.3. Proof-of-Concept Demonstration

6.3.1. Device Characterization and Array Fabrication

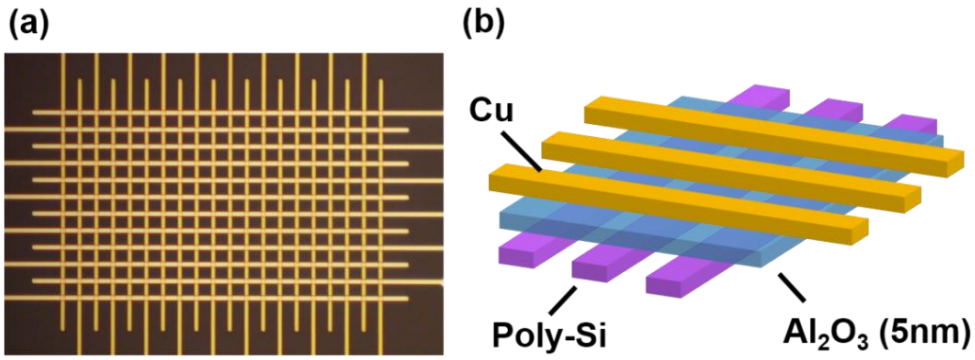


Figure 6.5. (a) Optical microscopic image of the as-fabricated RRAM array, (b) Structure schematic of single device and crossbar array. Reproduced from [16].

To reduce the system power consumption and allow as large arrays as possible, the $Cu/Al_2O_3/Poly\text{-}Si$ based CBRAMs discussed in Chapter 4 with low operation current (<100nA), high on/off ratio (>10^3) are used to constitute the crossbar arrays for the in-memory computing architecture (Figure 6.5-6.6). The detailed fabrication flow is discussed in Section 4.2. The experimental I-V curves can be well fitted using a compact dynamic

SPICE model [15] (Fig. 6.6a), which enables realistic simulations of large-scale computing system. For reliable array application, the device-to-device uniformity is also verified (Figure 6.6b).

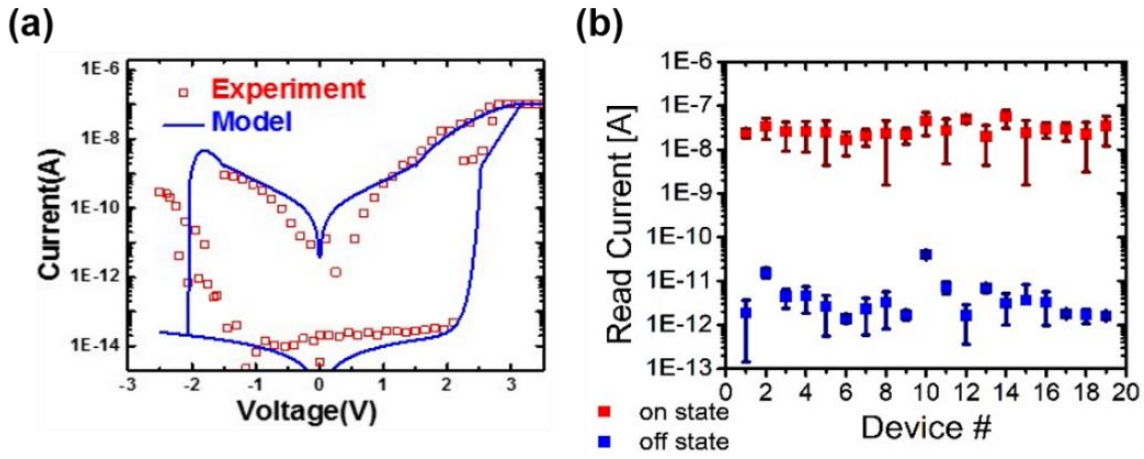


Figure 6.6. (a) Measured and simulated device I-V. The RRAMs exhibit very low (< 100nA) programming current. (b) HRS and LRS resistance distribution measured from 20 randomly chosen devices. Each device is DC cycled for 5 times. Excellent uniformity and high on/off are achieved. Reproduced from [16].

6.3.2. Prototype Circuit Setup

A prototype circuit is setup as follows to verify the in-memory computing functionality. After the microfabrication, the RRAM crossbar array is wire-bonded onto the chip carrier (Spectrum LCC 8423) and tested on a special-purpose PCB board (Figure 6.7). Four digital-to-analog converters (DACs) on board provide 0-5V voltages independently. 16 matrix switches are incorporated to support up to 32x32 routing. The on-board current is collected using a 12-bit analog-to-digital converters (ADCs). The entire system is controlled with a Spartan 6 XC6SLX9 FPGA. High-level programming tools based on Python and C++ are implemented for board operation. The circuit block diagram is illustrated as Figure 6.8.

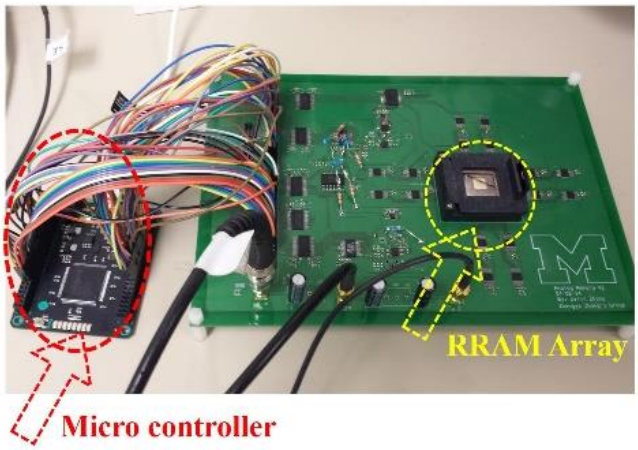


Figure 6.7. Circuit demonstration of the in-memory computing based on RRAM array. Reproduced from [16].

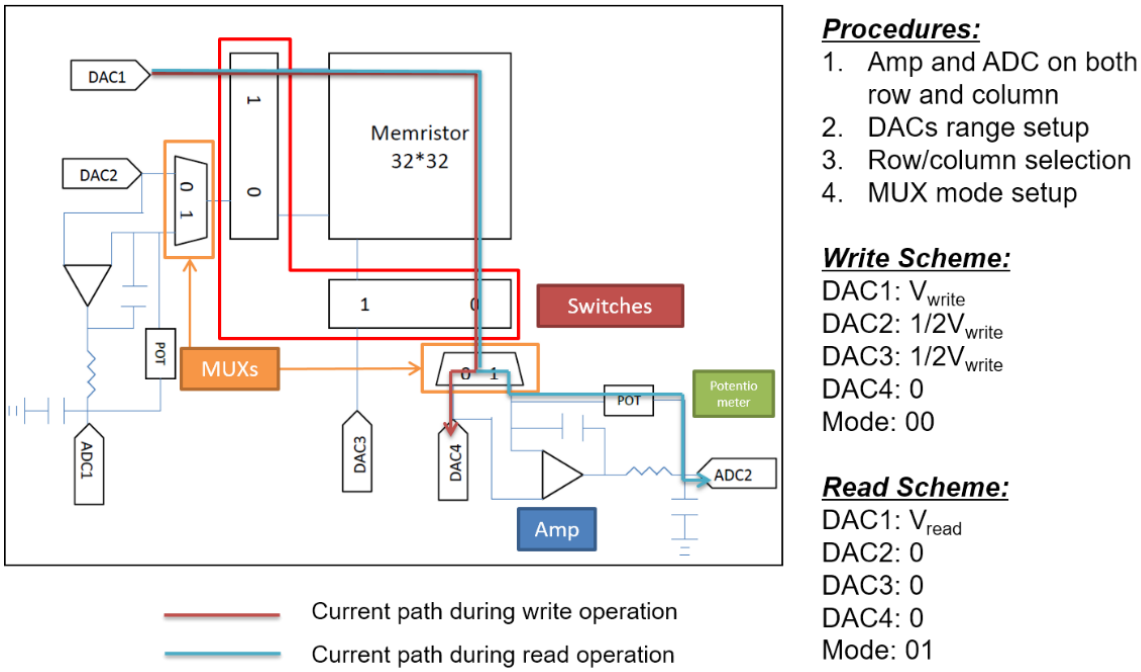


Figure 6.8. Block diagram of the designed PCB and its working setup.

6.3.3. Implementation of NOR Logic and 1-bit Full Adder

As the most fundamental operation in the computing architecture, NOR is first realized experimentally in three neighboring cells in the crossbar array (Figure 6.9), corresponding

to the input cells (A, B) and the output cell (C) in Figure 6.3. 4V V_{dd} is applied on the WL of C; while ½ V_{dd} (2V) is applied on the WLs of A and B. The output value of cell C after program is then determined by the resistance states of the input cells A B, following the NOR logic $C = \overline{A + B}$.

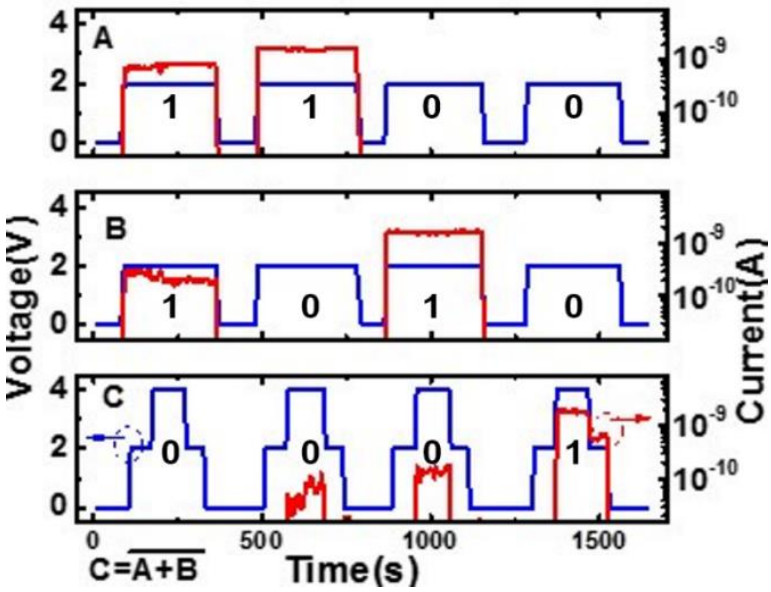


Figure 6.9. Measured current response of NOR logic while applying ½V_{dd} on Inputs A&B and V_{dd} on Output C. The test method is based on the approach described in Figure 6.3. Reproduced from [16].

Further, a 1-bit full adder is demonstrated as a proof of principle of the proposed computing architecture. Figure 6.10 shows the simulated programing process of the 1-bit FA in an 8x16 crossbar array. The FA function is parsed into NOR operations (Figure 6.10a) and stored in the array using the pulse trains shown in Figure 6.10b. The cell values of RRAM array after the programming step are shown in Figure 6.10c. After writing the function, the desired output can be directly read out for any given input during the computing stage.

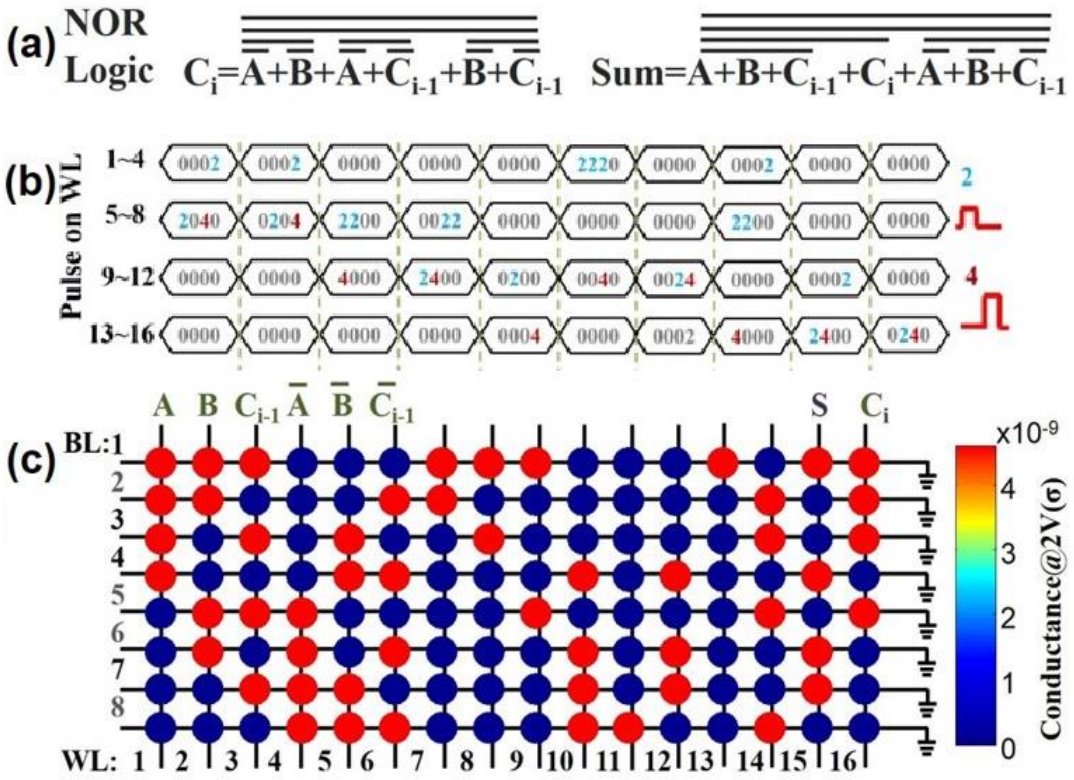


Figure 6.10. Proposed programing process of a 1-bit full adder. (a) Decomposed NOR logics for the 1-bit FA. (b) Programming pulses on the WLs. (c) Resistance mapping after programming. Reproduced from [16].

Figure 6.11 shows the simulated current on BL/WLs during the read operation using the proposed data identification protocol. During this stage, an identification code designed for each input value and its complement is first applied to the input WLs (WL1~6). Taking input ABC=101 for instance, the identification code 020202 will be applied on the first 6 WLs. Here "2" stands for a read voltage of 2V and "0" means grounded. On one hand, if the identification code matches the input pattern (and its complement) stored in a BL, the current in this BL will be low since read pulses will point to devices at HRS and all devices at LRS are grounded. A low current $I_{match}=n*V_{read}/R_{HRS}$ is obtained for the matching BL, where n represents the number of HRS leakage paths biased at V_{read}. On the other hand,

any other BLs will output much higher current, since at least one input cell will mismatch the identification code and produce a high through-current of $I_{non\text{-}match}=V_{read}/R_{LRS}$. As a result, the matching BL can be directly identified by a simple comparator without additional address decoding schemes. As shown in Figure 6.11, BLs 1, 3 4, 8 are correctly identified during four searches (input values 111, 101, 100, 000). The target output values of C_i and S are obtained by direct read-out as 10, 01, 11 and 00 respectively.

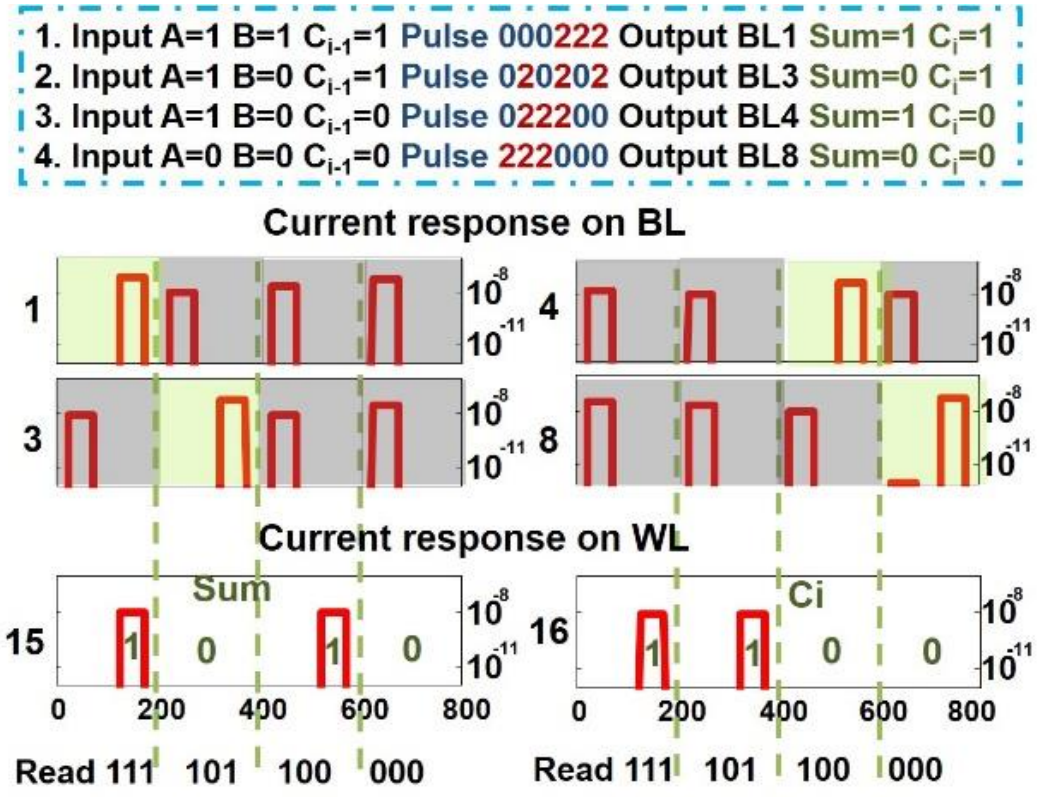
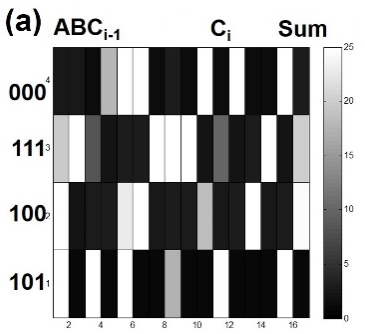


Figure 6.11. Simulated read-out process of the 1-bit FA. The function results can be correctly located (middle panel, highlighted by the green windows) and read out (bottom panel). Reproduced from [16].



(b) $I_{CAM}<I_0$ $I_{CAM}>I_0$

Input signal	CAM Current on BL				Readout results Current on WL			
	BL (111)	BL (101)	BL (100)	BL (000)	Sum		Ci	
111	8.03	30.2	35.6	53.1	20	"1"	9.78	"1"
101	27.3	3.42	51.1	27.3	1.17	"0"	25	"1"
100	50.9	30.3	8.23	30.2	24.6	"1"	2.36	"0"
000	204	79.2	21.7	6.52	2.99	"0"	1.52	"0"

Figure 6.12. (a) Measured resistance mapping after writing the functions into the RRAM array for 4 input cases. (b) Measured CAM current from the BLs during data identification process (left) and the output results directly read from the WLs (right). Reproduced from [16].

Parallel function write and direct output read are then experimentally performed on the fabricated RRAM crossbar array. The resistance mapping in Figure 6.12a confirms the successful write of desired logics. Figure 6.12b verifies the response current correctly read from the BL/WLs for identifying input and reading output.

6.3.4. Implementation of 4-bit Multiplier

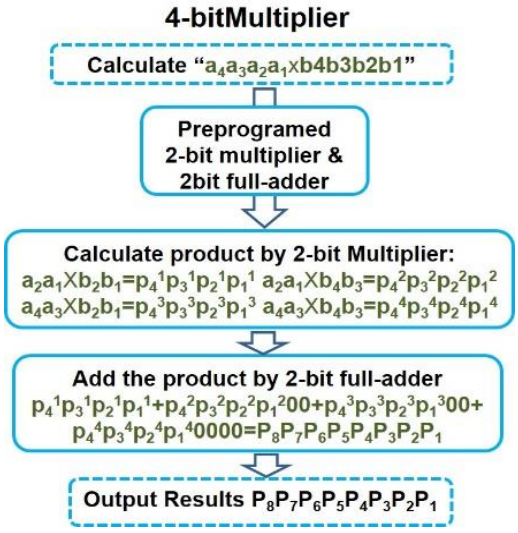


Figure 6.13. Building a 4-bit MULT from a 2-bit MULT and a 2-bit FA. Reproduced from [16].

A 4-bit multiplier (MULT) is further developed to show that multiple programed arrays with different stored logic functions can be constructed and pipelined to complete more complex tasks. Based on the same working protocol, the 4-bit MULT combines a 2-bit MULT with a 2-bit FA (Figure 6.13). The 2-bit MULT and the 2-bit FA here are pre-programmed using the same parallelism NOR logic discussed earlier. Their resistance mappings are illustrated in Figure 6.14.

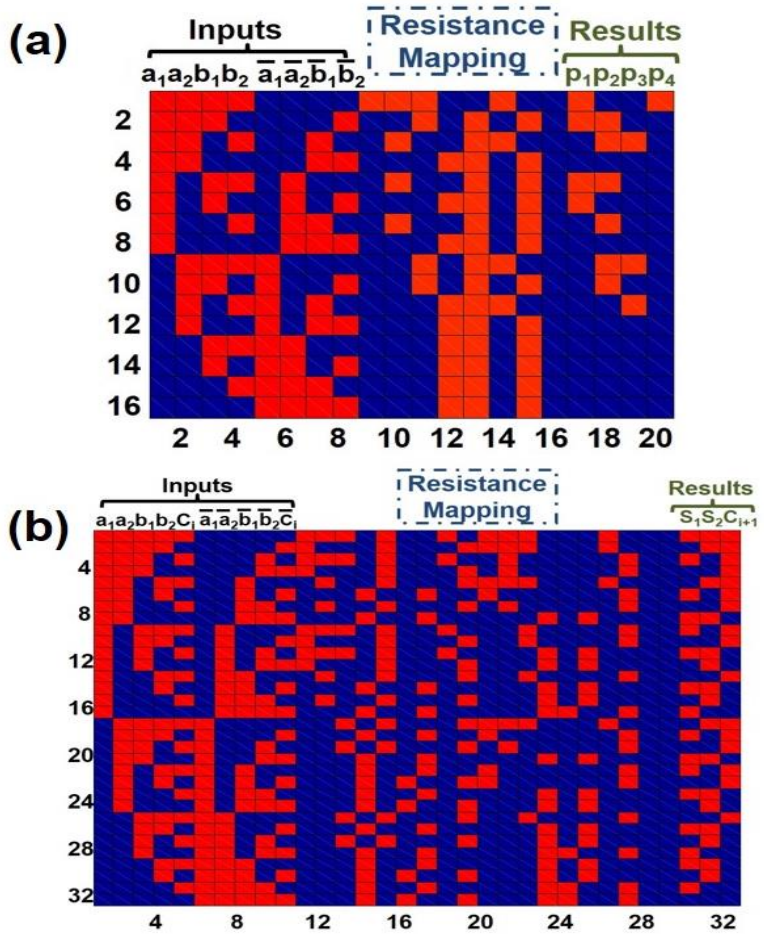


Figure 6.14. Resistance mappings of the pre-programmed 2-bit MULT (a) and 2-bit FA (b). Reproduced from [16].

The process of solving 1011x1101 is simulated as an example. First, the calculation results of 11x01 11x11 10x01 10x11 are obtained by reading out results from the 2-bit MULT. These products are then shifted and serially added by the 2-bit FA, eventually

leading to the results of 4-bit multiplication, as shown in Figure 6.15. Similarly, complex functions can be parsed into several sub-functions stored in different RRAM arrays. However, the required array size depends on the complexity of each sub-function. Therefore, proper tradeoff needs to be handled for speed (operation steps) and area cost (array size) in this crossbar array based computing architecture while complex functions are involved.

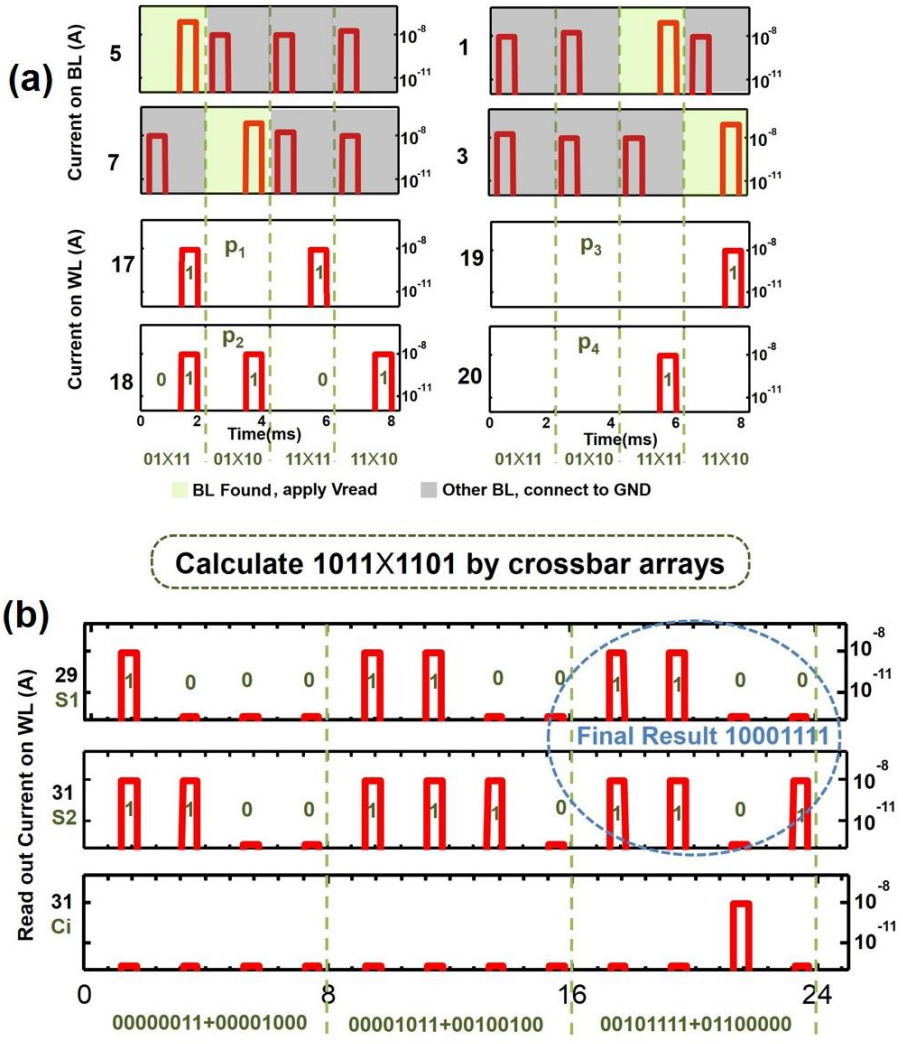


Figure 6.15. Simulated read-out process of the 4-bit MULT: (a) Current response during the read of the 2-bit multiplication. (b) Current response during the read of the 2-bit summation. Reproduced from [16].

6.4. Performance Analysis

Critical circuit performance benchmarks such as power consumption and speed are analyzed through numerical simulations based on the dynamic model with realistic device parameters. Figure 6.16 shows the average energy cost (per operation) during the programing and read-out stages. The identification method employed in this novel architecture significantly reduces the power and a low energy consumption of 3.67nJ is obtained for completing a 4-bit multiplication operation. Since many devices will be programed or read at the same time, the maximum current through a BL/WL in the worst case scales with the array size. Therefore, low operation current of RRAM device is particularly important for the feasible circuit realization. Figure 6.17 highlights the essential role of RRAM R_{on}/R_{off} ratio in this in-memory computing scheme. Neglecting the parasitic resistance from electrodes, the maximum array size that ensures reliable computing operations is found to be proportional to the R_{on}/R_{off} ratio. As a result, RRAM cells that offer low programming/read current, high R_{on}/R_{off} ratio and fast operation speed are strongly desired for the proposed computing architecture.

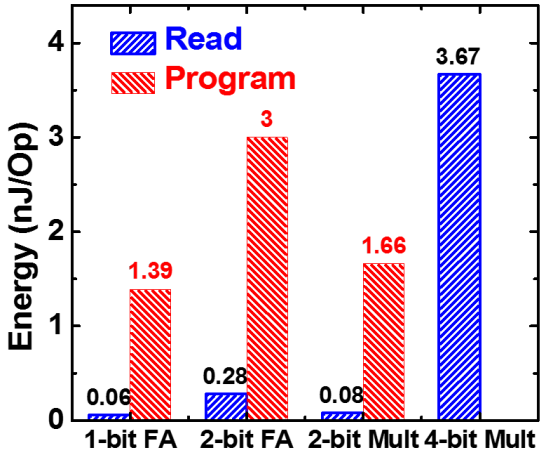


Figure 6.16. Average energy cost of the proposed FA/MULT implementation. Reproduced from [16].

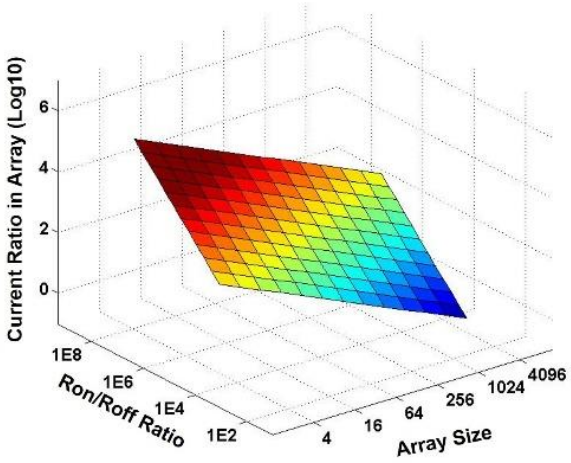


Figure 6.17. Read-out current ratio ("1" vs "0") as a function of R_{on}/R_{off} ratio and array size. Reproduced from [16].

6.5. Conclusion

A novel RRAM crossbar-based computing architecture is proposed and demonstrated for low current, highly parallel and reconfigurable computing. Key operations and design methods are introduced for prototype circuit realization. 1-bit FA and 4-bit MULT are implemented by experiment and simulation to verify the proposed architecture.

6.6. References

[1] J. von Neumann, "First draft of a report on the EDVAC," *IEEE Ann. Hist. Comput.*, vol. 15, no. 4, pp. 27–75, 1993.

[2] R. Nair, "Evolution of Memory Architecture," *Proc. IEEE*, vol. 103, no. 8, pp. 1331–1345, 2015.

[3] W. a. Wulf and S. a. McKee, "Hitting the memory wall," *ACM SIGARCH Comput. Archit. News*, vol. 23, no. 1, pp. 20–24, 1995.

[4] P. Francisco, "The Netezza data appliance architecture: A platform for high performance data warehousing and analytics," *IBM Redbooks*, 2011.

[5] R. Balasubramonian, J. Chang, T. Manning, J. H. Moreno, R. Murphy, R. Nair, and S. Swanson, "Near-Data Processing: Insights from a MICRO-46 Workshop," *IEEE Micro*, vol. 34, no. 4, pp. 36–42, 2014.

[6] J. Draper, C. W. Kang, I. Kim, G. Daglikoca, J. Chame, M. Hall, C. Steele, T. Barrett, J. LaCoss, J. Granacki, J. Shin, and C. Chen, "The architecture of the DIVA processing-in-memory chip," in *Proceedings of the 16th international conference on Supercomputing - ICS '02*, vol. 7, p. 14, 2002.

[7] C. E. Kozyrakis, S. Perissakis, D. Patterson, T. Anderson, K. Asanovic, N. Cardwell, R. Fromm, J. Golbus, B. Gribstad, K. Keeton, R. Thomas, N. Treuhaft, and K. Yelick, "Scalable processors in the billion-transistor era: IRAM," *Computer*, vol. 30, no. 9, pp. 75–78, 1997.

[8] M. Gokhale, B. Holmes, and K. Iobst, "Processing in memory: the Terasys massively parallel PIM array," *Computer*, vol. 28, no. 4, pp. 23–31, 1995.

[9] P. Kogge, "EXECUBE-A New Architecture for Scaleable MPPs," in *1994 International Conference on Parallel Processing (ICPP'94)*, vol. 1, pp. 77–84, 1994.

[10] J. H. Lau, "Evolution, challenge, and outlook of TSV, 3D IC integration and 3d silicon integration," in *2011 International Symposium on Advanced Packaging Materials (APM)*, pp. 462–488, 2011.

[11] J. Jeddeloh and B. Keeth, "Hybrid memory cube new DRAM architecture increases density and performance," in *2012 Symposium on VLSI Technology (VLSIT)*, pp. 87–88, 2012.

[12] J. Borghetti, G. S. Snider, P. J. Kuekes, J. J. Yang, D. R. Stewart, and R. S. Williams, "'Memristive' switches enable 'stateful' logic operations via material implication," *Nature*, vol. 464, no. 7290, pp. 873–876, 2010.

[13] A. Siemon, T. Breuer, N. Aslam, S. Ferch, W. Kim, J. van den Hurk, V. Rana, S. Hoffmann-Eifert, R. Waser, S. Menzel, and E. Linn, "Realization of Boolean Logic Functionality Using Redox-Based Memristive Devices," *Adv. Funct. Mater.*, vol. 25, no. 40, pp. 6414–6423, 2015.

[14] P. Chi, S. Li, Z. Qi, P. Gu, C. Xu, T. Zhang, J. Zhao, Y. Liu, Y. Wang, and Y. Xie, "Processing-in-Memory in ReRAM-based Main Memory," *SEAL-lab Tech. Rep.*, vol. 2015, no. 001, 2016.

[15] P. Sheridan, K.-H. Kim, S. Gaba, T. Chang, L. Chen, and W. Lu, "Device and SPICE modeling of RRAM devices.," *Nanoscale*, vol. 3, no. 9, pp. 3833–3840, 2011.

[16] B. Chen, F. Cai, J. Zhou, W. Ma, P. Sheridan, and W. D. Lu, "Efficient in-memory computing architecture based on crossbar arrays," in *2015 IEEE International Electron Devices Meeting (IEDM)*, pp. 17.5.1-17.5.4, 2015.

Chapter 7

Summary and Future Work

7.1. Simulation Study of Selector Element

In Chapter 2, the read operations of RRAM crossbar array have been systematically studied for the 1S1R configuration with nonlinear selector device and storage element. The read margin of crossbar array was found to strongly depend on the nonlinearity and the conductivity of the selector device. We optimized the selector benchmarks to better characterize the device behavior and balance the parameter requirements for device development. Several bias schemes have been analyzed to accommodate memory cells with different features and improve the entire array performance (e.g. read margin and power consumption). Influences from peripheral sensing circuit, parasitic resistance and memory resistance were also evaluated to shed light on the inherent voltage divider effect in the crossbar array. This simulation work provides a theoretical guidance for future RRAM device optimization and circuit design.

Publications resulting from this chapter

- **J. Zhou**, K.-H. Kim, and W. Lu, "Crossbar RRAM Arrays: Selector Device Requirements During Read Operation," *IEEE Trans. Electron Devices*, vol. 61, no. 5, pp. 1369–1376, 2014.
- S. Kim, **J. Zhou**, and W. D. Lu, "Crossbar RRAM Arrays: Selector Device Requirements During Write Operation," *IEEE Trans. Electron Devices*, vol. 61, no. 8, pp. 2820–2826, 2014.

7.2. Tantalum Oxide Selector Device

In Chapter 3, the conduction mechanism of a Pd/TaO_x/Ta/Pd selector device with high non-linearity and asymmetric I-V characteristics was investigated. The trapezoidal barrier model was adopted to consistently explain the observed transport behaviors at different temperatures, revealing dominant thermionic emission at positive bias and dominant tunnel emission at negative bias. Based on this TaO_x selector, a self-rectifying RRAM cell with a HfO_x switching medium was further developed. The proposed RRAM structure offers great advantages of high LRS selectivity (5×10^3) and simple fabrication processes. Numerical simulations show that up to 1Mb array integration is possible with this novel RRAM device.

Publications resulting from this chapter

- M. Wang†, **J. Zhou†**, Y. Yang, S. Gaba, M. Liu and W. D. Lu, "Conduction mechanism of a TaO_x-based selector and its application in crossbar memory arrays," *Nanoscale*, vol. 7, no. 11, pp. 4964–4970, 2015.

7.3. Ultra-low Power Conductive Bridge RAM (CBRAM)

In Chapter 4, we demonstrated the sub-nA operation of RRAM devices for the first time using a Cu based CBRAM with polysilicon in-cell resistor. A novel Cu/Al_2O_3/aSi/Ta RRAM cell with self-rectifying characteristics was further developed to improve the device reliability and support CMOS-compatible BEOL integrations. With an amorphous silicon barrier/rectifying layer, the proposed CBRAM device exhibits low current (~nA), high on/off ratio (>100x) and pronounced nonlinearity without sacrificing the retention and endurance. These excellent device characteristics can potentially lead to device applications in next-generation low-power large-scale crossbar memory arrays.

Publications resulting from this chapter

- S. Gaba, F. Cai, **J. Zhou**, and W. D. Lu, "Ultralow Sub-1-nA Operating Current Resistive Memory with Intrinsic Non-Linear Characteristics," *IEEE Electron Device Lett.*, vol. 35, no. 12, pp. 1239–1241, 2014.
- **J. Zhou**, F. Cai, Q. Wang, B. Chen, S. Gaba, and W. D. Lu, "Very Low-Programming-Current RRAM with Self-Rectifying Characteristics," *IEEE Electron Device Lett.*, vol. 37, no. 4, pp. 404–407, 2016.

7.4. Field-Induced Conductance Switching in the $LaAlO_3/SrTiO_3$ Interface

In Chapter 5, we investigated the conductivity switching in the LAO/STO heterojunction and demonstrated a reconfigurable top-gated transistor structure. By applying an external electric field, the 2DEG in the transistor channel can be modulated in a non-volatile manner. The experimental findings and simulation results indicate that the conductance switching observed in the oxide interface may originate from competing effects of ionic migration process and electron trapping. Our reconfigurable transistor device provides a novel structure in developing nanoelectronics component based on emerging complex oxides.

7.5. In-Memory Computing Using Crossbar RRAM

In Chapter 6, we proposed a novel in-memory computing architecture based on crossbar arrays. By exploiting the co-located memory-processing functionality of RRAM and the high parallelism of crossbar, we aimed to build high-performance systems free from the von Neumann bottleneck in conventional computing architectures. Basic design principles and operation protocols have been proposed and demonstrated. A 1-bit fuller adder and a 4-bit multiplier have been implemented through proto-type circuits and dynamic simulations to verify the architecture concept.

Publications resulting from this chapter

- B. Chen, F. Cai, **J. Zhou**, W. Ma, P. Sheridan, and W. D. Lu, "Efficient in-memory computing architecture based on crossbar arrays," in *2015 IEEE International Electron Devices Meeting (IEDM)*, pp. 17.5.1-17.5.4, 2015.

7.6. Future Work

Despite these progresses on device optimization, mechanism analysis and circuit application, work on resistive switching memory and reconfigurable devices is by no means complete. There are still several technical and scientific challenges that deserve continued exploration.

- Efficient array-level AC analysis for RRAM devices

In this work, the SPICE simulations for crossbar RRAM array focus on DC analysis. Given the transient response of RRAM device and the widespread parasitic RC delay, the conclusions based on the static bias condition may need to be modified for practical pulse operations. Although several dynamic RRAM device models have been developed [1], [2], conducting AC simulation is extremely time-consuming and limited to relatively small array size[3]. Therefore, the trade-off between accurate results and affordable computing resources needs to be carefully handled by introducing efficient simulation methods for large-scale crossbar array (e.g. Mega-bit).

- Improved material combination of CBRAM thin films

In Chapter 4, the low operation current of CBRAM device results from the combined effects of the switching medium and the rectifying/barrier layer. To constitute robust RRAM cells, the breakdown electric fields that cause device failure of the two functional

layers and the inherent voltage divider effect between them have to be considered, thus to avoid film degradation due to unbalanced voltage drop. Compared with the present amorphous silicon solution, a more reliable barrier layer that has stronger capability to suppress ion migration and sustain high electric field (e.g. graphene) may also be desirable [4]. For higher LRS nonlinearity, the Ta bottom electrode can be replaced with other metals with proper work function and reactivity with the barrier layer[5].

- Optimized program/erase algorithm and peripheral circuit

In addition to material optimization, excellent RRAM device performance requires precise electrical control on the filament growth-rapture processes. Several reported studies have implemented pulse-train schemes to minimize the distribution of resistance state after programming and achieve multi-level storage on metal-oxide RRAMs [6], [7]. In the case of sub-nA CBRAM, it is even more challenging to tune the filament shape with external electric control, since very limited numbers of metal cations are involved in the SET/RESET processes. Besides, low switching current is likely to cause technical issues on signal detection during the read. Therefore, adaptive program/erase algorithms and sophisticated peripheral circuits (e.g. alternative sensing schemes) need to be developed to support the low-power RRAM operations.

7.7. References

[1] P. Sheridan, K.-H. Kim, S. Gaba, T. Chang, L. Chen, and W. Lu, "Device and SPICE modeling of RRAM devices.," *Nanoscale*, vol. 3, no. 9, pp. 3833–3840, 2011.

[2] X. Guan, S. Yu, and H.-S. P. Wong, "A SPICE Compact Model of Metal Oxide Resistive Switching Memory With Variations," *IEEE Electron Device Lett.*, vol. 33, no. 10, pp. 1405–1407, 2012.

[3] H. Li, P. Huang, B. Gao, B. Chen, X. Liu, and J. Kang, "A SPICE model of resistive random access memory for large-scale memory array simulation," *IEEE Electron Device Lett.*, vol. 35, no. 2, pp. 211–213, 2014.

[4] L. Li, X. Chen, C. Wang, J. Cao, S. Lee, A. Tang, C. Ahn, S. Singha Roy, M. S. Arnold, and H.-S. P. Wong, "Vertical and Lateral Copper Transport through Graphene Layers," *ACS Nano*, vol. 9, no. 8, pp. 8361–8367, 2015.

[5] L. Zhang, B. Govoreanu, A. Redolfi, D. Crotti, H. Hody, V. Paraschiv, S. Cosemans, C. Adelmann, T. Witters, S. Clima, Y. Y. Chen, P. Hendrickx, D. J. Wouters, G. Groeseneken, and M. Jurczak, "High-drive current (>1MA/cm^2) and highly nonlinear (>10^3) TiN/amorphous-Silicon/TiN scalable bidirectional selector with excellent reliability and its variability impact on the 1S1R array performance," in *2015 International Electron Devices Meeting (IEDM)*, pp. 6.8.1–6.8.4, 2015.

[6] F. Alibart, L. Gao, B. D. Hoskins, and D. B. Strukov, "High precision tuning of state for memristive devices by adaptable variation-tolerant algorithm," *Nanotechnology*, vol. 23, no. 7, p. 075201, 2012.

[7] L. Zhao, H.-Y. Chen, S.-C. Wu, Z. Jiang, S. Yu, T.-H. Hou, H.-S. P. Wong, and Y. Nishi, "Multi-level control of conductive nano-filament evolution in HfO_2 ReRAM by pulse-train operations.," *Nanoscale*, vol. 6, pp. 5698–702, 2014.