Summer 2017

The Efficacy of Programming Energy Controlled Switching in Resistive Random Access Memory (RRAM)

David Malien Nminibapiel
Old Dominion University

THE EFFICACY OF PROGRAMMING ENERGY CONTROLLED SWITCHING IN RESISTIVE RANDOM ACCESS MEMORY (RRAM)

by

David Malien Nminibapiel
B.S. December 2010, Old Dominion University
M.S. August 2012, Old Dominion University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

ELECTRICAL AND COMPUTER ENGINEERING

OLD DOMINION UNIVERSITY
August 2017

Approved by:

Helmut Baumgart (Director)

Kin P. Cheung (Member)

Gon Namkoong (Member)

Zhili Hao (Member)

Christopher Bailey (Member)

ABSTRACT

THE EFFICACY OF PROGRAMMING ENERGY CONTROLLED SWITCHING IN RESISTIVE RANDOM ACCESS MEMORY (RRAM)

David Malien Nminibapiel
Old Dominion University, 2017
Director: Dr. Helmut Baumgart

Current state-of-the-art memory technologies such as FLASH, Static Random Access Memory (SRAM) and Dynamic RAM (DRAM) are based on charge storage. The semiconductor industry has relied on cell miniaturization to increase the performance and density of memory technology, while simultaneously decreasing the cost per bit. However, this approach is not sustainable because the charge-storage mechanism is reaching a fundamental scaling limit. Although stack engineering and 3D integration solutions can delay this limit, alternate strategies based on non-charge storage mechanisms for memory have been introduced and are being actively pursued.

Resistive Random Access Memory (RRAM) has emerged as one of the leading candidates for future high density non-volatile memory. The superior scalability of RRAMs is based on the highly localized active switching region and filamentary conductive path. Coupled with its simple structure and compatibility with complementary metal oxide semiconductor (CMOS) processes; RRAM cells have demonstrated switching performance comparable to volatile memory technologies such as DRAMs and SRAMs. However, there are two serious barriers to RRAM commercialization. The first is the variability of the resistance state which is associated with the inherent randomness of the resistive switching mechanism. The second is the filamentary nature of the conductive path which makes it susceptible to noise.

In this experimental thesis, a novel program-verify (P-V) technique was developed with the objective to specifically address the programming errors and to provide solutions to the most challenging issues associated with these intrinsic failures in current RRAM technology. The technique, called Compliance-free Ultra-short Smart Pulse Programming (CUSPP), utilizes sub-nanosecond pulses in a compliance-free setup to minimize the programming energy delivered per pulse. In order to demonstrate CUSPP, a custom-built picosecond pulse generator and feedback control circuit was designed. We achieved high (10^8 cycles) endurance with state verification for each cycle and established high-speed performance, such as 100 ps write/erase speed and 500 kHz cycling rate of HfO_2-based RRAM cells. We also investigate switching failure and the short-term instability of the RRAM using CUSPP.

This dissertation is dedicated to my parents, Clement Kojo Nminibapiel and Victoria Assibi Nminibapiel and to my siblings Victor, Edna, Elsie and Shequita.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Helmut Baumgart for giving me the opportunity to pursue both a Masters and a Ph.D. degree under his guidance. On countless occasions in my graduate school career, he has been supportive and has provided me opportunities that I will be forever grateful for.

I would like to thank Dr. Kin P. Cheung for his guidance and patience in my dissertation research. He accepted me into his group at the National Institute of Standards and Technology (NIST) and has always encouraged me to be the best researcher I can be.

I would like to thank the other members of my Ph.D. committee, Dr. Zhili Hao, Dr. Gon Namkoong and Dr. Christopher Bailey for their guidance.

I want thank Dr. Jason Campbell, Dr. Jason Ryan and Dr. Dmitry Veksler for their unwavering willingness to assist me during my research work. I would have not been successful otherwise.

I am thankful to Dr. Pragya Shrestha, Dr. Ji-Hong Kim, Dr. Asahiko Matsuda and Vasileia Georgiou at NIST for their assistance during my research work. Particularly, I would like to express my gratitude to Dr. Shrestha for her support and guidance.

I would like to thank Dr. Yaw Obeng and Dr. Albert Jones at NIST for his advice and assistance, and Dr. Kai Zhang at the Applied Research Center in Newport News, Virginia for his support and assistance during the course of my graduate career.

I would like to thank my parents and siblings for their never-ending love, support and encouragement and lastly, I would like to thank Ms. Safia Hurst for her support and encouragement.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

1.1 Background

Resistive switching (RS) refers to the phenomenon where insulators, under an applied electrical field, can undergo reversible resistance change from their normal insulating state to an abnormal conductive state. This phenomenon was first discovered in the 1960s [[1-4] and was proposed for use in memory applications. However, memory technology based on RS remained primarily a field of scientific study for the next 30 years, mainly because of the advancements of Si-based memory technology. It saw a re-emergence in the late 1990s in perovskite oxides [5-7]. Shortly thereafter, Baek et al. [8] demonstrated a robust memory device based on a binary transition metal oxide insulating layer. This breakthrough renewed interest in memory devices based on the RS mechanism [9-12]. These devices are called Resistive Random Access Memory (RRAM).

In its basic configuration, an RRAM cell comprises an insulating layer sandwiched between two conducting layers in a capacitor-like structure. The conducting layers are typically either a metal or a conductive nitride; the insulating-layer material, however, varies across a wide range of organic, inorganic, and hybrid materials [13, 14]. The working principle of these devices is based on modulating the conductivity of the metal-insulator-metal (MIM) structure by applying the appropriate bias. When this bias is removed, the cell retains the resistance level, making RRAM a non-volatile memory. The active region of RS can be as large as the entire insulating layer or a 1-dimensional (1-D) conductive path that electrically connects the two conductors [15]. The former is usually observed in Phase Change RAM (PCRAM) [16]. RRAM devices based on a 1-D conductive path are desirable because of the promise of extremely high

scalability enabling continued growth in density for future technology nodes [17]. The conductive filament (CF) is widely accepted to be composed of metal atoms or reduced oxides [18, 19]. The localized nature and composition have been verified through various methods, which include conductive atomic force microscopy (C-AFM) [20], high resolution transmission electron microscopy (TEM) [21], electron energy loss spectroscopy (EELS) [22] and X-ray absorption spectromicroscopy [23].

In a single bit cell, two distinct states are programmed: high resistance state (HRS) and low resistance state (LRS), representing the low (0) and high (1) states, respectively. The transition from HRS to LRS is referred to as SET while the opposite transition from LRS to HRS is referred to as RESET. Prior to performing SET or RESET or cycling, a FORMING step is required. V_{FORM}, V_{SET} and V_{RESET} represent the voltages required for FORMING, SET and RESET. The read-out operation is performed with a low voltage (V_{READ}), which should not disturb the resistance state.

Traditionally, switching modes are classified as either unipolar or bipolar. This classification is based on the sensitivity of the switching direction on the bias polarity. Unipolar devices can SET and RESET at the same fixed polarity, while bipolar devices can only SET and RESET at opposite polarity. The switching mode is not an intrinsic property of the cell but can be controlled by the stack properties and programming currents [24-26]. This explains the appearance of a so-called non-polar mode where switching direction is independent of polarity [13]. In regard to implementation, unipolar devices have the advantage of requiring a simpler programming algorithm since polarity is fixed. All switching modes are schematically shown in Figure 1.

Filament formation in the FORMING and SET step is accompanied by an inrush of current which must be controlled to avoid permanent damage [27]. Usually, a current-limiting element such as a resistor or transistor is placed in series with the RRAM to set a current compliance (I_C) limit. I_C is not only necessary to prevent thermal damage; it also serves to control the post-FORMED/SET state resistance value [28]. The inrush of current is a result of the abrupt change of the RRAM cell from high resistance to low resistance under the programming voltage. Ideally, the compliance element should limit the maximum allowable current that flows through the cell. However, this function can be undermined by the response time of the system. If the transition time of the RRAM cell is faster than the RC time of the system, a current overshoot is inevitable. This overshoot current exceeds the current limit set by the compliance element and therefore defeats the purpose of the compliance element [29]. To minimize current overshoot, strict control of parasitics in the system is required. For this reason, advanced RRAM cells either included an additional layer to serve as a series resistor or are integrated directly on the transistor gate.

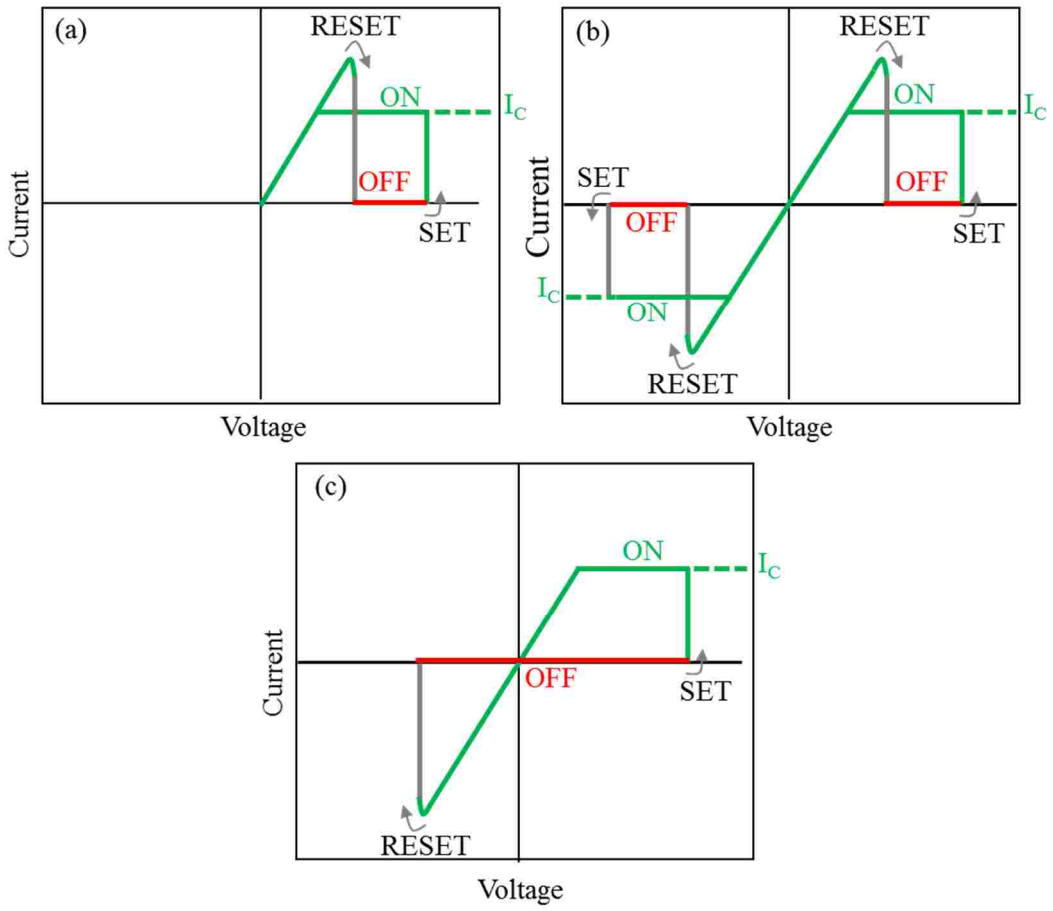

Figure 1. Schematic of a) unipolar, b) nonpolar and c) bipolar switching modes of RRAM cells under voltage sweep.

Cell level

Several aspects of the resistive switching mechanism in RRAMs have led to its emergence as a serious candidate for future nonvolatile memory. Conversely, these aspects also have given rise to most of the issues still hindering the full commercialization of RRAM-based memory chips. The wide variety of material systems which display RS enables an equally wide application space for RRAM [14]. RRAMs based on novel hybrid insulator materials like metal organic framework (MOF) films [30], and novel transparent electrode materials like Indium tin oxide (ITO) [31] have been demonstrated. In addition, switching performance is highly sensitive

to the cell stack and structure, which allows for stack engineering strategies to improve performance [32]. For instance, TaOx-based devices tend to exhibit superior retention [33], while conductive-bridge RRAM devices using electrochemically active electrodes tend to display large, resistance-memory windows [34]. On the other hand, because there are several stack options, it is challenging to construct models that can fully explain the different aspects of resistive switching. Without these models, device optimization and performance predictions become difficult. Also, the issue is reflected in the literature, where there are conflicting reports of the switching behavior even within the same RRAM material system. The confusion can be related to the interplay of multiple physical processes present during cell operation and their dependence on the material stack and programming environment.

Several features of the RRAM conductive path make it highly attractive for high density and low-power NVM. At the same time, these features are the root cause of its most severe challenges. Theoretically, the size of the conductive path can reach atomic scales without degrading performance while in some cases, scaling has been shown to improve performance [35, 36]. Generally, resistance levels increase with scaling, which is advantageous for power considerations. Also, ionic transport at the nanometer scale can cause sufficient resistance change; therefore, overall programming energy for RRAM is low.

However, it is this highly localized active region and filamentary conductive path which generate severe reliability and variability issues because of the reduced number of participating defects [37]. At these scales, the charge state or motion of a single defect within or close to the CF can dramatically change the resistance of the CF and give rise to instability [38-40]. Uniformity is affected by this behavior because the exact arrangement of defects changes for each cycle causing the so-called cycle to cycle variability. Read operation is prone to Random

Telegraph Noise (RTN) signals which not only degrade the resistance window [41, 42] but also disrupt program-verify techniques [43]. In a Multilevel (MLC) the number of states is limited by instability in the resistance state [44-46]. Therefore, thin conductive paths, which are desirable for high density and low-power applications, are susceptible to large fluctuations that affect stability and uniformity.

Array level

Ideally, a RRAM memory array can be a true cross bar architecture where the device area is defined by the cross point of the bitline (BL) and wordline(WL), which are arranged orthogonal to each other [47-50]. Due to the scalability of RRAMs, the linewidths and line spacing can be minimized to achieve a high density array. An RRAM cell footprint is $4F^2$, where F is the smallest feature size of a given process [51]. The memory cell area can be further reduced by utilizing the MLC capability of RRAMs or multi-stack cell structure [52]. With this approach, the cell area becomes $4F^2/n$, where n is the number of levels or stacks per device [53]. In addition to the cycle to cycle variability, device to device variability is also present at the array level. This variability is in part unavoidable due to the intrinsic randomness of the RS mechanism. However, cell to cell variability is more prone to additional extrinsic sources of variability such as poor process uniformity across the chip.

Sneak current is another serious issue at the array level. In the crossbar architecture, multiple cells share common BL and WL. Therefore, in addressing a device in HRS, current can sneak past the higher resistance device and flow through devices in LRS. Sneak currents are problematic because they increase the thermal budget, decrease array size, and reduce read-out margins [54]. Sneak current solutions involve integrating the RRAM cell with a selector device such as a diode or transistor in a 1 Diode - 1 RRAM (1D1R) configuration or a 1 Transistor - 1

RRAM configuration (1T1R), respectively [55, 56]. Although both configurations can eliminate sneak currents, the additional element increases the cell footprint.

An alternate solution relies on the non-linearity of the RS mechanism to minimize sneak current. This strategy, referred to as 1 RRAM (1R) configuration, eliminates the need for an area-consuming selector device. Dimin Niu et al. [57] investigated the performance of all three configurations – 1R, 1T1R, 1D1R – and determined that 1D1R is suitable for devices operated at high operating current and low nonlinearity, while 1R is suitable for low operation current and high nonlinearity. An anti-serial connected TaOx-based cell has also been demonstrated to minimize sneak current [54, 58].

Programming solutions to RRAM issues

Strategies to improve reliability and uniformity can be separated into two groups: programming and non-programming. In the latter approach, common strategies to improve performance involve material selection, stack engineering, doping and cell-structure optimization [33, 59-64]. H.Y. Lee et al. improved the resistance window (>1000) and achieved low operation voltage by incorporating a Ti layer to a HfOx cell [11]. Traore et al. observed better retention in HfOx-based devices doped with Al and Ti [65]. However, these strategies can be nullified by improper programming.

Programming current control remains one of the key parameters responsible for improving several key performance markers of RRAMs. High endurance and retention are linked to optimizing the programming currents since excessive and insufficient currents can result in failure [66, 67]. The former case results in reversible failures [68], while the latter results in permanent failures [69]. The RS mechanism is highly sensitive to temperature which in turn is generated by the current flowing through the conductive filament. As will be discussed in

Section 1.2, thermochemical cells rely on temperature for filament formation and rupture [70]. Therefore, it is particularly important to limit the maximum current or energy dissipated in the CF.

As previously stated, current control is traditionally executed with a series current limiting element such as a resistor or transistor [71, 72]. However, current control without the additional series element can be achieved by pulse width controlled current duration [73] and by current sweep programming [74, 75]. The former achieves current control by limiting programming energy and has been demonstrated in forming studies and endurance studies [43, 76]. The latter approach takes advantage of the gradual filament formation of the current sweep to eliminate excessive programming current. Improved retention has been demonstrated in the current sweep approach [77].

Although constant stress (DC) and sweep programming modes are commonly used in RRAMs [69, 74, 78], pulse programming is preferred for practical studies of switching performance [79]. This is because pulse programming is fast, energy efficient and mimics the true programming operation at the chip level. Chen et al. demonstrated a high endurance obtained by balancing V_{SET} and V_{RESET} [68], indicating the need to optimize pulse parameters such as rise/fall time, width and amplitude to improve endurance [80-82].

Variability remains one of the major challenges in RRAMs because it is intrinsic to the nature of the resistive switching process [83]. Along the filament, the exact location and extent of filament formation and rupture cannot be reproduced from cycle to cycle or from device to device. Therefore, post programmed states are represented as distributions instead of well-defined values. To enable accurate state determination during read-out operation, the distributions must never overlap. Program-verify techniques have been developed primarily to

ensure distinguishable resistance distributions or sufficient resistance window (RW). Generally, program verify algorithms adjust the pulse number [43, 84], width [85] and amplitude [86-88] until the desired RW is achieved. Program-verify has also been used to tighten the resistance distribution which is highly desirable for MLC applications of RRAMs [86].

1.2 Origins of Filamentary Resistive Switching

The memory function in RRAM is based on the reversible change of the cell resistance between at least two distinct resistance states. As described earlier, the memory cell structure is composed of two conductive layers separated by an insulating layer, in this case an oxide. Therefore, to change the resistance of this otherwise insulating cell would require the presence of some conductive path through the oxide connecting the two electrodes: top electrode (TE) and bottom electrode (BE). The conductive path may be related to an interface effect such as Schottky barrier modulation or a localized path adjoining TE and BE. The former is usually observed when a semiconductor is used as the insulating layer [14]. Therefore, based on the type of semiconductor and electron metal work function, an Ohmic contact or Schottky barrier that can be modulated is formed [51]. The switching mechanism may also be due to charge trap/detrap in the insulator. In which case, the presence or absence of charge can affect conduction at the interface [89, 90]. Cells that rely on interface modulation to switch are referred to as interface-type cells. In LRS (On-State), conduction ideally occurs homogeneously across the entire interface, making the LRS of interface-type cells strongly sensitive to area scaling [91].

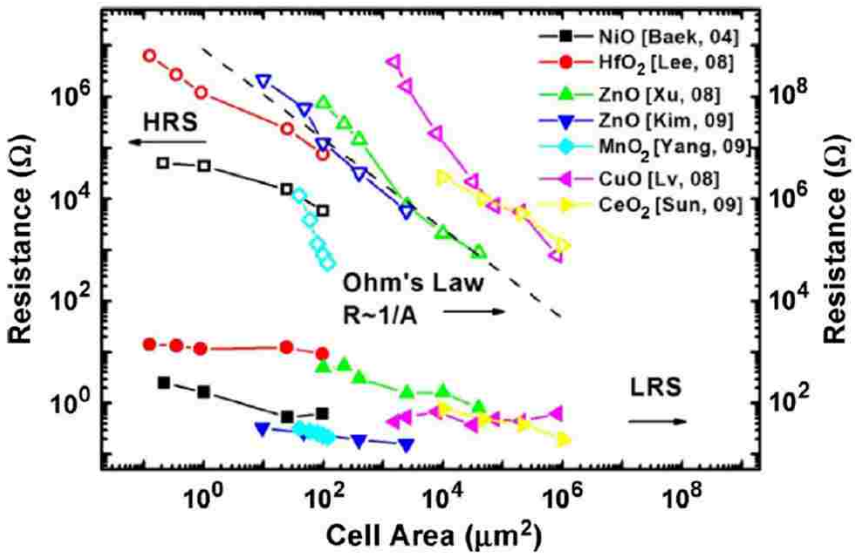

Figure 2. LRS and HRS dependence on cell area for different RRAM stacks [13].

On the other hand, some RRAM cells show a very weak, sometimes negligible, dependence of LRS on the device area. Figure 2 shows the dependence of LRS and HRS on the cell area. It is evident that HRS decreases considerably with scaling, while LRS shows negligible change. This suggests that the conductive path is localized with lateral dimensions much smaller than the device area. These cells are referred to as filamentary-type cells. In HRS, the filament is ruptured, disconnecting TE and BE, while in LRS the filament is intact. Ideally, in filamentary-type cells, a single, localized, conductive filament is responsible for RS. Due to the highly localized filament, device performance frequently remains, more or less, unchanged; however, in some cases, performance actually improved by area scaling [35, 36]. The filament dimensions will limit the extent to which area scaling does not impact LRS. In this thesis work, we examine filamentary-based RRAMs.

The conductive filament (CF), which is electrically more conductive than its surrounding insulating matrix, is a localized path in the insulating layer [92]. Therefore, conduction is unaffected by area scaling until the device area approaches the lateral dimensions of the filament. The resistance state of the memory cell is determined by the physical and geometric properties of the CF. As depicted in Figure 3, higher resistance states are accessed when the CF shrinks in width. Even higher resistance states are accessed when the CF is completely severed, i.e., electrically disconnecting the two electrodes.

Prior to direct observations of the localized conductive path in the switching matrix, the localized nature was observed through C-AFM [93-95] techniques. In LRS, nanoscale conductive paths are present, while in HRS these paths disappear or are reduced in size or number. The appearance of multiple conductive paths suggests multi-filament formation. However, direct observations of the CF, particularly non-metallic filaments, are challenging because of the localized nature and the difficulty in differentiating the CF from its surrounding insulator matrix. Transmission electron microscopy (TEM) has been used to identify conductive paths in metal oxides [21]. Using X-ray absorption spectroscopy, Strachan et al. observed a localized reduced region in TiO_2, which was responsible for the large increase in conductivity in LRS [96].

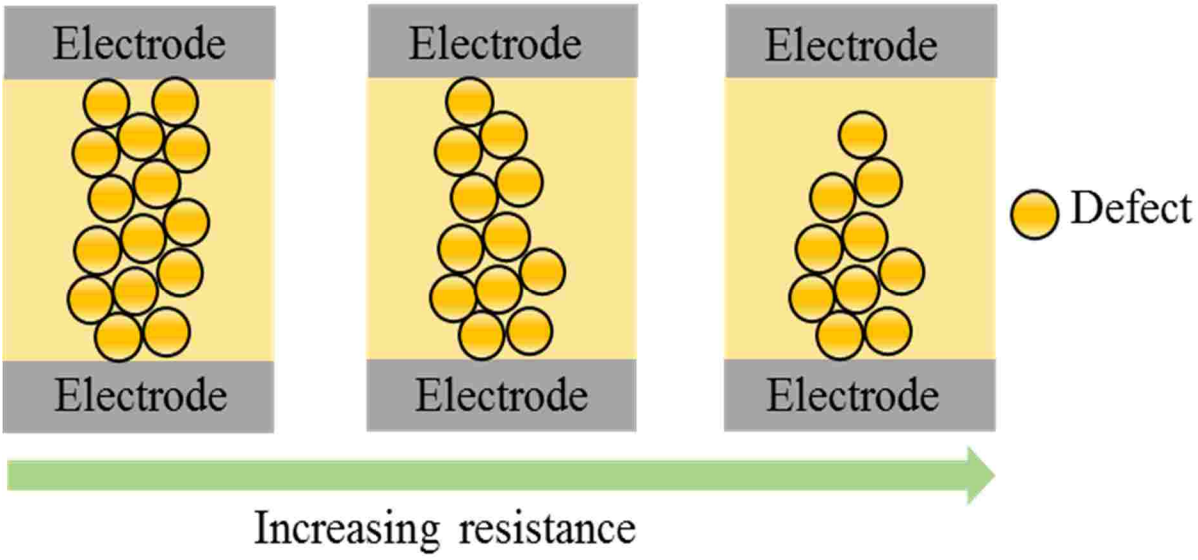

Figure 3. Illustration of the possible physical configuration of the CF in an RRAM cell. CF resistance increases as length and/or width of CF decreases.

The CF may display metallic-like or semiconductor-like conduction based on its composition, which is ultimately determined by the cell material stack and the programming currents. Metallic-like conduction is usually observed in RRAM cells whose electrochemically active metal electrode is dissolved and deposited in the insulating matrix. In these cells, the CF is composed of metal atoms. However, metallic-like conduction was observed in a Pt/NiO/Pt cell which has two inert electrodes [97, 98]. Therefore, metallic conduction is not limited to cells which include an electrochemically active electrode. Semiconductor-like conduction occurs where defects (oxygen vacancies, cation vacancies) are generated and accumulated in the insulator to form a conductive path. In such cases, the CF is not composed of metal precipitates from an electrochemically active electrode. Rather, they are usually composed of sub-oxides, which undergo a redox reaction due to the concentration of oxygen vacancies, in the switching

layer. F. Pan et al. describe two switching mechanisms based on ion migration: cation migration and anion migration mechanisms [14].

In the cation migration model, an electrochemically active electrode such as Ag is positively biased (anode) leading to the dissolution of the electrode ($Ag \rightarrow Ag^+ + e^-$). The resultant cations (Ag+) migrate under the applied field towards the negative biased electrode (cathode). At the cathode, the cations are reduced back to metal ($Ag^+ + e^- \rightarrow Ag$) and begin to form the metallic dendritic filament that eventually bridges the two electrodes. The nucleation and mobility of the cations in the insulating layer determine the growth direction of the metallic filament. In Ag/SiO2/Pt [99], Ag/a-Si/W [25] and Ag/PEDOT:PSS/Pt [100] the filament has been shown to grow from cathode to anode, anode to cathode, and in both directions, respectively. RRAM cells where cation migration is responsible for RS are referred to as electrochemical metallization cells (ECM), conductive bridge RAM (CBRAM) or programmable metallization cells (PCM). The conductive filament in CBRAM-type devices have been confirmed with TEM analysis [101]. Using electron energy loss spectroscopy (EELS) and energy dispersive X-ray (EDX), Xing Wu et al. determined that the CF in HfOx cell with a Ni TE was composed of Ni that migrated from the TE into the HfOx layer [22].

In the anion migration model, the electrodes do not contribute species to form the CF. Usually, the CF is composed of a localized sub oxide that is more conductive than the parent oxide material. Redox reactions driven by the migration of oxygen ions (anions) in the oxide are attributed to the RS. Under the applied electric field, anions migrate to the positively biased electrode leaving behind oxygen vacancies and a subsequent valance state change of the cation sublattice. This modifies the electronic conductivity and forms the conductive path during FORMING and SET [102]. By applying the positive electric field, anions migrate back into the

oxide and recombine with oxygen vacancies during RESET. These oxygen-vacancy defects result in the generation of sub stoichiometric oxides which form the conductive path. It is common for RS to be described by oxygen vacancy migration analogous to hole movement. RS based on oxygen ion/vacancy migration is referred to as Valence Change Memory (VCM). Typical oxides which exhibit anion-migration induced redox reaction include, Ta_2O_5, HfO_2 and ZrO_2. It is common for the oxides to be represented as TaOx, HfOx, ZrOx as they are purposefully deposited with significant oxygen content deficiency. In TiO_2, for example, the conductive path was observed to be a result of the formation of Ti_4O_7 in the TiO_2 matrix [21, 96].

The oxygen vacancy concentration distribution is critical to the switching kinetics in VCM cells. Cells that have uniform distribution of oxygen vacancies tend to exhibit both bipolar or unipolar switching modes. In the bipolar case, filament formation and rupture are caused by anion migration induced by the applied field [103]. In the unipolar case, temperature drives both filament formation and rupture [104]. Cells can be structured intentionally to have a non-uniform distribution of oxygen vacancies to make the switching direction more sensitive to polarity. This can be achieved by using a top metal electrode (eg., Ta, Al or Ti) with a high oxygen affinity [105] or introducing a layer (eg., TaO_{2-x}) with high concentration of oxygen vacancies. This creates a disproportionate number of oxygen vacancies close to TE which favors a positive bias on TE for SET since negative bias would initially drive oxygen ions back into the vacancies.

RS based on field-induced ion processes (generation, migration, redox) were described above; however, these processes may be thermally induced. A third class of cells dominated by thermal effects are referred to as Thermochemical memory (TCM) [106] or fuse-antifuse cells [22]. During FORMING/SET, filament formation is driven by thermal decomposition of the oxide layer. The deficiency of oxygen results in CF of either oxygen vacancies [107] or metal

precipitates [108] or a mixture of both [22][31]. The driving force for the formation of metallic CFs is the energetically favored, lower-valence states of these metal oxides at high temperatures [14]. TCM cells usually show metallic conduction in the ON state and semiconductor-like conduction in the OFF state [97]. During RESET, thermally activated O^{2-} diffusion driven by the concentration gradient of the oxygen deficient CF and the oxygen rich surrounding matrix oxidizes the CF to create a dielectric barrier. Lee et al. [95] observed evidence of O^{2-} migration in a Pt/NiO/Pt cell by secondary mass ion spectroscopy. Because joule heating is independent of the polarity of the applied voltage, the nonpolar switching mode is usually observed. TCM switching behavior has been shown in CoO, TiO_2, NiO, ZnO based oxides sandwiched between two inert electrodes in a symmetric cell [109].

Regardless of the switching mechanism, ECM, TCM or VCM, the as-deposited RRAM oxide typically requires an initial step to construct the conductive path. This is the FORMING step. FORMING can be interpreted as a soft breakdown event [110], which generates the minimum number of species needed to electrically connect TE and BE. Time-dependent, dielectric-breakdown measurements show that the relationship between FORMING voltage and FORMING time is exponential [111]. Therefore, the percolation theory of dielectric breakdown can be used to explain the FORMING process [112]. RS generally involves an interplay of electrochemical and thermochemical processes, where one may dominate depending on the specifics of the stack and the programming procedure. V_{FORM} decreases with oxide thickness suggesting that FORMING is a field driven process [113]. At high (> 10 MV/cm) electric fields, metal oxide bonds can be broken creating both ions and defects. However, the temperature due to leakage current is expected to contribute significantly to bond breaking and filament formation. For example, in a polycrystalline film, grain boundaries act as precursor sites for the

CF because of enhanced leakage current paths via defects [97]. Electron transport through such vacancies is accompanied by energy dissipation resulting in local temperature increases [27], which can enhance vacancy generation. This, in turn, lowers the activation energy barrier for subsequent vacancies. Coupling this lower activation energy with the exponential dependence of vacancy generation on both the electric field and the temperature leads to positive feedback that can accelerate CF formation.

High voltages are typically needed for FORMING since it generates the prerequisite number of defect species to construct the entire CF. Because of the required voltages, FORMING is usually undesirable because it increases power consumption and the high voltage can easily initiate thermal damage. Consequently, several strategies have been employed to obtain so-called 'FORMING free' [60] [114, 115] devices or, at a minimum, 'decreased FORMING voltages'. Some strategies take advantage of the linear relationship between V_{FORM} and oxide thickness, where simply decreasing the oxide thickness will decrease V_{FORM}. Other strategies to decrease V_{FORM}, involve purposefully depositing defect rich oxides by controlling oxygen content or anneal ambiance during oxide deposition [116]. The higher leakage current of these defective oxides dominates the FORMING process while keeping V_{FORM} low. Hot forming has also been used to reduce the FORMING voltage and time [78]. At elevated programming temperatures, the electron transfer rate increases, and an associated increase in dissipated power during charge trap/detrap accelerates the generation of defects. This sensitivity of the forming step to temperature of the device corroborates the role of temperature in resistive switching.

RESET and SET do not scale with oxide thickness. This suggests that only a small portion of the CF participates in RESET and SET. This is also evident by the lower V_{SET} and V_{RESET} voltages compared to V_{FORM}. The SET operation is a scaled down version of FORMING

as it reforms the rupture portion of the CF. Therefore, current compliance is required during SET. RESET operation makes the CF less conductive. This can be explained either by a reduction of CF dimension or the creation of a dielectric barrier in the CF. In the former scenario, the filament is still intact while in the latter, the filament is ruptured. Two prominent models have been proposed to explain the RESET process in filamentary RRAM cells: the thermal-dissolution model [117] and ionic-migration model [118]. In the thermal-dissolution model, the applied voltage causes current to flow through the CF. Joule heating leads to a temperature rise that leads to dissolution of the CF. This process is self-accelerated because the hottest spot on the filament shrinks the fastest, enhancing the electric field and current density. Since this process is independent of the polarity of the applied voltage, it describes unipolar or nonpolar RESET. However, in the case where voltage polarity is required such as in bipolar RESET, the electric field drives the dissolution of the CF. Oxygen ions (O^{2-}) drift to recombine with oxygen vacancies [119].

The physical and chemical properties of the CF are affected by the material stack, the interface, the film structure, and the thermodynamics associated with programming. Because of the direct impact of these parameters on the CF, switching performances such as endurance and retention can be tuned. For instance, the higher performance of some material systems is associated with the stable oxides that constitute the CF. Like in TaOx-based systems, two thermodynamically stable solid phases are achievable, resulting in more stable CF compared to TiOx [120, 121]. Programming also directly influences the physical and chemical nature of the CF.

1.3 HfO_2-based RRAM

As previously described, several materials can be used as the insulator or storage layer. Of these materials, transition metal oxides are favored based on several key factors. In transition metal oxides, the stoichiometry can be adjusted to control material properties and improve switching performance [95]. Most metal oxides (CuOx, WOx, HfOx, SiOx) already in use in complementary metal oxide semiconductors (CMOS) show the resistive switching phenomenon. This simplifies and dramatically reduces the cost of implementing memory technologies based on RRAM. For some stacks, such as WOx and CuOx, a simple, additional oxidation step is enough to integrate RRAMs with a transistor [122]. For RRAM arrays that require a selector diode or transistor, the thermal stability of metal oxides becomes an indispensable feature as well. The popularity of oxide-based RRAM cells is also based on the merit of their switching performance in comparison to other material systems [13].

One of the more mature metal oxides explored for RRAM is HfO_2. RRAMs based on HfO_2 switching layers have shown excellent switching scalability and performance. In addition, HfO_2 has proven compatibility with CMOS technology, where it is employed as a high-κ gate dielectric for high-performance MOSFETs. Figure 4 shows a SEM and TEM image of two different HfOx devices. Govoreanu et al. demonstrated a cross point (Figure 4 (a)) 10 x 10 nm^2 TiN/Hf/HfOx/TiN cell with good endurance (10^7 cycles) and stability (30 hrs @ 200^OC) [35]. Zhang et al. fabricated devices, shown in Figure 4 (b) with an active area of a few nanometers (< 10nm) with multilevel functionality and 10^8 endurance [36]. In addition to MLC capability, S. Yu et al. fabricated vertical cells for use in a three-dimensional (3D) cross point architecture which would further increase the storage capacity of the memory chip [52]. In regard to

switching performance, ultra-low (pJ) switching energy, sub ns switching speeds, 10^{10} endurance has been demonstrated in HfO_2 based devices [11, 43, 123].

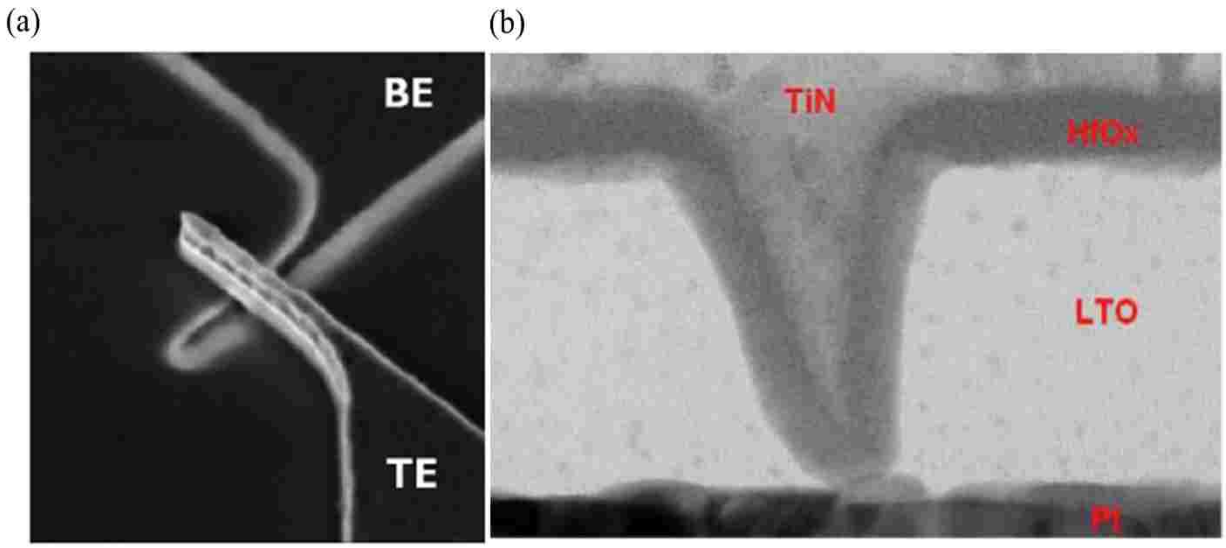

Figure 4 (a) Top view SEM image of TIN/Hf/HfOx/TiN cell. Cell area is defined at the crossing of TE (10 nm line width) and BE (10 nm line width) [33]. (b) Cross-sectional TEM image of TiN/HfOx/Pt device. The active area is < 10nm, defined by TE (TiN) [34].

The RS mechanism in HfO_2 is widely accepted to be attributed to the formation and rupture of a conductive filament due to oxygen ion and vacancy processes. [66-68]. Unless an electrochemically active electrode is used [22], most HfO_2-based devices are categorized as VCMs. However, unlike CBRAM where the CF is composed of metal atoms, the CF in VCM is mostly composed of a reduced valence state of the oxide. Therefore, observing CF using TEM is not easily accomplished [124]. However, the conductive path has been evidenced by U. Celano et al. using an AFM based tomography technique [20]. Kumar et al. [23] observed lateral oxygen migration using X-ray absorption spectromicroscopy. The conductance of the CF may display

semiconductor-like or metallic-like properties depending on the programming currents and material stack [125].

Usually, amorphous or polycrystalline HfO_2 layers are deposited for RRAM using various deposition techniques including atomic-layer deposition (ALD) [126], rf-magnetron reactive sputtering [127] and atomic-vapor deposition [128]. Generally, single crystalline films show lower V_{FORM} compared to polycrystalline and amorphous oxide films. This is attributed to grain boundaries in the crystalline films, which provide favorable sites for CF growth [93]. The grain boundaries also serve as low-energy diffusion paths for oxygen ions to easily diffuse through [129]. For polycrystalline films, the oxygen vacancies form an extended defect or percolation path as observed in NiO [130]. In amorphous oxide material, however, there are no preferable sites for CF formation; hence, the CF growth location is extremely random [131] and V_{FORM} is high [132]. During deposition, the ambient [132] and annealing conditions [127] can impact switching parameters. This is due to the role of oxygen ions in RS as these processes control the oxygen content during deposition.

HfOx-based RRAM cells can exhibit unipolar, nonpolar, and bipolar switching modes [109, 127, 133]. As discussed earlier, the switching mode is not intrinsic to the insulating layer but is dependent on the complete material stack of the cell. Unipolar and nonpolar are usually achieved in symmetric cells, for example a Pt/HfOx/Pt cell, since both TE and BE interfaces are the same. However, unipolar switching has also been observed in asymmetric cells. L. Goux [134] et al. observed both unipolar and bipolar switching in TiN/HfO2/Pt asymmetric cells. For TiN-HfO2-Pt stacks, poor RESET was observed for both unipolar and bipolar when positive bias was applied to TiN electrode. RESET and cycling were better with a positive bias applied to the Pt electrode. This behavior suggests that the switching mode classification should be based on

switching yield since some devices can display multiple modes depending on the chosen bias. Bertaud et al. [135] observed that the switching mode did show a dependence on electrode choice. Metals like Pt and Cu, with high enthalpy of formation of oxides lead to unipolar switching, while metals such as Al, Hf, Ti with low enthalpy of oxide formation lead to bipolar switching.

Most advanced HfO_2 devices have at least one additional layer in the MIM structure which is inserted between TE and HfO_2. To explain RS, it is generally accepted that mobile O^{2-} ions diffuse back and forth to modulate the defect rich CF [24]. This movement of mobile oxygen ions is central to the RS mechanism. The additional layer is purposefully incorporated in the MIM structure to store and release O^{2-} ions during programming. These functional layers are commonly referred to as oxygen-exchange layers (OEL), scavenging layers, or insertion layers [136, 137]. Multilayer cells are also fabricated not for oxygen scavenging, but to create an internal series resistance for low-power operation [63][[64, 84, 138].

Typical materials used as OEL include Ti, Ta, Hf and Zr [120]. These metal layers easily form metal oxides in the memory stack by extracting oxygen from the HfO_2 layer. For example, a Ti OEL layer leads to the formation of TiOx at the interface [139]. This process leaves behind a defect rich, non-stoichiometric oxide layer. The extent of this extraction process is determined by the choice of metal cap material and its thickness. HfOx cells with Ta OEL typically exhibit better retention compared to Ti and Hf [120]. This is attributed to the higher oxygen diffusion barrier of the Ta-O compared to Ti-O. Yuzheng Guo et al. [33] used defect energy calculations to determine that OEL can control both vacancy concentration and its charge state. However, this means release of O^{2-} for RESET is harder and affects repeated cycling. This endurance-retention trade-off was reported by Y.Y Chen et al. [140]. The OEL thickness also affects cell resistance

and switching voltages. In a study of the cell resistance on OEL to oxide ratio, D. Walczyk et al. [55] found that the pristine resistance of a Ti/HfO_2/Ti/TiN cell decreased with decreasing HfO_2 to Ti ratio. This is explained by the degree of oxygen gettering from the HfO_2 layer by Ti. Thicker or more reactive OEL layers extract more oxygen, which dramatically lowers the resistance and voltages required for switching. Additional benefits of OEL include, improved read disturb, increased HRS, and improved window [134] [141].

Cells with OEL generally support the bipolar mode because of the imbalance distribution of defects along the thickness of the oxide. This is because oxygen gettering is more aggressive close to the OEL [142]. The VCM model postulates that oxygen ions drift to the positively biased electrode, leaving behind defects (oxygen vacancies) to create the conductive path. Due to the spatial distribution of defects, it is preferred to apply a positive voltage at the OEL – oxide layer to initiate FORMING. It is expected that the filament grows from TE to BE and should therefore assume a conical shape as confirmed by C-AFM [20] and scanning TEM [143]. During repeated cycling, it is conceivable that over time, the location of the constriction might change, giving rise to different shaped filaments [144]. Forming-free HfO_2 based devices have been demonstrated [60, 114, 132]. These devices are usually functionalized by stack engineering and doping methods. However, Butcher et al. [78] demonstrated a programming solution to decreasing V_{FORM} , where FORMING was performed under elevated temperatures.

RESET is expected to occur at the thinnest portion (constriction) of the CF. This is because current density and temperature are highest at the constriction. For a conical shaped filament, the RESET location should be located at the BE (opposite from OEL interface). L. Goux et al. [134] observed a high sensitivity of RESET on the BE interface, corroborating the role of BE in the RESET operation. Because the CF resistance increases in RESET, the process

is assigned to a decrease in CF dimensions or dielectric barrier in the filament. Usually, a complete rupture results in deeper HRS. To RESET the device, a negative bias is applied to the TE with BE grounded. The O^{2-} stored in the OEL drifts under the applied field into the bulk oxide and recombines with the vacancies to re-oxidize the filament, either partially or completely. Note, the RESET may also be initiated thermally due to joule heating of the CF. Both unipolar and bipolar RESET has been observed in HfOx-based devices. In a model proposed by B. Gao et al. [119], the electric field creates Vo^{2+} by depleting oxygen vacancies of electrons and O^{2-} diffuse to the positively charged vacancy and recombine. However, in the bipolar case, O^{2-} migration is field controlled, while in the unipolar case, O^{2-} migration is thermally controlled.

L. Zhao et al. [145] proposed that the driving force for RESET changes during the process of filament rupture. In LRS, before RESET begins, a continuous filament exists. The applied voltage causes a temperature rise due to joule heating. O^{2-} then diffuses according to the concentration gradient towards the oxygen vacancy rich CF. Once a dielectric barrier (gap) is created, the electrical field dominates since current and temperature significantly decreases. Therefore, thermal diffusion dominates the RESET operation when a continuous filament is present while drift dominates when a gap is created. The SET process is similar to FORMING; however, to re-construct the continuous filament only the portion ruptured (gap) in the RESET operation has to be re-formed. Consequently, V_{SET} is lower than V_{FORM}. In HfOx memory, conduction in HRS at low bias regime has been shown to be dominated by trap assisted tunneling[146], hopping [71] and Poole-Frenkel [147]. Su et al. investigated current conduction as a function of resistance and observed that current conduction transforms from hopping to

Ohmic as resistance decreases [71]. Conduction in LRS has been shown to be metallic-like for low resistance values (< 103 Ω).

Like most other RRAM systems, the switching performance of HfOx based cells can be enhanced using different strategies, which include interface and stack engineering, doping and programming. In [59], the authors showed that implanting Gd in the HfOx layer improved switching speed, enlarged the On/Off ratio, and improved uniformity. This improvement is attributed to Gd ions decreasing the energy of vacancy formation in the oxide. This is desirable for systems in which the CF is composed of oxygen vacancies. Likewise, CFs that are composed of metal, like in ECM, can be improved by incorporating metals of the same electrochemically active electrode into the HfOx layer [148]. Oxygen plasma treatment of the HfOx has been shown to improve switching performance by incorporating interstitial oxygen ions for repeated cycling [149].

1.4 Dissertation Objective

Reliable RRAM memory operation requires that the ON (LRS) and OFF (HRS) state distributions are distinct enough to enable accurate state identification. However, the stochastic nature of the resistive switching mechanism causes incomplete or failed write/erase operations, which result in an overlap of the resistance distributions. To ensure that each operation is complete, state-verification during programming is mandatory. The primary objective of this experimental thesis work is to develop a program-verify technique purposefully to address the variability induced overlap of the state distributions. This special technique is uniquely based on utilizing low-programming energy for write/erase to promote gradual tuning and suppress over-programming – the major cause of irreversible switching failure. The combined benefit of the developed program-verify technique would therefore enable a high cycling lifetime with a

resistance memory window. This current work provides a feasible solution to one of the most challenging issues in filamentary-based RRAM devices. Prior to demonstrating this novel program-verify technique and as an integral part of this thesis work, a measurement setup was developed with two key components: (i) a novel home-made 100 picosecond pulse generator with adjustable pulse amplitude, repetition rate and pulse polarity and (ii) a custom-built control circuit board with maximum operation speed at 10 MHz.

CHAPTER 2

EXPERIMENTAL SETUP

This chapter describes the customized, pulse-programming setup called Compliance-free Ultra-Short Smart Pulse Programming (CUSPP), that was developed for studying the switching performance of RRAM devices. The setup integrates a custom-built, ultra-short, pulse generator with a custom-built control circuit. By implementing a new current-compliance approach and read operation, the programming setup enables unique, intrinsic studies of the RRAM high-speed switching characteristics. These studies include 1) the short-term (microsecond) instability of the post-programmed resistance state and 2) the intrinsic recovery behavior of RRAMs after failure. To date, some of the highest operational speeds of HfOx-based RRAM have been demonstrated with our new measurement setup: specifically, 100 ps write/erase speed and 500 kCycles/sec cycling rate have been achieved. The description is divided into three parts: (1) the custom-built 100 ps pulse generator, (2) custom- designed control circuit and (3) the fully integrated program-verify system (CUSPP).

2.1 HfO_2 RRAM Device

All RRAM devices used in the CUSPP setup were provided by SEMATECH. The oxide layer is polycrystalline HfO_2, deposited using Atomic Layer Deposition. Three different oxide thicknesses were available: 4.5 nm, 5.8 nm and 6.8 nm. The top electrode (TE) and bottom electrode (BE) is TiN. A 5 nm thick Ti oxygen exchange layer (OEL) is deposited between the top TiN and HfO_2 layer. As previously discussed, OEL are purposefully deposited to create defects in the oxide layer and lower the switching voltages. Also, the asymmetry in the stack promotes the bipolar switching mode. In this work, voltages are applied to the TE, while the BE is grounded. A large range of device sizes were available, from 2000 x 2000 nm^2 down to 20 x

20 nm^2. The device size is defined by the cross point of the top and bottom electrode because of the cross-bar architecture of the devices. Unless otherwise stated, all the results reported in this work were obtained using the 200 x 200 nm^2, 5.8 nm thick HfO_2 devices because these devices generally showed higher switching yield.

2.2 100 picosecond Pulse Generator

In a previous study, it was determined that the forming energy was just as critical to proper control of the post-programmed resistance as the forming current [73]. This determination was based on the observation that devices, which were formed with the same maximum current but different energy levels, lead to different resistance states. High energy results in low resistance, low energy results in high resistance. Therefore, a new approach was devised to limit the current duration instead of the more conventional approach, which limits the current amplitude. In the latter approach, a current limiting element, such as a transistor or resistor, is placed in series with the RRAM to limit current amplitude [51].

Since programming energy is $\int V(t) * I(t)\, dt$, the approach of this thesis uses the pulse width (PW) to restrict the range over which *dt* can vary. This range, which is the current duration of the pulse, gives the upper and lower limits for that integral. However, the current duration is not necessarily equal to the PW of the applied voltage pulse. This is because the switching event in RRAM occurs randomly at some time during the pulse [85]. This pulse variability means that both the current duration and the programming energy are also variable. It was concluded that the key to minimizing *dt* and, consequently, the energy range is to minimize the programming PW. By using the ultra-short pulse approach, the range of energy (partially pulse width dependent) delivered per pulse is limited, and the maximum energy delivered per pulse is minimized. There are two major benefits of our approach. First, resistance

programming or tuning is expected to be more gradual since the maximum energy delivered per pulse is small. Second, any potential adverse effects associated with excessive currents or energy are minimized.

Normally, RRAM devices are programmed using PW in the nanosecond range [72, 150]. However, sub nanosecond programming has also been demonstrated [151] [152] [62]. These 100 ps switching time scales suggest that the processes responsible for RS are extremely short and demonstrate the feasibility of RRAM as a replacement for DRAM and possibly SRAM. In this work, we propose to use 100 ps pulses for programming. To generate these ultra-short pulses, a 100 ps (Full Width at Half Max) pulse generator was developed using various radio-frequency (rf) and microwave components. These components include a source, diodes, switches, attenuators, an inverter, pick-off tees, power splitters and a dc block. This system was designed to generate both negative and positive pulses with tunable pulse amplitudes and rates for RRAM programming.

The 100 ps pulse is generated by a commercial impulse generator that transforms a sinusoidal wave input into a negative output pulse with 100 ps rise and fall time. The impulse generator was more suited for programming the RRAM cells because of the higher output voltages (- 18V) for sub nanosecond PW compared to commercial picosecond pattern generators.

Figure 5 (a) and (b) show the respective input and output signal of the pulse generator, commonly referred to as a comb generator. The working principle of the comb generator is based on a Step Recovery Diode (SRD). In forward bias, the P-N junction conducts (turn on) and is full of charge carriers. When the device is reverse biased, the conduction remains high due to the stored charges in the junction. However, this reverse conduction lasts only until all the stored charge is removed. After that, the conductance decreases (turn off). The doping profile in SRD is

optimized so that all the stored charge is swept out almost instantaneously. Thus, the turn-off time is extremely short. The associated sudden drop in current produces a voltage spike with a wide range of harmonics. Based on the intended use, filters are used at the output to select the desired harmonics. Both the amplitude and spacing of the voltage spikes can be adjusted by the input sine wave. Here, a dual-channel pulse generator is used as the initial source of a sinusoidal shaped signal. This signal is amplified before it is delivered as the input to the comb generator.

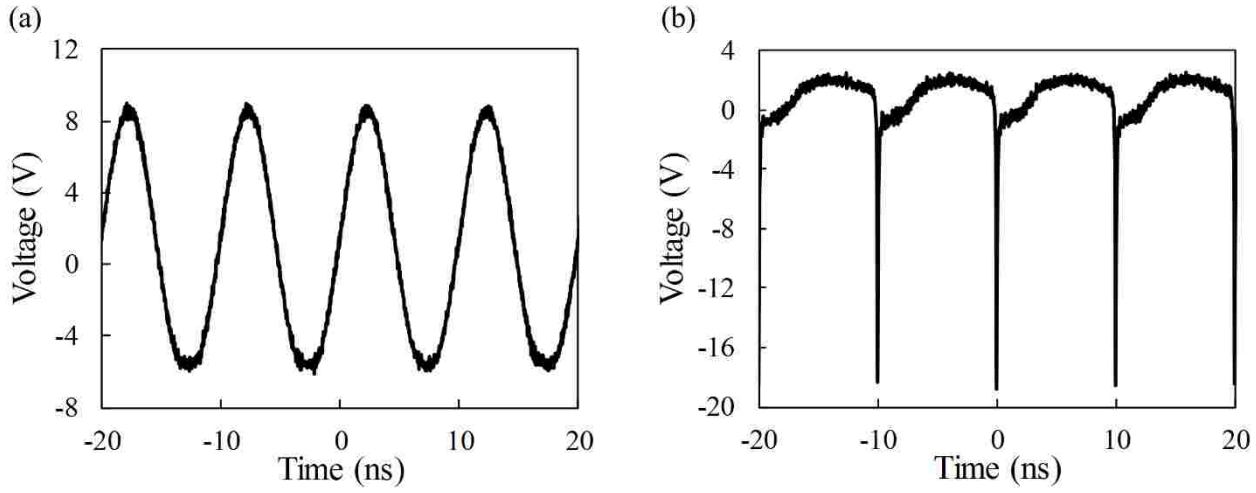

Figure 5. (a) Input to and (b) output from a step recovery diode comb generator. The comb generator outputs voltage spikes with amplitude – 18 V and PW ~ 100ps. Note, the presence of a significant residual sinusoidal signal at the output.

As shown in Figure 5 (b), the output of the comb generator contains a residual signal from the input sine wave. This signal is undesirable because it is large enough (> +0.5 V) to program or, alternatively, disturb the device resistance state. Therefore, it must be removed. The output signal goes through multiple stages of pulse shaping to achieve this. With emphasis on

preserving the voltage spike amplitudes and bandwidth (3.5 GHz), several strategies were considered and tried.

A simple solution was to use a serial connection of attenuators and high-speed diodes. Although the attenuation decreased both spike (desired) and background signal, it greatly improved the ability of the diode to “clamp” the background signal. A fast Schottky diode was used for this purpose because of its low capacitance (pF) and low minority carrier lifetime (100 ps at 80 mA forward bias). These features allowed for minimal degradation of the pulse rise and fall time. The diode, which is packaged as a chip (SOT-323/SC-70), had to be housed in a module compatible with Subminiature Type-A (SMA) connectors. This was necessary because SMA connections are made throughout the system as shown in Figure 6. The diode module is engineered by carefully drilling a hole through a SMA connector to separate the center conductor and placing the diode to reconnect the signal path. The hole size was made as small as possible to avoid reflections and distortion of the fast pulse as shown in Figure 6.

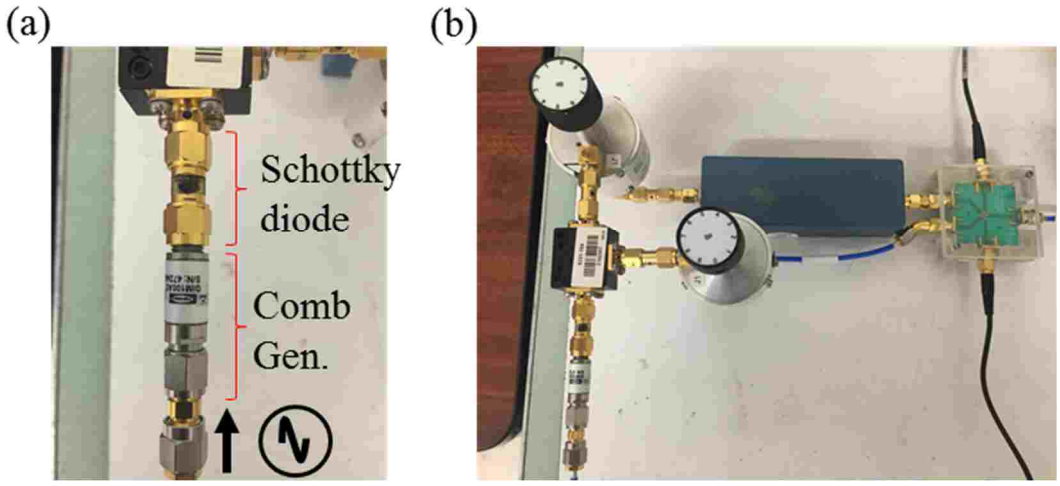

Figure 6. (a) Commercial comb generator and engineered connector to house the Schottky diode. SMA connectors are used throughout to connect individual microwave components (b) Splitter, attenuators, inverter and switch assembly used to generate independently adjustable positive and negative pulses.

A limitation of the comb generator source is that it only outputs negative pulses as shown in Figure 5 (b). However, to operate bipolar devices, both negative and positive pulses are required. For the TiN/Ti/HfO_2/TiN devices used in this work, SET and FORMING require positive polarity while RESET requires negative polarity. The voltage polarity is referenced to the BE which is grounded during programming. To produce positive pulses, an inverting stage was included. For cycling (SET and RESET) measurements such as endurance, alternating positive and negative pulses are needed. Therefore, both inverted and non-inverted signals must be produced. To achieve this, the pulse is first split into two halves. Then, a high-speed (15 ps) inverting transformer inverts one side to positive pulses. The power splitter and inverter are shown in Figure 6 (b). To allow for separate amplitude adjustment of the positive and negative

pulse trains, a digital high-speed (DC – 8 Ghz) attenuator with 1 dB step adjustments from 0 dB to 9 dB is used.

The inverter stage introduces a significant loss to the pulse amplitude because of the power splitter attenuation (6 dB) required to produce two separate pulse streams. For SET voltage (V_{SET}) and RESET voltage (V_{RESET}), the loss is minimal and does not affect programming; however, this loss is significant for FORMING voltage (V_{FORM}). V_{FORM} is typically up to 3 x larger than V_{SET} and V_{RESET}. We explored several strategies geared towards bypassing the splitter stage; this includes using a multiplexer to direct the initial negative pulse train towards an inverting or non-inverting path and amplifiers to compensate for the loss in amplitude. Although both remedies were feasible, both have their drawbacks. Because of this, we chose a different option. Since FORMING initiates cycling and is required once, we chose a simple solution: to exclude the splitter from the setup completely for just the FORMING step.

At this point, two separate pulse trains with opposite polarity have been generated on two separate lines. The signals must be recombined to construct a pulse train of alternating positive and negative pulses. However, both V_{SET} and V_{RESET} have the same clock cycle and repetition rate because they share a common source. To solve this problem, we used a high-speed switch (switch A) as a multiplexer. The switch selects between V_{SET} and V_{RESET} at a preset rate. This way, the separation (period) of the negative and positive pulse can be tuned by the switching rate of switch A. The resulting combined pulse train is sufficient for cycling the device. However, as will be discussed later, it is not adaptive and cannot avoid SET/RESET failures. In Section 2.2, we propose a complementary control circuit that uses feedback to control pulses such that failures can be dealt with.

Figure 7 shows switch A, which is a high-speed (DC – 18 GHz) single-pole double-throw (SPDT) FET switch in its customized housing. The switch is an extremely small die - 0.83 mm (L) x 1.11 mm (W) x 0.1 mm (D) - requiring wire bonds to make contact to all the pads (input, output and control) except rf ground. We used a microstrip transmission line with 50 Ω characteristic impedance, designed on a Printed Circuit Board (PCB), for this purpose. To ensure optimum insertion and return loss, single thin (0.001 in) wire bonds were made while bond lengths were kept short (~2.5 mm) so as to appear as a lumped element to the 100 ps rise and fall time.

A multi-functional heat sink was designed to make electrical contact to the rf ground located on the backside of the switch, and to provide mechanical support. The latter is important because the PCB is not rigid and therefore wire bonds can be easily detached during use. The switch is attached to the heat sink using a conductive silver epoxy, because it provides both thermal and electrical conductivity. A brass metal piece was designed using Autodesk ArtCAM and machined using a Computer Numeric Control (CNC) milling machine. The overall loss of the customized switch module is roughly -4 Db (up to 4 GHz). Although this is significant, the output voltages remain sufficient for the RRAM devices examined in this work. Because of our deliberate, transmission-line design, the output pulse after the switch maintains the 100 ps transition times.

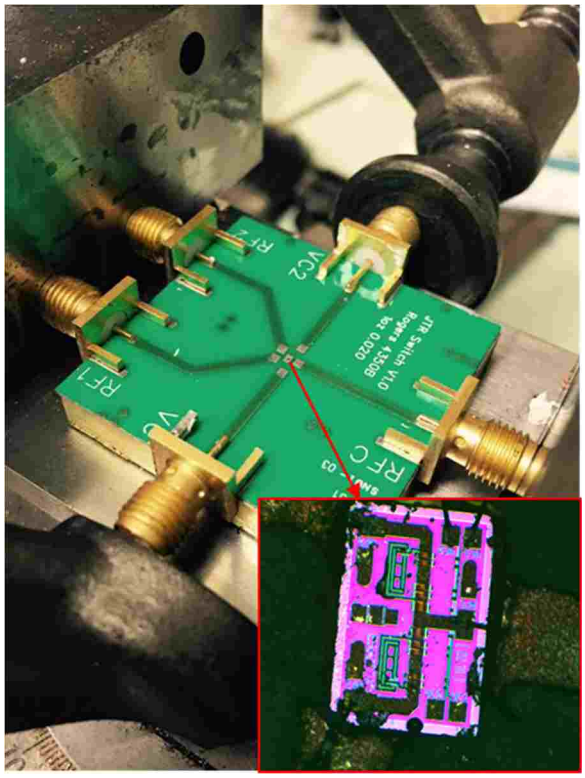

Figure 7. PCB layout for the high-speed switch with a brass base designed to provide mechanical support, electrical and thermal contact to the backside of the switch. Inset shows image of chip with wire bonds connecting signal and control pads to microstrip transmission lines.

Figure 8 shows a representative V_{RESET} programming pulse and high-speed, measurement probe. The noticeable offset (indicated by the red dashed line in Figure 8 (a)) is supplied by a programming DC source through a pick-off tee. The low DC voltage (V_{READ}) is used to read/sense the state of the device during programming. A RRAM is a two-terminal device meaning that both program and read operations are performed using the same two terminals. This can cause what is called “read disturb”. To minimize the risk of read disturb, the read operation is commonly performed using a low voltage pulse. Because of this, the read pulse amplitude and length must be optimized to ensure that they can determine the resistance state of

the RRAM accurately. Typical read-pulse durations are in the microsecond to nanosecond range and read-pulse amplitudes are 10 x lower than the programming pulse [87, 153, 154]. Positive or negative read polarity can be used, although some novel read operations utilize a combined positive and negative pulse read to avoid any dielectric polarization effects [155]. In this work, negative read polarity is used always.

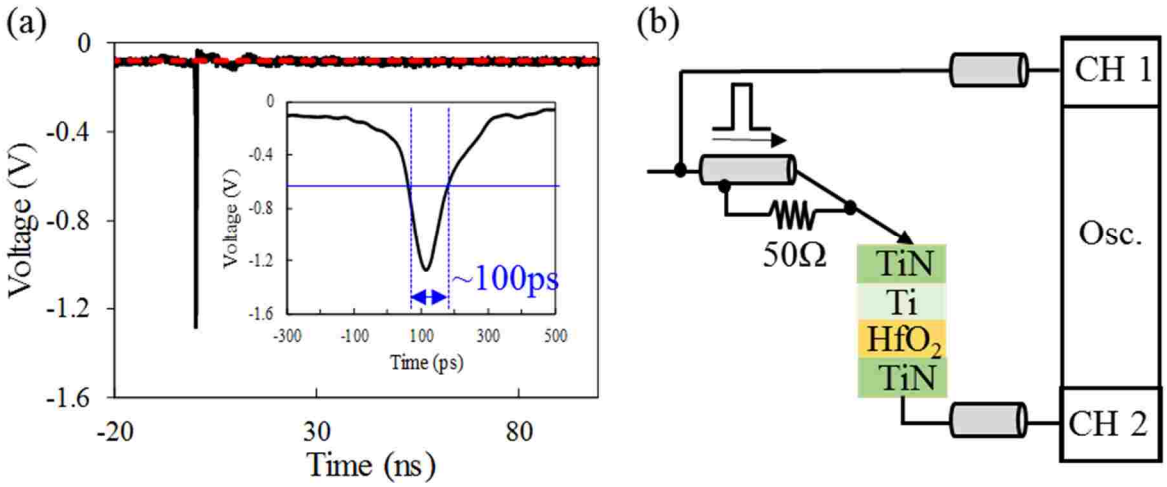

Figure 8. (a) Typical V_{RESET} pulse offset by a low DC voltage(V_{READ}) used for read. (Inset) Pulse FWHM is approximately 100 ps captured on a high-speed oscilloscope (Rise time contributions of osc. has not been deconvolved). (b) High-speed measurement setup with proper 50Ω termination directly at the probe tip. Notice the absence of a series current limiting element.

The DC read operation adopted in this work is based on traditional, electrical stress measurements that utilize constant voltage to study stability. Typical voltage amplitudes used in such measurements are > |0.1| V. Here, read voltages are usually < -0.1 V. For even longer V_{READ} duration (> 10^3 s), voltages as low as -0.008 V are used. Stable LRS and HRS have been demonstrated in HfOx RRAM cells using +0.3 V for 10^3 sec [156]. Therefore, we do not expect

to disturb the resistance states using the constant voltage V_{READ}. By utilizing a DC read during programming in CUSPP, we essentially combine a stress measurement with cycling for stability and endurance studies, respectively. As will be discussed in Chapter 4, fluctuation behavior is observed in the short-term post programming in a manner that cannot be achieved using pulse read.

Delivering fast pulses to the RRAM cell without rise-time degradation and signal distortions requires transmission-line design at all stages of the setup. The entire system, therefore, is designed to have a 50 Ω characteristic impedance and to support frequencies higher than 3.5 GHz (100 ps rise). These requirements are maintained throughout the system but particularly at the launch site (probe needle to RRAM) where reflections and ringing can distort switching analysis. Traditionally, a Ground-Signal-Ground (GSG) probe is used for such high-speed measurements [151]. However, this requires a matching, probe-pad layout of the device under test (DUT). Since the devices in this work were not designed to fit a GSG probe setup, this is not a feasible option. Thus, we had to design a high-speed probe based on a fundamental transmission line to assure high signal fidelity.

Since a RRAM is a two-terminal device, two high-speed probes are sufficient for operation. To design the high-speed probes, three main strategies were employed. First, the probe needle is made as short as possible to appear as a lumped element to the fast, pulse-transition time. Second, a 50 Ω rf resistor is strategically placed in parallel to the RRAM as shown in Figure 8 (b) to guarantee an impedance matched load. This assures that for all targeted RRAM resistance values, which typically range from 10^3 Ω to 10^6 Ω, the signal "sees" a 50 Ω termination. The 50 Ω termination also provides a fast RC to discharge any parasitic capacitance quickly. This is important because parasitic capacitance, which is ubiquitous in any setup, would

delay the response of the system to any abrupt resistance change. Any such delay is detrimental to our goal of minimizing programming energy. Third, a short ground loop connecting both input and output probe tips is soldered to minimize inductance. Total inductance at the probe tip is minimized by the mutual inductance of the signal and ground path. Without this mutual conductance component, total inductance increases and causes severe signal distortion.

Data from the probes are sent to an oscilloscope where two output currents are measured: one resultant from the ultra-short programming pulse (I_{PULSE}) and the other from the read DC voltage (I_{READ}). I_{PULSE} contains switching current information, which can be used to deduce switching energy and temperature. Figure 9 (a) shows the measured current from the RRAM cell before ($I_{PRE\text{-}F}$) and after ($I_{POST\text{-}F}$) FORMING. Note, both current waveforms resemble a typical displacement-current waveform. This is a result of the device capacitance inherent in its capacitor-like structure. The displacement current waveform is modified when the device undergoes a resistance change as shown by the decrease from $I_{PRE\text{-}F}$ (blue) to $I_{POST\text{-}F}$ (red) in Figure 9 (a). Because identical pulses are used during FORMING, the change in the displacement current is due to current flowing through the cell. Therefore, I_{PULSE} contains both the displacement current and device current (I_{DEVICE}). Pragya et al. developed a technique to successfully extract I_{DEVICE} [157]. Figure 9 (b) shows the extracted I_{DEVICE} using similar methodology.

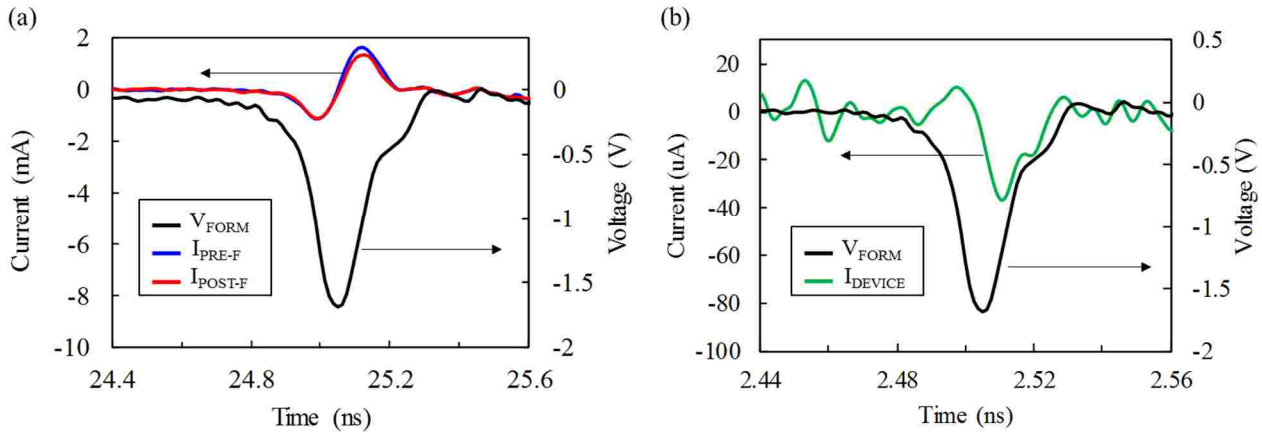

Figure 9. (a) Displacement current before (blue) and after (red) the RRAM forming event. The applied voltage is shown in black. (b) The extracted I_{DEVICE} which flows through the device.

The maximum deliverable pulse amplitudes for FORMING, SET and RESET are +5 V ($V_{FORMING}$), +2 V (V_{SET}) and – 2.77 V (V_{SET}), respectively. The lower V_{SET} amplitude compared to V_{RESET} is due to additional attenuation by the inverter (-0.6 dB). For the devices studied in this work, these pulse amplitudes are sufficient. In fact, the voltages are usually attenuated to avoid over-programming, which degrades performance. The digital step attenuator described earlier is frequently used to decrease the pulse amplitude during programming. However, this attenuator can only provide 1 dB step change to the pulse amplitude. For finer attenuation, an analog attenuator is also included in the setup.

2.3 CUSPP Control Circuit

Nonuniformity of the post-programmed resistance states remains a tremendous hurdle to overcome in RRAMs. As discussed in Chapter 1, the stochastic nature of the resistive-switching mechanism creates this non-uniformity. Resistance states in RRAMs are therefore represented as a distribution instead of a fixed value. For reliable operation, the resistance distributions of each state must be sufficiently separated to avoid ambiguity during operation. Because the resistance

change is not deterministic, programming without some sort of verification cannot guarantee separated distributions.

To provide this verification, we designed a custom, stand-alone, active, feedback-control circuit to complement the 100 ps pulse generator setup (see Figure 10). The control circuit is designed with high speed components to achieve a read resolution of 100 ns and a comparator speed up to 100 MHz. The main stages of the circuit include: a three-stage amplification, a comparator, switch and a Series Voltage Reference. The populated CUSPP circuit board is shown in Figure 11.

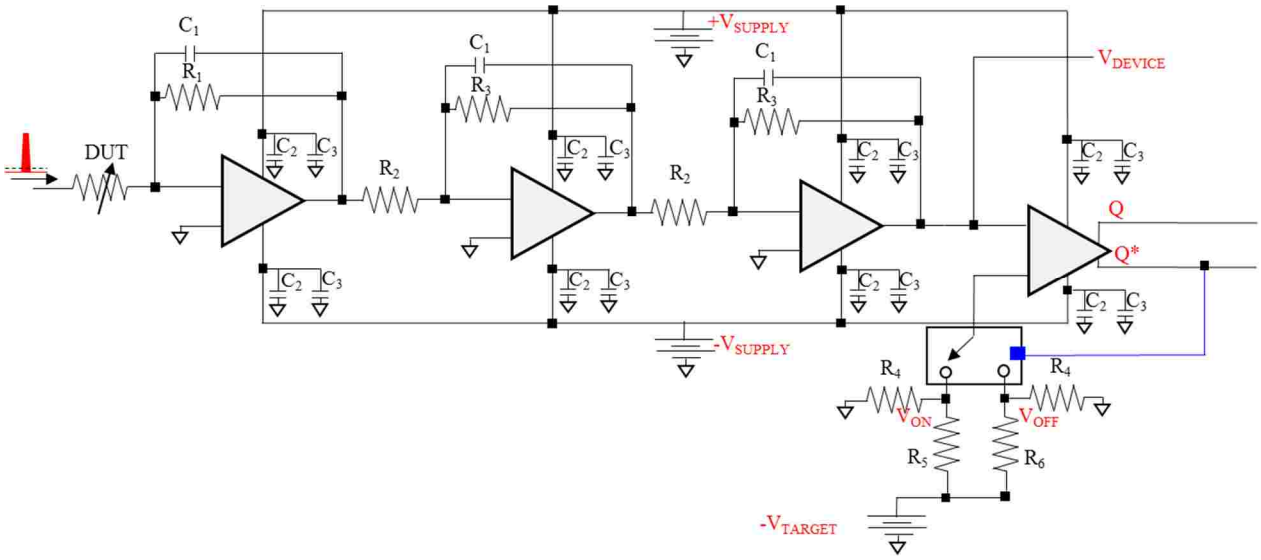

Figure 10. Schematic circuit diagram of CUSPP control circuit. Specific R values are dependent on the desired amplifier gain, V_{ON} and V_{OFF}. Feedback capacitor C1 (18 pF) is used to decrease bandwidth, C2 (0.1 µF) and C3 (1 µf) are bypass capacitors. Feedback control to switch B is shown with blue colored line.

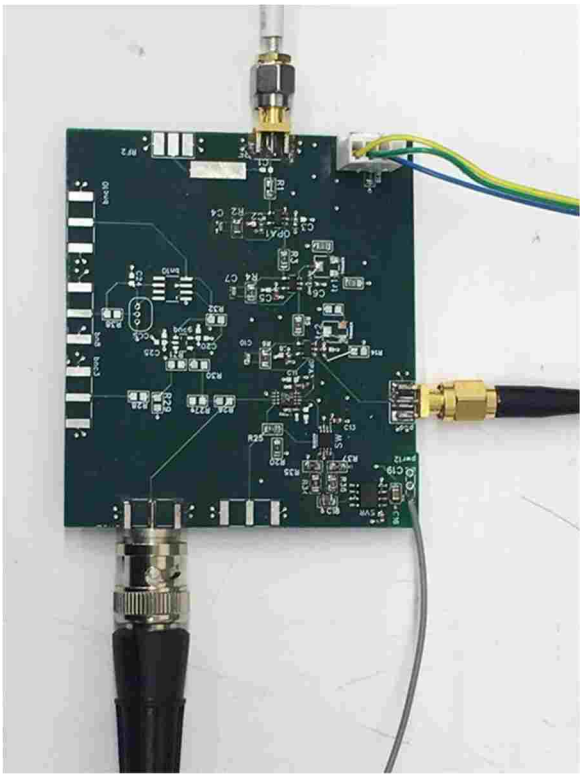

Figure 11. Picture of CUSPP control circuit.

The first function of the circuit board is to measure and to amplify the read-out current from the DUT. This read-out current is due to the DC voltage read discussed in Section 2.1. Since the RRAM displays fast switching transients during programming and read, high-speed operational amplifiers (230 MHz Gain Bandwidth product) are needed. In addition to the bandwidth requirement of the op amps, we included noise considerations as well. Program-verify algorithms rely on an accurate reading of the device's resistance state to determine the appropriate action. Hence, it is critical to minimize the noise at the amplifier stage. The amplifier choice is, therefore, based on its noise parameters such as input voltage noise (7 nV/(Hz)$^{1/2}$) and input bias current (2 pA). Additionally, large-value resistors should not be used in the inverting

amplifier circuit to minimize the thermal (Johnson) noise component, JN of the total noise at the amplification stage:

$$JN = \sqrt{4kTBR}, \qquad (2.1)$$

where k is Boltzmann's constant, T is temperature of resistor, B is bandwidth of measurement and R is resistor value.

To maintain sufficient gain and decreased noise, we implemented a three-stage amplification system. The first amplifier performs a trans-impedance amplification while the second and third amplifiers perform voltage amplification. The read-out current is converted to a voltage (V_{DEVICE}), which is measured on the oscilloscope. Since V_{READ}, Gain and V_{DEVICE} are known, the resistance of the DUT can be calculated. Note, we also accounted for noise gains due to amplification. For each amplifier, we used a bypass procedure for the power supply to support the high-speed operation. We implemented this procedure by carefully placing bypass capacitors close to both power supply pins of each amplifier in a manner to minimize impedance.

As previously stated, the output voltage measured at the oscilloscope contains two current signals; I_{PULSE} and I_{READ}. The former is due to the programming pulse and the latter is due to the DC read voltage. Since this work focusses on I_{READ}, the control circuit is designed purposefully to filter out I_{PULSE}. To achieve this, the amplifier bandwidth is decreased by using feedback capacitors of appropriate value. This strategy also impacts the fast transients in the I_{READ} signal. Still, we maintain an acceptable fast RC time of 30ns in this work.

State verification during program-verify is performed by a high-speed comparator with fast propagation delay (4.5 ns), fast rise and fall time (2 ns) and dual output. The comparator

functions to output logic low (0 V) or high (+ 5V) when V_{DEVICE} exceeds or falls below a preset target voltage level referred to here as V_{TARGET} (see Figure 10). V_{TARGET} represents the target resistance level of the RRAM. In this work, V_{DEVICE} is applied to the inverting input (-IN) of the comparator while V_{TARGET} is applied to the non-inverting input (+IN). The dual outputs, Q and Q* (complementary output), supply control signals to two separate switches, an off-board switch (switch A) and an on-board switch (switch B).

As described earlier, switch A determines between SET and RESET operation because it controls whether V_{SET} or V_{RESET} is applied to the DUT. Since the comparator controls switch A, it serves as the arbiter to select between V_{SET} and V_{RESET} based on the comparison of V_{DEVICE} to V_{TARGET}. To enable dual state verification, V_{TARGET} has two levels: V_{ON} representing the LRS target resistance level (LRS_{TARGET}) and V_{OFF} representing the HRS target resistance level (HRS_{TARGET}). V_{ON} and V_{OFF} are supplied by a series voltage reference which outputs a low noise +2V DC. A resistor network is used to adjust V_{TARGET} and consequently, LRS_{TARGET} and HRS_{TARGET}.

Figure 12 summarizes the control waveforms and the resultant V_{SET} or V_{RESET} pulse trains using the combined pulse generator and circuit setup. In a closed-loop fashion, the custom circuit monitors the device state and provides the control logic that determines both the polarity and the number of pulses required to SET the device to LRS_{TARGET} and RESET the device to HRS_{TARGET}. To enable automatic cycling between HRS and LRS, we included a second switch (switch B) in the circuit (Figure 10). Switch B toggles the voltage at the comparator +IN input between V_{ON} and V_{OFF}. For example, during RESET, when the target resistance level is set by V_{OFF}, V_{RESET} pulses are applied to the DUT until V_{DEVICE} reaches V_{OFF}, then the comparator changes logic level. Since the comparator output controls both switches; both switches change

their outputs. Switch A changes V_{RESET} pulses to V_{SET} pulses, and switch B changes V_{OFF} to V_{ON}. This initiates the SET operation. The CUSPP protocol adaptively changes the pulse number (identical pulses) until the target state is reached. Once reached, the pulse polarity changes to reverse operation. This sequence can be performed within 10 ns - enabling high cycling rate measurements.

The CUSPP program-verify algorithm utilizes identical pulses for programming since their amplitudes and widths are fixed. This type of algorithm is known to be effective in resistance tuning [145]. Most other algorithms automatically change the pulse amplitude and width in predefined increments during resistance tuning [85, 86, 154]. Although these algorithms offer an additional degree of control, they are more complex to implement.

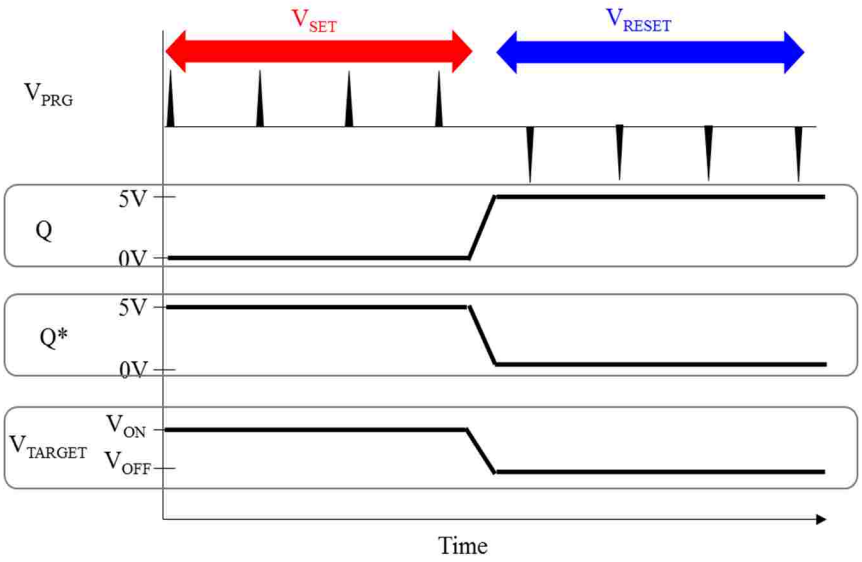

Figure 12. Timing diagram schematic of control signals and programming pulse during CUSPP operation. The high-speed design of the control circuit enables transitions times less than 10 ns during SET and RESET cycling.

2.4 Compliance-free Ultra-short Smart Pulse Programming (CUSPP)

Traditionally, FORMING is performed with constant voltage bias or sweep while cycling is performed with pulses [86, 158]. Current-based sweep forming has also been demonstrated [74]. Although sweep-mode forming is common, these programing modes are impractical. Figure 13 shows the post-formed resistance of several devices using the 100 ps pulse. Typical V_{FORM} values are in the range of +2.5 V to +4V. Initially, all pristine RRAM devices have very high resistance (> 10^{10} Ω) values and require a forming operation to initiate cycling. The resistance change during forming can span 9 orders of magnitude. Under V_{FORM} an abrupt resistance change randomly occurs (like dielectric breakdown) and the process proceeds until the driving forces, such as voltage, current, and temperature, are terminated. Figure 13 shows the distribution of post-FORMED resistance states. The ability to program to such low values using the 100 ps would suggest that even shorter pulse widths could be sufficient to provide tighter control of the resistance change.

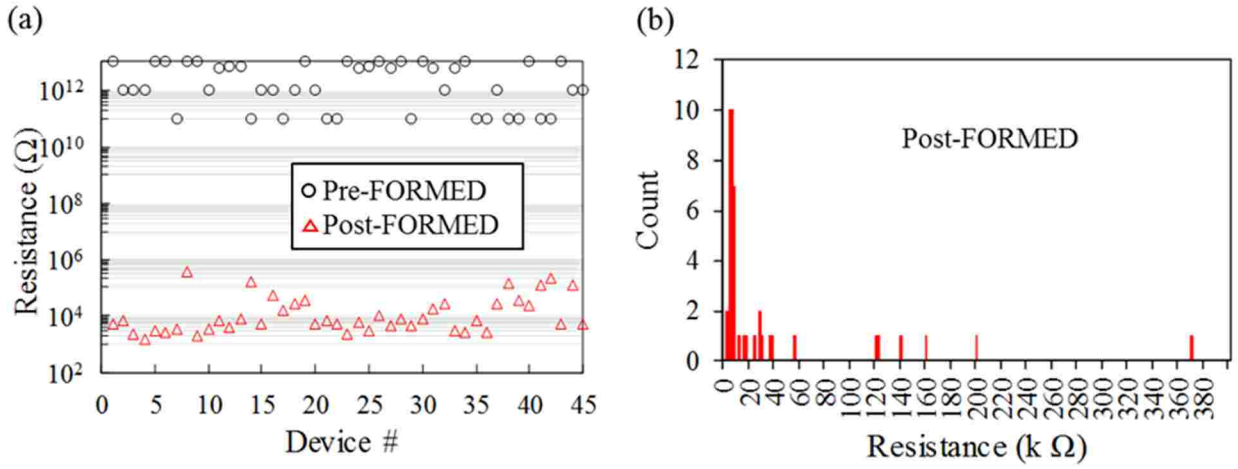

Figure 13. (a) Typical Pre-FORMED and Post-FORMED resistance levels. All devices are FORMED with a 100 ps pulse with VFORM ranging from +2.5V to +4 V. (b) Resistance distribution of Post-FORMED devices.

Figure 14 shows a representative normal RRAM resistance evolution using CUSPP. To program a 10x resistance window, the target LRS (R_{ON}) and HRS(R_{OFF}) are 10 kΩ and 100 kΩ respectively. The target levels are set by the value of V_{ON} and V_{OFF} as described in 2.2. Two distinct operational modes are shown in Figure 14: single pulse and multi-pulse. In its single-pulse operation (Figure 14 (a)), the device reaches the LRS_{TARGET} and HRS_{TARGET} after a single V_{SET} and V_{RESET} pulse, respectively. Therefore, both SET and RESET require single pulses. This single-pulse behavior would suggest that there is no need for a read-verify since both targets are met after each pulse. However, the randomness of the resistive switching mechanism can result in programming failures. This is observed in Figure 14 (b) where the device resistance requires multiple pulses to reach HRS_{TARGET}. Notice that the pulse (schematic) amplitude is unchanged in either the single pulse mode or multi-pulse mode.

Assume in Figure 14 (b) that program-verify is not activated, the resistance (~ 10 kΩ) after the first RESET pulse would be recorded as HRS, which is essentially the same as the target LRS. This would produce an error since LRS and HRS cannot be distinguished. However, CUSPP adaptively increases the V_{RESET} number until the resistance reaches 100 kΩ. During RESET, the device resistance might transition to other states before reaching HRS. One such state, referred to as Intermediate Resistance State (IRS) is shown in Figure 14 (b). This exhibits the multilevel cell (MLC) capability of RRAMs. If programming was stopped after the second V_{RESET}, a third state would be addressed. Figure 14 also illustrates the usefulness of a DC read. The resistance stability between each programming pulse is continuously monitored.

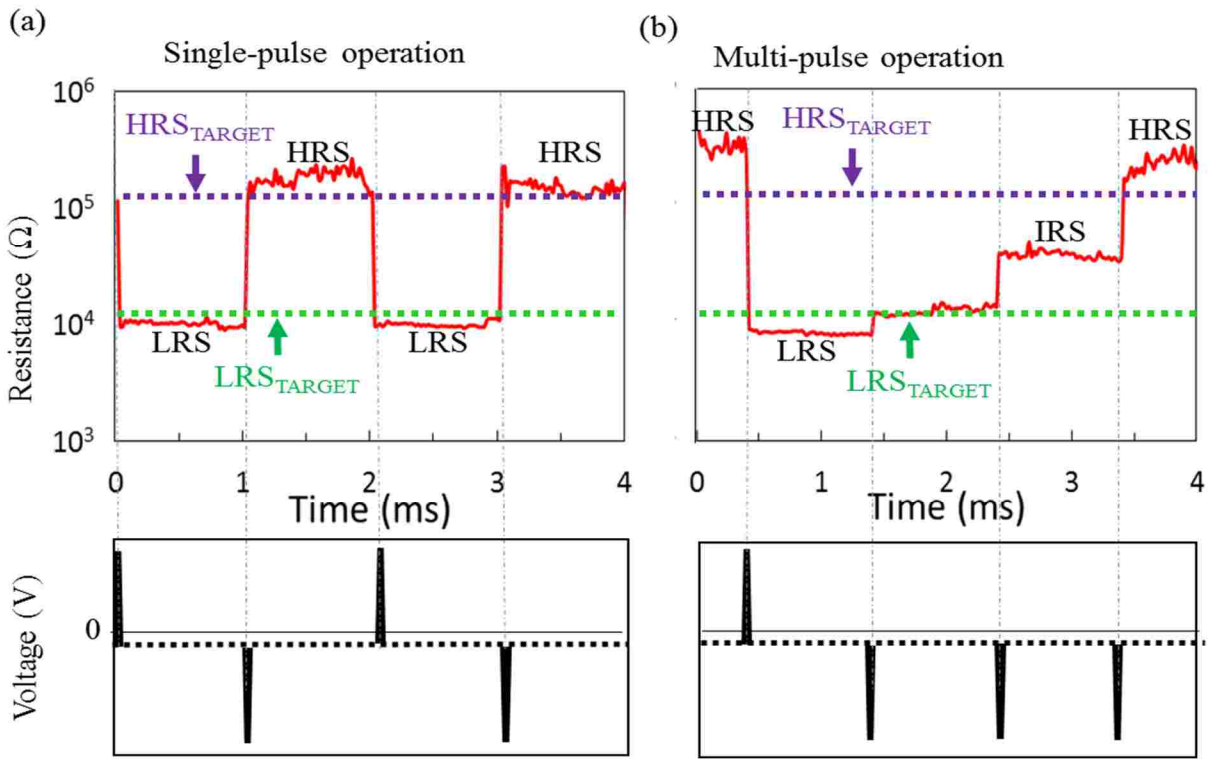

Figure 14. Representative time traces of the resistance evolution (red) during (a) single pulse operation and (b) multi-pulse operation modes. In the same device, the randomness in the resistive switching mechanism results in resistance changes which are not repeatable from cycle to cycle. However, program-verify guarantees that each operation is complete.

Figure 15 shows the distribution of LRS and HRS obtained after 2000 cycles. Since a low DC voltage is used during the read operation in CUSPP, each state is read for a duration determined by the repetition rate of the programming pulse. For example, in Figure 14, the read duration is 1 ms as determined by the 1 KHz repetition rate of the programming pulse train. Therefore, the resistance distributions can be plotted as a function of time, t. In Figure 15, the resistance values are chosen at t = 0. At t = 0, the device resistance reaches the target level signifying the completion of the SET or RESET operation and start of read. Therefore, as expected, each state begins or ends at the predefine target level; for LRS, this level is 10 kΩ and

for HRS, this level is 100 kΩ. Therefore, a 10 x RW is obtainable using CUSPP. This translates to a proportionate read-out current demonstrating the usefulness of program-verify techniques to guarantee sufficient separation of the memory states. As oppose to a conventional read operation which uses voltage pulses, CUSPP also obtains a complete recording of the resistance during read. By simply choosing the appropriate time delay, we can mimic a delayed pulse read. In Chapter IV, we discuss the usefulness of a constant read in a post-programmed stability study of RRAM.

As discussed in Chapter 2, high-resistance states (thinner filaments or tunneling gaps) are highly dispersed compared to low-resistance states. Program-verify does not function primarily to tighten the resistance distribution. However, Puglisi et al. demonstrated a dispersion-aware program-verify algorithm designed to tighten the distribution [86].

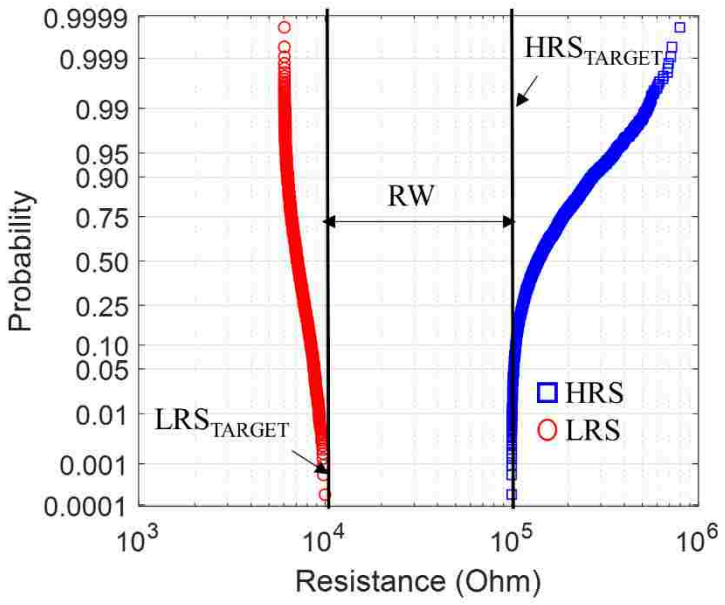

Figure 15. Probability distribution of LRS and HRS obtained after 2000 cycles using CUSPP. The resistance values are taken at t =0 (start of LRS and HRS). The HRS distribution is bounded by HRS_{TARGET} (100 kΩ) and LRS distribution is bounded by LRS_{TARGET} (10 kΩ).

In Chapter 4, we examine the resistance stability of the post-programmed states at microsecond to millisecond time delays from the programming event. The novel read operation adopted in this work, provided critical information in determining the cause of instability during programming. Utilizing a low DC read, we obtained a continuous record of the resistance evolution during programming at enhanced time resolutions. This enabled us to understand the impacts of that evolution on the post-programmed distribution and on program-verify algorithms.

Another unique feature of the CUSPP setup is the data-acquisition procedure used in this work. In this work, in a typical high-endurance measurement, the data is acquired on the oscilloscope 10 million samples at a time. The acquisition time is determined by the sampling rate of the oscilloscope (acquisition time = (# of samples) / (sampling rate). When this time elapses, the oscilloscope stores the data and re-arms for acquisition. A LabVIEW program was written to control the number of acquisitions performed by the oscilloscope. A downside of the procedure is that, during the time the oscilloscope is storing data, data acquisition is disrupted. Data acquisition is, therefore, performed in time blocks that are separated by times when acquisition is unavailable. However, statistically speaking, the extremely large number of endurance measurements tends to minimize the effects of these dead times.

Figure 16 shows 10^6 endurance measurements of an RRAM programmed with CUSPP. To construct the endurance plot shown in Figure 16 (a), only the HRS and LRS values are needed. However, this representation, which is popular in RRAM studies, does not show any information on the cycling dynamics. In this work, every cycle state is recorded (within the acquisition time) because of the acquisition technique used. We observe the device resistance evolution after each programming pulse regardless of the success of the operation.

This allows for detailed analysis of the switching dynamics during cycling as shown in Figure 16 (b), where a failure occurred because the device failed to reach the LRS_{TARGET} (3 kΩ) within the normal time or number of pulses. During this failure time, CUSPP continues to apply V_{SET} until the SET criteria is met. Eventually, this occurs and cycling commences. The recovery behavior is examined further in Chapter 3. As will be discussed, no noticeable degradation of the switching behavior was observed after recovery. These findings, not only highlight characteristics of RRAMs but also demonstrate the effectiveness of CUSPP to program and to recover switching.

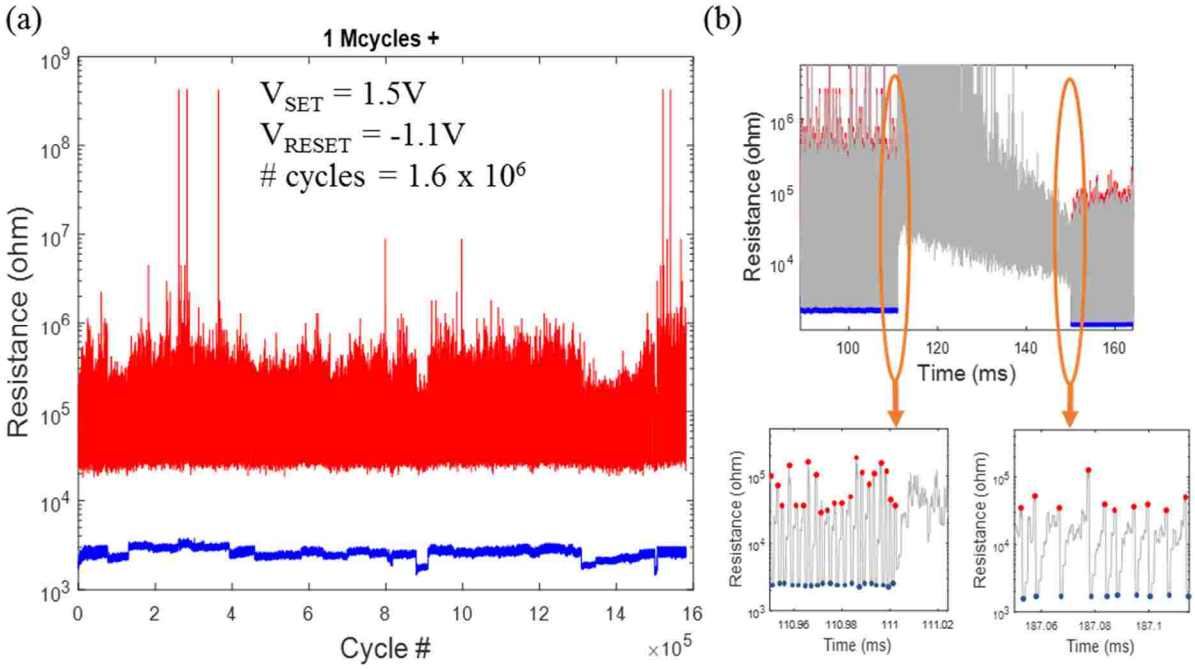

Figure 16. (a) 10^6 endurance between 3 kΩ and 30 kΩ target LRS and HRS, respectively using CUSPP. Each cycle is shown to demonstrate the functionality of the program-verify algorithm (b) Time trace of resistance evolution during programming. The device suddenly fails to SET but after sustained pulsing, it recovers and continues to cycle.

2.5 Summary

In this chapter, we proposed a program-verify technique called Compliance-free Ultra-short Smart Pulse Programming (CUSPP). CUSPP relies on a novel, current-control scheme to achieve low-programming energy operation. CUSPP is a custom-built, pulse-train, measurement system with two integral parts: an ultra-short (100 ps) pulse generator and a custom-built control circuit. The 100 ps pulse generator was realized by assembling several rf components together to achieve dual-polarity short pulses with an adjustable pulse rate and amplitude. The control circuit was designed to actively monitor the device resistance state and provide feedback control to the pulse generator. The CUSPP programming protocol adaptively changes the number and polarity of the pulses to switch the device resistance between two distinct preset targets. Because of the high-speed design of the control circuit, pulse rates with state verification up to 10 MHz have been achieved. Therefore, CUSPP effectively demonstrated the high-speed (switching speed and cycling rate) operation capability of HfO_2-based RRAMs previously unreported.

Two additionally unique features of the CUSPP measurement setup are the absence of a series current limiting element such as a resistor or transistor [29] and the use of a low DC voltage for read. Current control was accomplished by utilizing the ultra-short pulse width of the voltage pulse to limit the current duration. By doing so, the energy delivered per pulse is greatly minimized, thus minimizing programming induced variability in RRAMs [73]. The low DC voltage read approach utilized in this work is a deviation from the more traditional pulse read operation [153]. However, its usefulness was demonstrated as it provided a continuous record of the state evolution during programming. This enabled unique analysis of the post-programmed resistance state instability.

CHAPTER 3

CYCLING ENDURANCE

In this chapter, we used the newly developed CUSPP technique to perform cycling endurance studies on HfO_2-based RRAM devices. The endurance performance and analysis serves as an indicator of both the performance of the memory stack and the adopted programming technique. We achieve 10^8 endurance with state verification using CUSPP, thereby demonstrating its ability to sustain reversible resistive switching. Due to the randomness of the resistive switching process, unavoidable switching failures occur during extending cycling. However, we show that the failures are reversible. The recoverable failures are attributed to the delicate nature of the CUSPP technique, which avoids permanent failure.

Typically, high endurance is achieved by optimizing the programming pulse parameters for an individual device and avoiding failure completely. However, to achieve high endurance in a memory chip, each memory cell requires its own unique optimum conditions, which is impractical. In this chapter, we propose a possible solution based on the analysis of the switching performance pre- and post-recovery. We observed that the recovery operation does not degrade the switching performance of the device. We conclude, therefore, that recovery may be deployed indefinitely to achieve a high endurance yield. Our findings are in stark contrast to previous studies on recovery behavior where the device degrades after DC [68] and pulse recovery [159, 160].

3.1 Experimental Details

We generated experimental endurance measurements on 200 nm x 200 nm crossbar RRAM devices with TiN\HfO_2\Ti\TiN stacks. Ultra-short programming pulses (~100 ps) were applied directly to the device with a 50 Ω terminated probe but without a current limiting

element connected serially to the RRAM. Programming and read voltages were applied to the top electrode with the bottom electrode grounded. Devices were FORMED with voltages ranging from +2.8V to +3.2V.

A programming pulse rate of 500 kHz was used in these experiments. Thus, the read duration was limited to 2 µs with a measurement speed of 10 MS/s set by the sampling rate of the oscilloscope. A DC read is utilized instead of the conventional pulse read. The target HRS and LRS were 100 kΩ and 10 kΩ respectively; we chose these values to achieve a 10 x resistance window (RW). The CUSPP programming protocol continuously pulses the RRAM with identical pulses (fixed voltage and width) until the device resistance reaches the HRS/LRS target during RESET/SET. The bipolar HfO_2 RRAM devices require negative polarity to RESET and positive polarity to SET.

3.2 Results and Discussion

Figure 17 shows the endurance of a representative cell. 10^8 endurance was achieved with V_{SET} and V_{RESET} set to +1 V and -1 V, respectively. The device did not fail after 10^8 cycles; instead, the experiment was stopped. While not every cycle is shown in Figure 17, each cycle is verified. It is common practice not to plot every single cycle because of data-size limitations and the difficulties in displaying billions of data points/cycles on a figure. However, this representation may not be valid for programming protocols without state verification as it cannot be assured that endurance failure did not occur. A failure in cycling is defined by the minimum acceptable RW, which sets the read-out margin. Without state verification, it is incorrect to assume that each cycle met this criterion.

On the other hand, in program-verify programming, each cycle starts at the intended target value; therefore, endurance represents a true count of the number of successful cycles. Program-verify (P-V) algorithms usually change pulse number [43, 84], amplitude [86, 87] or width [85] until the targets are met. A failure using P-V is defined by the number of iterations or time required to successfully program the device. Figure 17 also shows that the RW is lower (5x) than the target 10x window. As will be discussed in Chapter 4, this is attributed to the instability of the post-programmed state. HRS dominates the window closure since non-uniformity of the resistance states increases with increasing resistance [37].

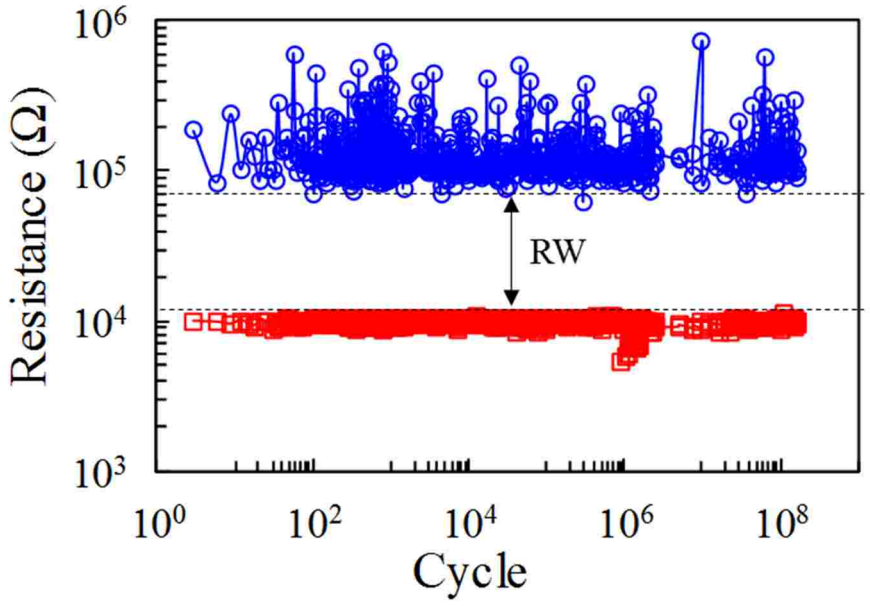

Figure 17. Endurance performance of the bipolar HfO_2 RRAM device using CUSPP. V_{SET} and V_{RESET} were +1 V and -1 V respectively and read voltage was -0.05 V.

As described in Chapter 2, data acquisition is performed using an oscilloscope in 1 second long acquisition times. The pulse rate was 500 kHz, which means that for single pulse operation, where each SET/RESET requires a single pulse, the maximum attainable cycling rate is 250 kcycles/sec. However, because of the innate randomness of the resistive switching process in RRAMs, each operation may require multiple pulses (multiple pulse operation), in which case, the cycling rate is less than 250 kcycles/sec.

This is indeed the case in this endurance study. Figure 18 shows the cycling dynamics of the RRAM programmed with CUSPP. It is apparent that a complete cycle required multiple pulses, which significantly decreased the cycling rate. We believe that the specific rates can be influenced by V_{SET} and V_{RESET} amplitudes since switching time is exponentially dependent on voltage [113, 161]. The drawback of increasing the programming voltage is the increased probability of excessive currents and over-programming, which has been shown to degrade endurance [69]. However, low-pulse operation (multi-pulse operation) is desirable because resistance tuning is more precise [154]. Thus, it is conceivable that per pulse amplitude, there is a preferred cycling rate for a given device. Unfortunately, this preferred rate is clearly not constant as shown in Figure 18.

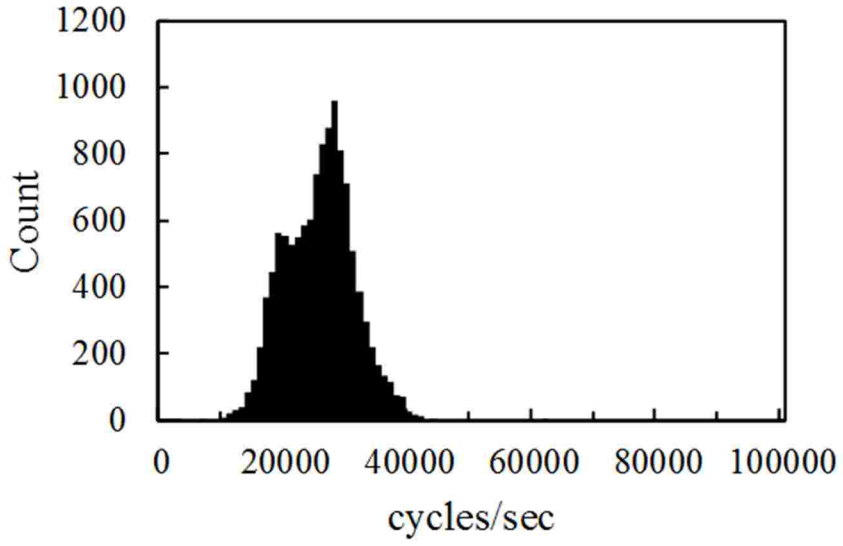

Figure 18. Cycling rate distribution acquired for 12,000 sec of cycling using CUSPP.

Figure 19 shows the temporal evolution of the cycling rate, which does not appear to degrade or improve over time. As suggested earlier, there seems to be an optimum cycling rate for the device that is approximately 10 x lower than the single-pulse limit (red line). It is tempting to assume that the cycling rate is primarily controlled by V_{SET} and V_{RESET}; if true, this would mean that other devices would display the same dynamics for fixed programming conditions. However, because of the device-to-device variability in RRAM arrays, this is incorrect. For this device, 28439 cycles/sec (mode) is considered the optimal cycling rate; from here on out, we will refer to this optimal rate as the normal rate. During the lifetime of the cycling measurement, large excursions from this normal rate can be observed. In some cases, the cycling rate decreases by two orders of magnitude (see Figure 19).

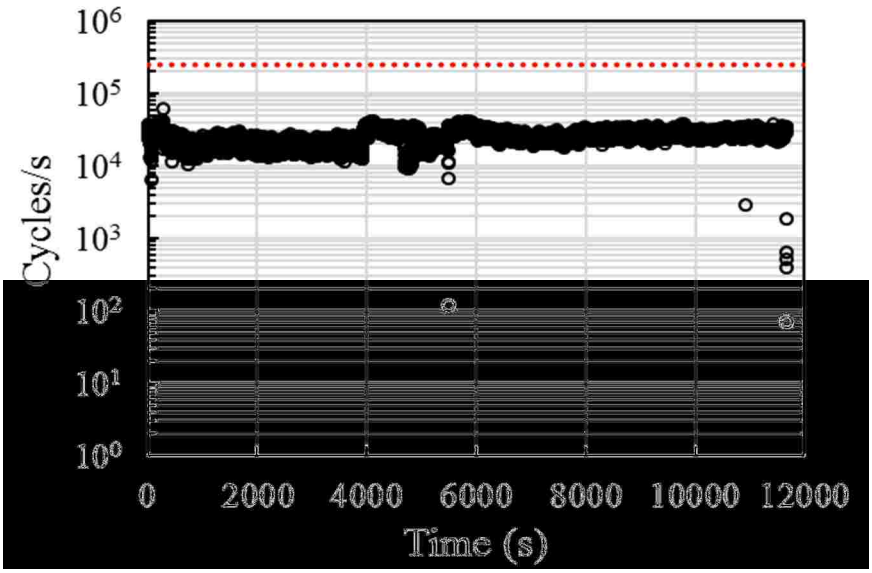

Figure 19. (a) Evolution of cycling rate during the endurance study. Red dashed line indicates the maximum attainable cycling rate of 250 kcycles/sec for a single pulse operation mode. Lower cycling rate indicates multi-pulse operation.

Figure 20 shows the distribution of the number of pulses for V_{SET} and V_{RESET} during normal operation. The SET operation requires fewer pulses (mode = 2) compared to the RESET operation (mode = 16). This suggests that the RESET operation is the rate-limiting process of the overall programming time. In addition, the RESET operation also shows more variability since the V_{RESET} pulse number is more dispersed. SET is abrupt under voltage operation because of the interplay of an increasing electric field, current, and temperature. Together, they create a positive feedback, which subsequently leads to a dielectric breakdown event. On the other hand, RESET is more gradual under voltage operation [74].

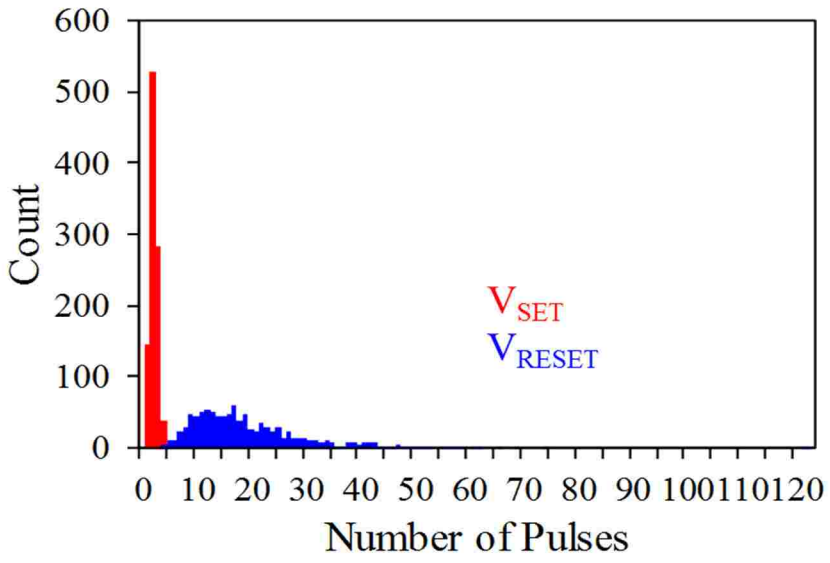

Figure 20. Distribution of number of V_{SET}/V_{RESET} pulses during normal operation for 1000 cycles. SET is abrupt while RESET is gradual and therefore shows a more dispersed distribution.

Figure 21 shows the experimental resistance distribution of LRS and HRS programmed during normal operation. The HRS distribution clearly shows the effect of the CUSPP program-verify on the post-programmed resistance states. For HRS, all resistance states are forced to be above the 100 kΩ resistance target. For LRS, all resistance states are forced to be below the 10 kΩ resistance target. Achieving these targets is the primary function of program-verify techniques. Moreover, these techniques ensure sufficient separation of the tail of both LRS and HRS distributions. However, because of short-term instability of the post-programmed resistance state, some states fall below and other states fall above their respective targets. This results in decreasing the memory window in that process. We discuss fully this short-term instability in Chapter 4.

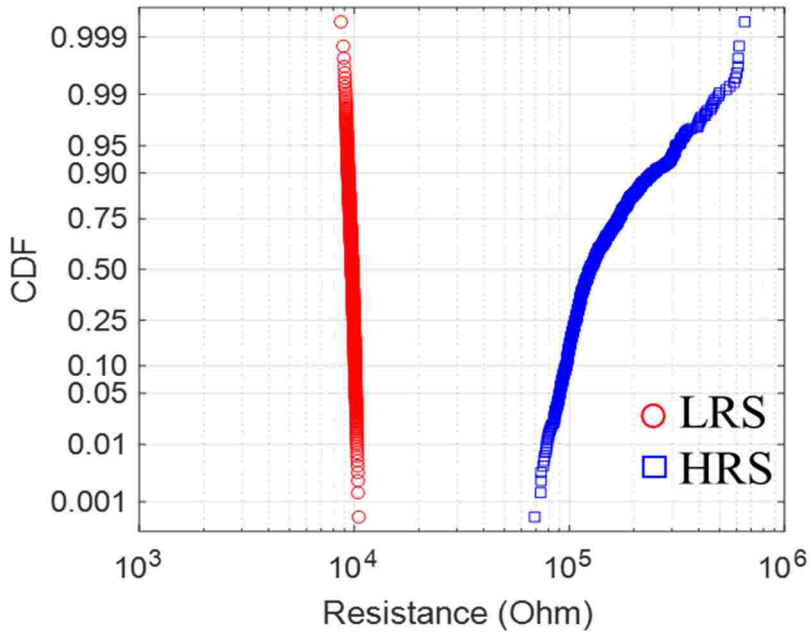

Figure 21. Cumulative distribution of 1000 cycles of LRS and HRS during normal switching operation.

Figure 22 shows the time evolution of the device resistance during programming. The resistance (gray trace) toggles between the LRS_{TARGET} (red dotted line) and the HRS_{TARGET} (blue dotted line). The SET operation is abrupt and well behaved as the resistance decreases after V_{SET} as expected. On the other hand, the RESET operation is more gradual (more pulses) and displays both a decrease (abnormal) and increase (normal) in resistance after V_{RESET}. This behavior could be a result of either high currents[26] or RESET current overshoot [162].

Figure 22 (b) captures the onset of cycling failure. The device goes into HRS and is unable to SET to LRS for nearly half of the entire data acquisition time. Cycling ceased after 400 ms of cycling. This is referred to as a Stuck HRS failure. Note, during STUCK HRS the device is continuously pulsed with V_{SET} until SET occurs. A high V_{RESET} or insufficient V_{SET} is typically associated with Stuck HRS [62]. At first glance, this does not seem to be the case here since SET requires a low number of pulses while RESET requires a high number of pulses as

shown in Figure 20. This suggests that, V_{SET} is sufficiently high while V_{RESET} is low. However, the occurrence of Stuck HRS can be explained by the randomness of resistive switching.

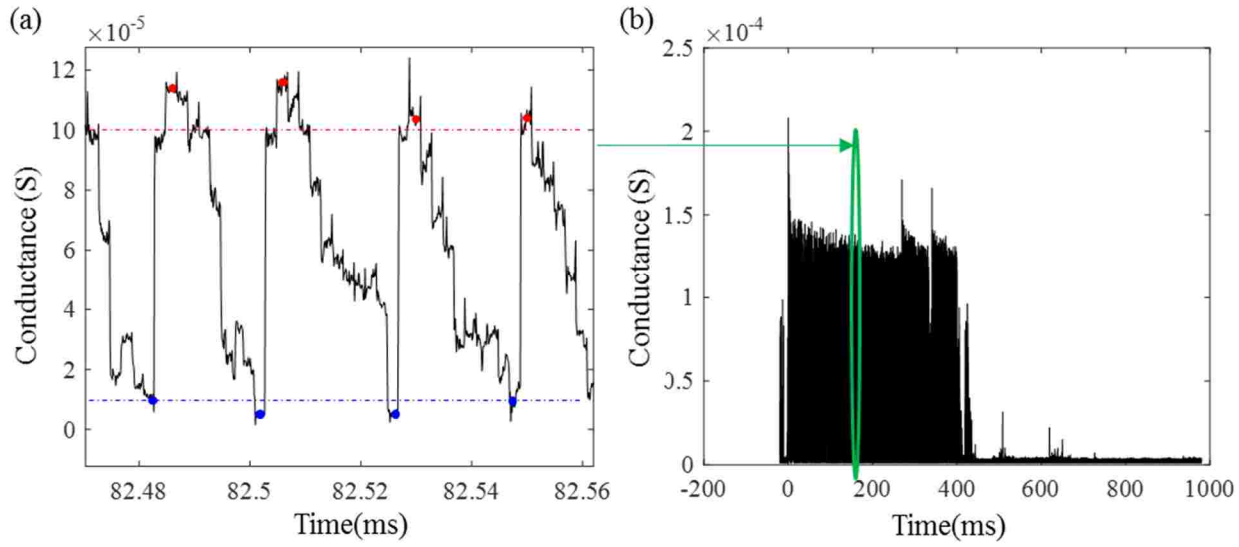

Figure 22. Conductance time trace showing the (a) normal cycling behavior of the RRAM and (b) occurrence of stuck HRS failure.

During programming, V_{SET} (+1 V) and V_{RESET} (-1 V) are fixed; however, LRS and HRS are not fixed. As shown in Figure 21, the HRS distribution is wide and spans a decade reaching deep resistance levels. While +1 V (V_{SET}) may be sufficient to SET the device for low to moderate HRS, it may become weak for deep HRS. Therefore, the randomness of the resistive switching process initiates the failure and dictates the strength/weakness of the V_{RESET} /V_{SET}. Thus, the failure is unavoidable. Additional failure modes reported in the literature include stuck LRS [81, 163] and stuck IRS (Intermediate Resistance State) [160]. Stuck LRS is marked by a failure to RESET due to a strong V_{SET} or weak V_{RESET} [81]. Stuck IRS is marked by the increase or decrease of the LRS or HRS to some intermediate resistance level. This is associated with the

formation of a parasitic oxide at the interface which acts as a series resistor with resistance equivalent to the intermediate state [80]. We only observed the stuck HRS failure during cycling.

Switching failures can be permanent and irreversible or temporary and reversible. Permanent failures are associated with over-programming, which causes thermal damage due to the excessive currents [69]. Temporary failures are associated with weak programming pulses and are, therefore, recoverable by utilizing a stronger pulse. Y.Y. Chen et al. demonstrated that by using a DC sweep after either a SET or RESET failure, the device can become responsive again, and cycling can continue. This procedure to resuscitate cycling is referred to as a recovery operation (RO). Pulse RO has also been demonstrated [159, 160]. However, the switching performance after recovery was not examined.

In this work, we demonstrate recovery using CUSPP and demonstrate switching performance post-recovery. The stuck HRS failures observed in this work are not permanent failures but temporary and reversible. When a failure occurs, CUSPP protocol continues to apply V_{SET} until SET occurs. We observed, on several occasions, that the switching begins randomly without changing pulse voltage. In some cases, the failure time becomes exceedingly high. In such a case, the algorithm is halted and V_{SET} is increased to SET the device. However, once SET is successful, V_{SET} is readjusted back to its lower initial value, and cycling continues. Currently, this process of disrupting programming and adjusting V_{SET} is performed manually.

Figure 23 shows the endurance of the device after three different recovery operations that required an increase in V_{SET}. This new V_{SET} is only used to force the device out of HRS. Endurance in Figure 23 (a) and (b) was achieved after a SET recovery using a V_{SET} recovery pulse of amplitude +1.17V (17% increase). A successful SET was achieved after roughly 15 pulses in both cases. This pulse number is expected to vary randomly. In previously reported

studies on recovery operation, the post-recovery endurance was either short-lived [68] or degraded [159]. However, we did not observe a degraded endurance post-recovery in our experiments. We achieved 10^6 endurance after the first RO and 10^8 after second RO, demonstrating a successful recovery. Figure 23 (c) shows 10^8 endurance after a third recovery. For the third RO, +1.17V could not SET the device after 10^3 pulses. The recovery V_{SET} was therefore increased to 1.29V and SET was successful after 6 pulses.

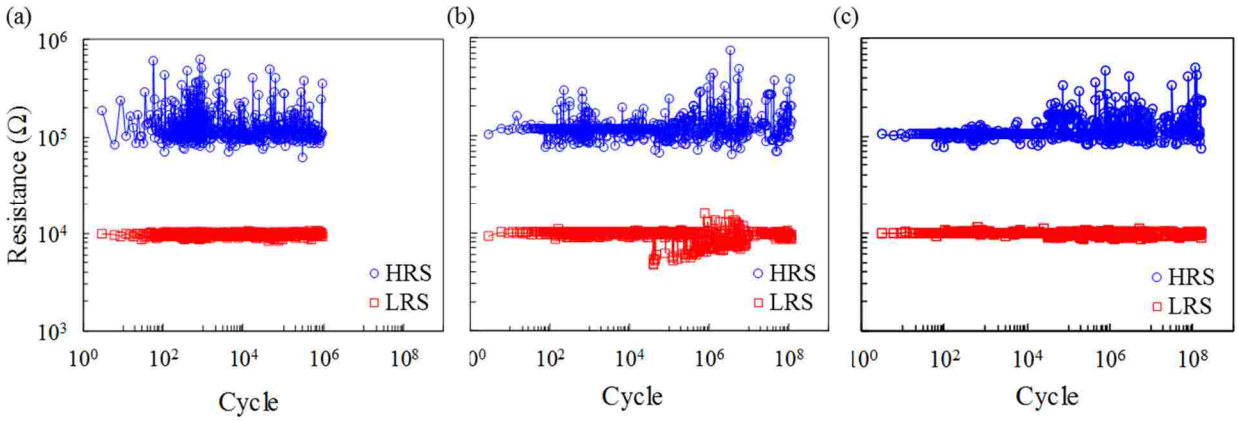

Figure 23. Endurance using +1V V_{SET} of a single cell after the (a) first, (b) second and (c) third recovery operation (RO). For the first and second RO, <15 pulses at 1.17V was needed to recover SET. While for the third RO, 6 pulses at 1.29V were needed to recover SET after failures using 1V and 1.17V.

We conclude that the endurance performance is not affected by the recovery operation. Wang et al. observed a degradation of some switching parameters with extended cycling [164]. We did not observe degradation to SET and RESET operations as shown by both the endurance and cycling-rate analysis. To further support this claim, we examined the impact of the recovery operation on the resistance distribution post-recovery as shown in Figure 24. The LRS and HRS distribution for 10^3 cycles, taken after all 3 recovery operations, shows no apparent sign of

degradation. Although 10^3 cycles represent a very small portion of the entire 10^8 cycles achieved in this study, the distributions neither improved nor worsened over time and after recovery operations.

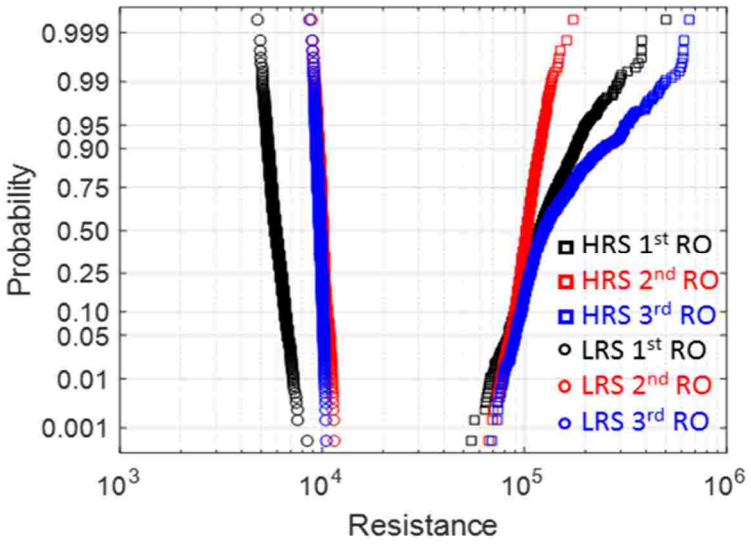

Figure 24. LRS and HRS distributions for 1000 cycles obtained after three different recovery operations using RP2. There is no trend of a degrading distribution with increasing number of recovery operations.

Based on the analysis of the endurance, distribution and cycling rate of the device, pre- and post-recovery, we propose a strategy to achieve high endurance yield in a memory chip across billions of cells. Traditionally, high endurance is achieved when the programming pulse is optimized for a single cell [58, 68]. However, the optimized conditions are unique to that individual cell. In a memory array with billions of cells, it becomes impractical to find the optimal programming pulse conditions for each cell.

The proposed scheme draws on two very important findings in this long endurance study. First, programming pulse parameters must be chosen such that permanent failures do not occur.

To achieve this, the programming voltages must be chosen such that the step changes in resistance during programming are small. Large abrupt resistance changes can easily lead to over-programming. The initial pulse conditions to achieve multiple operations can be determined during the developmental stage of an RRAM chip.

Second, a recovery operation algorithm should be a ramp voltage pulse (RVP) to achieve fast but controlled recovery. This was demonstrated earlier where V_{SET} was increased twice to expedite the recovery operation. RVP or incremental single pulse (ISP) are traditionally used for programming and not for recovery operation [87]. Here, we propose using RVP only for the recovery step.

The execution of the RVP in this work was performed manually due to the limitation of the current experimental setup. Without doubt, this algorithm can be deployed automatically and intelligently. Based on the failure criteria, the control circuit can increase V_{SET} accordingly. In a memory array, some devices would require RO more frequently than others. However, because there is no degradation of switching performance after RO, devices that require more recovery can still achieve high endurance. Figure 25 shows the flow chart of the proposed programming scheme.

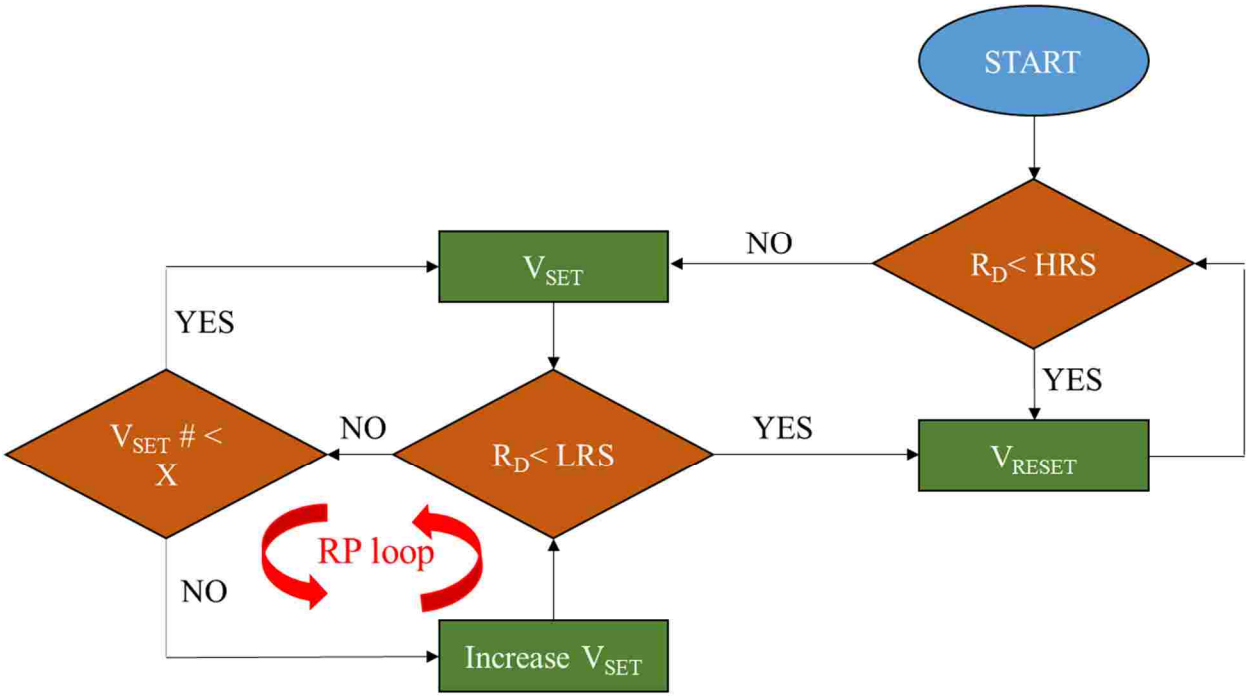

Figure 25. Flow chart of the proposed programming algorithm to achieve high endurance yield on a RRAM wafer. X represents the number of failed VSET pulses before recovery protocol is deployed.

3.3 Summary

In this chapter, we assessed the cycling stability of RRAMs during a long-endurance experiment in which the RRAM was programmed using CUSPP for 12000 sec to achieve ~ 3 x 10^8 cycles. The experiment was stopped after achieving 10^8; the device did not fail. The high endurance achieved here 1) compares favorably with reports of other program-verified endurance approaches [84, 158] and 2) demonstrates the effectiveness and functionality of the developed tool. V_{SET} and V_{RESET} were chosen such that programming required multiple pulses per operation. Although the multi-pulse operation is slower, its gentle programming reduces the risk of over-programming, which is known to cause endurance failure. We analyzed the cycling-rate behavior of the device under such a multi-pulse operation. During extended cycling, we

observed that this cycling rate randomly changes but does not degrade or improve over time. Analysis of the SET and RESET pulse numbers showed that RESET required ~ 8 x more pulses compared to SET and dictated the overall cycling rate. SET usually required a maximum of 3 pulses.

Program-verify algorithms continue to apply programming voltage until the resistance reaches its target. However, if this process requires an abnormally long time to complete, it is referred to as a failure. The failure may occur during the SET operation (Stuck HRS) or RESET operation (Stuck LRS) [80]. Failures can be permanent in which case the device will never reach the target; alternatively, failures can be temporary in which case the device will eventually reach its target. We observed only recoverable SET failures using the CUSPP protocol. Since permanent failures are associated with excessive currents due to over-programming [165], the absence of them demonstrates the efficacy of utilizing the engineering approach developed in this work.

During the stuck HRS, the device is unable to SET to the target 10 kΩ in a timely manner. Therefore, to expedite SET, a higher V_{SET} is required. This operation is referred to as a recovery operation (RO). A total of 3 RO were performed during this study. The first two required a 17% increase in V_{SET} (+1V) to SET the device; the last recovery required a 29% increase in V_{SET} to SET. Note, after a successful SET, V_{SET} is decreased to its initial value to resume cycling. Analysis of the cycling-rate dynamics and resistance states – pre- and post-recovery – do not show any degradation. These observations led to a proposed programming algorithm with RO capability to achieve high endurance in a memory array. The algorithm first requires that all failures are recoverable. To ensure this, the gradual approach of multi-pulse programming must be employed. When the device goes into a failure, a ramp voltage series

should be used to recovery cycling. Since recovery has been demonstrated to preserve performance, this algorithm can be executed indefinitely to achieve a high-endurance yield across a wafer.

CHAPTER 4

SHORT-TERM INSTABILITY

In this chapter, we utilize the newly developed programming setup, CUSPP, to investigate the short-term (sub second) stability of RRAMs during normal cycling operation. As discussed in Chapter 1, state verification during programming is necessary to overcome the intrinsic variability in RRAMs. However, program-verify techniques rely on short-term stability to successfully place the device's resistance at the target resistance level. By adopting an unconventional read operation, we directly identified two sources of instability: fluctuations and relaxation. We observed that although both mechanisms can confound program-verify techniques and render them ineffective, fluctuations dominated the instability behavior. This is in contrast to the first report on short-term instability by Fantini et al. [153], where fast relaxation was determined to be the main source of instability. Further examination of the characteristics of the fluctuations at shorter time delays and enhanced measurement time resolution (10 µs) revealed two key features. First, the probability of encountering a large fluctuation decayed with time to a non-zero value, ~ 100 µs after the verify time. Second, the amplitude of the fluctuations within the same time period did not decay, suggesting that the instability caused by fluctuations alone will not change the resistance distribution with time [43].

4.1 Experimental Details

The RRAM cells used in this study are crossbar TiN\HfO_2(5.8 nm)\Ti\TiN, with TiN as conductive electrodes, Ti as the oxygen-exchange layer (OEL), and HfO_2 as the active switching layer. Programming was performed in ambient conditions using the CUSPP setup. The devices require an initial forming step. Typically, forming voltage amplitudes varied between +2.5V and +3.5V, while switching voltages were $\leq$ |1.5 |V. The resistance of the RRAM states was

monitored by applying a low (-8 mV and -80 mV) DC voltage, V_{READ}, and measuring the output current with fast amplifiers and a real-time, high-speed oscilloscope. The measurement speed was limited by the sense circuit RC time (30 ns) and the oscilloscope sampling rate. The read duration was controlled by adjusting the pulse repetition rates to either 1 Hz or 1 KHz. To ensure sufficient separation between the ON state and the OFF state, the HRS and LRS targets were programmed to be 100 kΩ and 10 kΩ, respectively. These values were chosen to achieve a 10x resistance window (RW). The voltages (AC and DC) used for all operations were applied to the top electrode, while the bottom electrode was grounded.

4.2 Results and Discussion

Figure 26 shows a single representative HRS time trace monitored for 1 s. Notice from the inset that at t = 0 the resistance level exceeds the HRS target (100 kΩ), which means the device was successfully programmed to HRS. Because of the constant read operations and 1 Hz pulse rate, each state can be monitored for a maximum of 1 second. The measurement speed was set to 500 kS/sec (data point every 2 μs) by the sampling rate of the oscilloscope.

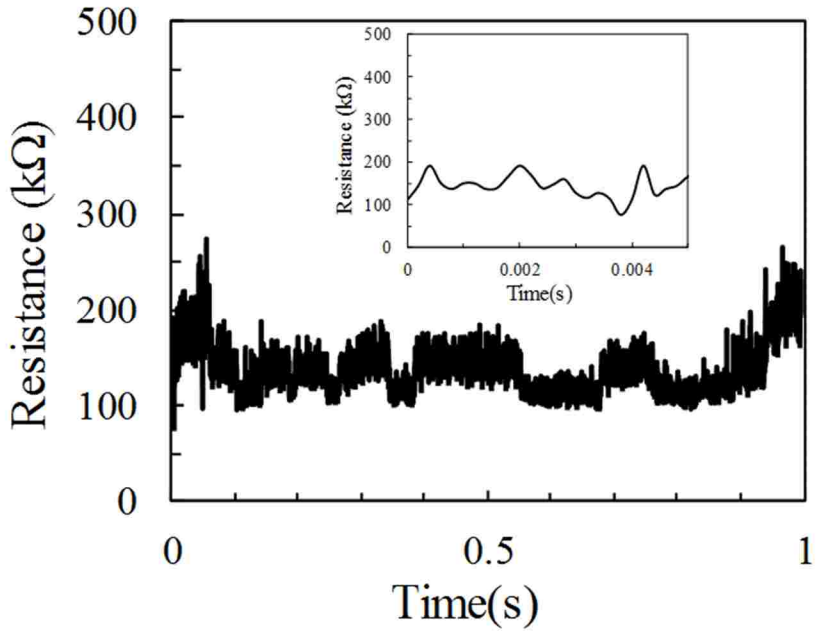

Figure 26. HRS read trace captured during normal cycling between LRS and HRS. 1 Hz programming pulse rate was chosen to enable 1 s long read duration. At t = 0 (immediately after RESET pulse) the resistance exceeds 100 kΩ (inset) target signifying completion of RESET.

Figure 27 shows the cumulative distributions of LRS and HRS at different time delays (10 µs, 100 µs, 1 ms, 1s) post programming. The resistance values are mean values of 26 µs integration times. Note, program-verify guarantees that at the start of each HRS/LRS, the device resistance is at the predefined target level as shown in Figure 26. Ideally, either distribution would be bounded by the target resistance level at 10 kΩ for LRS and 100 kΩ for HRS. However, in Figure 27, it is apparent that this is not the case since the distributions 10 µs after programming (t = 0 in Figure 26) extend past the target levels (indicated by dashed lines in Figure 27). The resistance window at the 0.01 percentile is 2x instead of the programmed 10x target. This would imply that either the program-verify failed or the resistance decreased in a very short time scale, thereby producing a distribution tail. In addition, similar to [153], the

distribution continues to widen in time, resulting in a decreasing RW that completely disappears after 1 second.

This rapid loss of the RW was attributed to relaxation processes in the switching layer [153, 166]. This attribution was based on 1) the appearance of a tail only microseconds after programming and 2) a continued spread of the tail with time. In those studies, a conventional, read-pulse operation was used to track the distribution with time. However, pulse-read operations, unlike constant-read operations, have a serious shortcoming. Pulse-read operations take snapshots of the resistance evolution at discrete points in time only. Therefore, such operations are blind to the resistance behavior between read pulses. The continuous read operation utilized in CUSPP is therefore advantageous for short-term stability since we can obtain a complete record of the filament behavior post-programming.

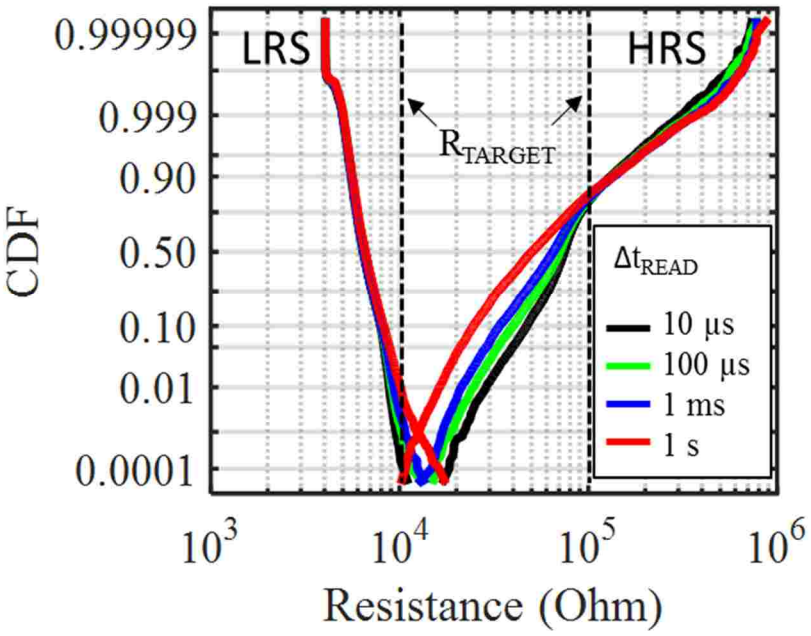

Figure 27. Cumulative distribution (10^4 cycles) of LRS and HRS measured 10 µs, 100 µs, 1 ms and 1 s post-programming. A tail is present (especially in HRS) only 10 µs after programming and continues to widen with time.

The resistance window is decreased to 2x in the first 10 µs post-verify and completely disappears 1 ms post-verify. This warrants a closer look at the resistance behavior within the first millisecond after programming. Therefore, read operation was decreased to 1 ms (from initial 1 s), and the sampling rate of the oscilloscope was increased to 10 MS/sec to provide higher temporal resolution. From Figure 27, it is clear that although both LRS and HRS distributions have tails, the HRS tail dominates window closure. The wider distribution of HRS compared to LRS is a result of reduced filament size for higher resistance, which are more susceptible to variations due to the lower number of defects making up the conductive path [37]. We, therefore, focus on the instability of HRS. This is consistent with earlier studies, where it was observed that the fluctuation amplitude increases with resistance [38, 167]. To observe the impact of the resistance level, both LRS and HRS targets are increased to 20 kΩ and 200 kΩ, respectively.

Figure 28 is a representative temporal trace of the resistance evolution for a complete cycle (SET and RESET). Note, it required 3 pulses to reach the HRS target (200 kΩ) and 1 pulse to reach the LRS target (20 kΩ). Immediately following each pulse (pulse location indicated by red markers), there is a large transient resistance fluctuation. In addition, during the read operation (time between pulses), we observe randomly occurring fluctuations. These random fluctuations are not measurement system artifacts and are absent in calibration measurements of reference surface-mount resistors of similar values.

The appearance of resistance fluctuations in RRAM during retention studies and electrical stress measurements have been reported [39, 156, 168, 169]. These studies typically involve a constant low voltage read over thousands of seconds at normal and elevated temperatures. In effect, by utilizing a constant read during programming, CUSPP combines a short-term, stress measurement with resistance cycling. Still, the observed fluctuations here

happen at time scales that are orders of magnitude shorter. Thus, if the origin of the fluctuations observed here is the same as reported for longer retention studies, then the mechanism must span a rather wide time range.

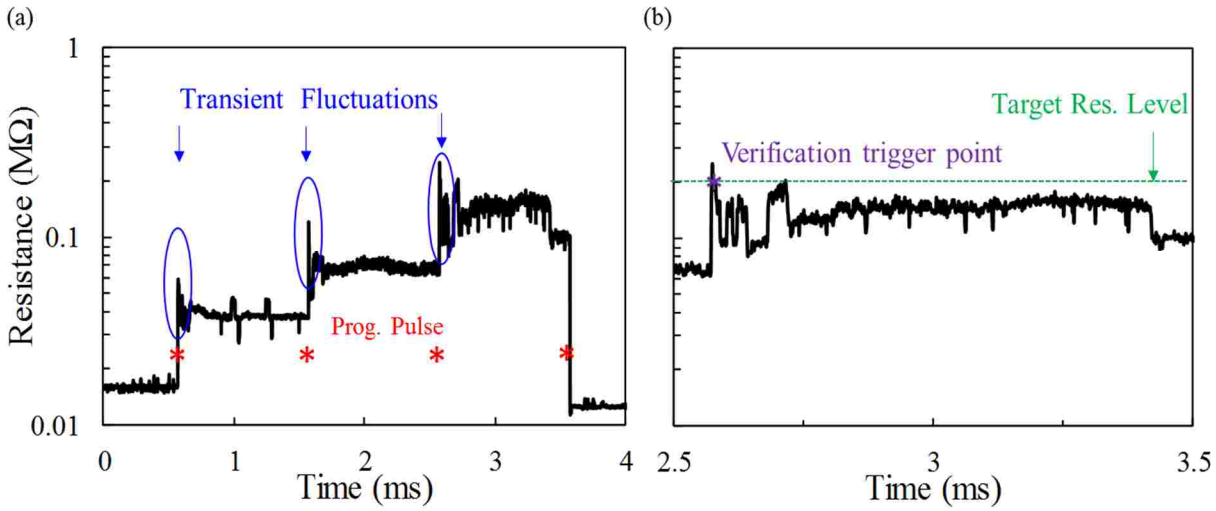

Figure 28. (a) Measured resistance (V_{READ} = -8mV) as a function of time for a complete cycle. The RESET operation required 3 pulses, while SET required a single pulse. Notice the presence of transient fluctuations immediately after the programming pulse (red markers) and at random times during the read operation. (b) Shows the HRS state monitored for 1 ms. Notice that the fluctuation reaches the target (200 kΩ) and triggers (purple marker) the fast comparator but the resistance decreases to a lower resistance level. This event causes the tail in the distribution.

During programming, the fast comparator is triggered when the resistance exceeds the HRS target-threshold. Once the comparator is triggered, the device resistance is designated as HRS and the operation is complete. This is illustrated in Figure 28 (b), where HRS was achieved and monitored for 1 ms. The comparator triggered on the transient fluctuation; however, almost immediately after, the resistance fluctuates between 100 kΩ and 200 kΩ before stabilizing. From the continuous read, we can directly observe that the initial tail after 10 µs is not due to failure of

the program-verify protocol. Indeed, that protocol did successfully register a RESET operation because the device resistance reached the target level. However, the stability of the resistance state after verification cannot be guaranteed by program-verify. The appearance of a tail, therefore, is a result of fast fluctuations that prematurely verify program completion.

A slower comparator might be the solution to this problem because it might be able to ignore such fast fluctuations, thereby allowing the program-verify to properly complete. Nevertheless, as will be discussed later, there are practical reasons why this solution will not work. To explore possible solutions to this serious fluctuation problem, further examination of the fluctuations themselves was needed.

As was discussed in Chapter 1, due to the inborn variability in the RRAM switching mechanism, we would expect that the characteristics of the fluctuations would differ from one cycle to the next. We statistically evaluated the fluctuation behavior for 4000 cycles. Figure 29 (a) – (c) illustrates three representative, HRS, temporal traces during SET and RESET cycling. The high resolution (100 ns) raw data (in gray) reveals the presence of fast fluctuations and slow fluctuations. In fact, these fast fluctuations produce the Random Telegraph Noise (RTN) signal often observed in RRAMs [168]. The slow fluctuations become clearer after a 10-point average (red trace) of the raw data; this decreases the read resolution from 100 ns to 1 μs. For comparison, Figure 29 (d) shows the raw and average resistance trace of a fixed value 260 kΩ surface mount resistor. Neither fast nor slow fluctuations observed in the RRAM cell are observed for the resistance measurement of a surface mount resistor. The fluctuations are therefore not a measurement system artifact but are directly linked to the physical nature of the conductive filament.

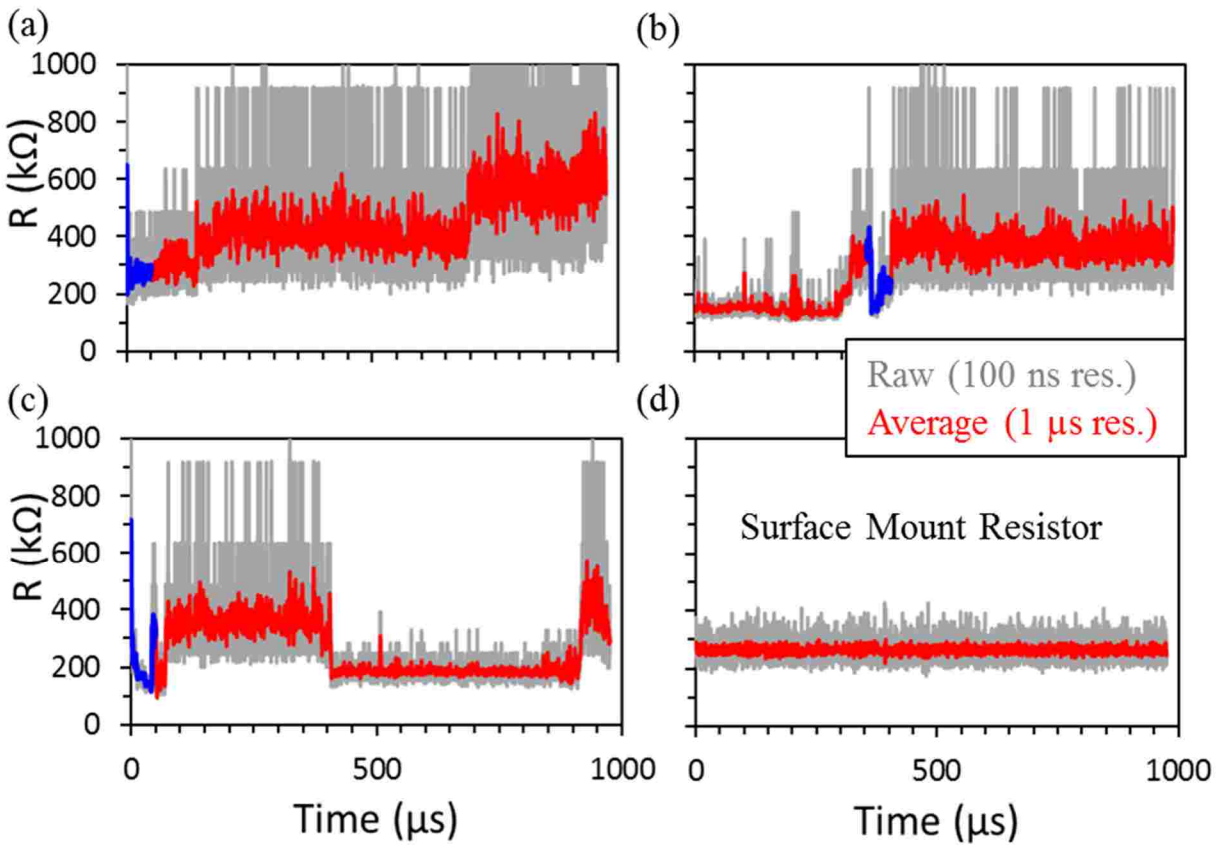

Figure 29. (a), (b) and (c) Representative HRS read traces captured during normal operation between HRS and LRS. Averaging (red) suppresses fast fluctuations present in the raw data trace (gray) but exposes slow fluctuations. For each HRS, the location of the largest fluctuation (blue) after averaging is recorded as $\Delta R_{LARGEST}$. (d) 260 kΩ surface mount resistor does not show similar fluctuations observed in RRAM devices indicating that the fluctuations are related to the device.

Although averaging successfully decreased the overall noise of the read operation as shown in Figure 29, averaging has its practical limits because it slows down the read operation, which in turn affects overall write/erase speed of RRAMs. To maintain reasonably fast operational speeds, we do not average further. The behavior of the fluctuations after 10 pt averaging is analyzed hence forth. We statistically investigated the occurrence rate of these fluctuations on 4000 HRS traces. Since larger fluctuations are more problematic, for each

trace/cycle, the largest fluctuation is recorded. To accomplish this, each cycle is divided into one hundred 10 μs long, time segments. For each HRS cycle, the time segment containing the maximum resistance change, $\Delta R_{LARGEST}$, is identified as shown in Figure 29 (blue data).

The distribution of these segments is represented as a histogram in Figure 30 accumulated for all cycles. This distribution can be interpreted as the probability that a large fluctuation will occur within a given time segment. The histogram reveals two key observations about that probability. First, the probability of encountering the largest resistance change or fluctuation in a 1 ms read is highest at the beginning of the read operation, immediately after the programming pulse. Second, the occurrence of these fluctuations does not vanish entirely within the duration of the read operation but saturates to some non-zero value.

This observation suggests that large fluctuations can occur randomly at any time during read. Therefore, introducing a read-delay after programming does not avoid fluctuations and subsequent read errors. Due to insufficient sampling, the functional form of the decay cannot be accurately determined. However, it is clear that there is a significant reduction in the number of occurrences of the largest fluctuation within the first tens of microseconds after RESET.

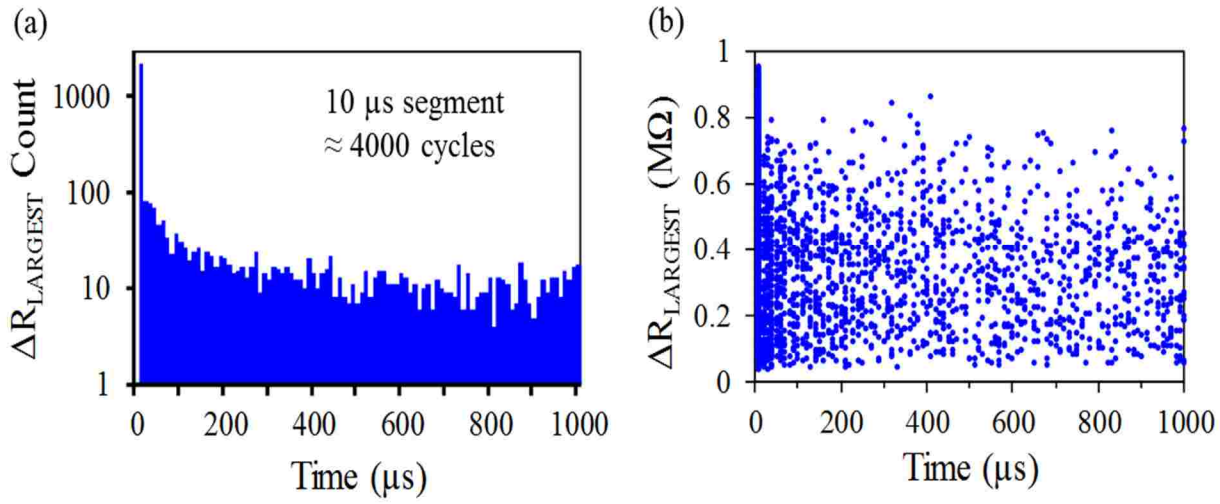

Figure 30. (a) Distribution of time segments (10 µs) containing the largest fluctuation, $\Delta R_{LARGEST}$. There is statistically higher probability of the largest fluctuation occurring closer to the programming pulse (t = 0). (b) The corresponding distribution of the fluctuation amplitude show that the fluctuation amplitude remains large over the entire read time. $\Delta R_{LARGEST}$ can occur randomly at any time during read.

Figure 30 (b) shows the corresponding $\Delta R_{LARGEST}$ amplitude of the same resistance fluctuations shown in Figure 30 (a). As expected, the data density decreases with time, highlighting the decaying nature of the fluctuation probability with time. Figure 30 also reveals that not only do large fluctuations persist, their amplitude remains unchanged for the entirety of the read operation. Consequently, regardless of the nature of the read operation, these large fluctuations will exist. Furthermore, they pose a serious vulnerability because they impact both read-out margin and level-placing programming algorithms.

Admittedly, the choice of the segment size may be responsible for the observed behavior of the fluctuations. Smaller time segments are not feasible because sufficient points are needed to capture an entire fluctuation. The 10 µs long segment, therefore, only captures fluctuations that occur within 10 µs. Slower rise or fall times of fluctuations cannot be captured by this method.

However, the bin size can be increased to account for slower changing fluctuations. Figure 31 shows the distribution of $\Delta R_{LARGEST}$ for 100 µs long segments. Similar to Figure 30, the fluctuation probability shows the same behavior. The largest fluctuations occur more frequently early in the read operation, and the probability does not decay to zero.

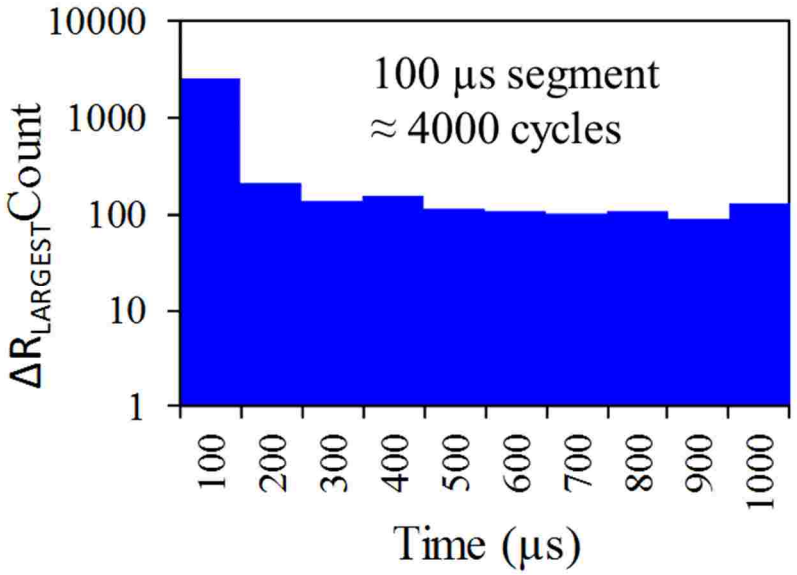

Figure 31. Distribution of time segments (100 µs long) containing the largest fluctuation. The fluctuation distribution decay is unchanged for longer time segments.

The analysis of the $\Delta R_{LARGEST}$ indicates that some fluctuations can increase the resistance level, while others can decrease it. From a practical standpoint, the impact of the direction of resistance change is determined by the intended RRAM usage. For multilevel cells, both types of fluctuations are serious obstacles to reliability. On the other hand, for single level cells, only fluctuations that close RW are critical. As shown in Figure 27, when HRS decreases the RW closes. However, if HRS increases, an unintended benefit is attained since RW increases. Similar to the construction of the distribution shown in Figure 30 (b), Figure 32 shows the distribution of

the segments in which the minimum HRS resistance value, HRS_{MIN}, occurs. This distribution can be interpreted as the probability that a given fluctuation will close the RW by decreasing HRS. We observe that this probability is highest immediately following the RESET pulse and slowly decays to a non-zero value with time. These fluctuations persist for the entire read operation. The HRS_{MIN} distribution is nearly identical to $\Delta R_{LARGEST}$ distribution, corroborated by a correlation factor of 0.98. This correlation is an indication that the majority of large fluctuations do act to close the RW.

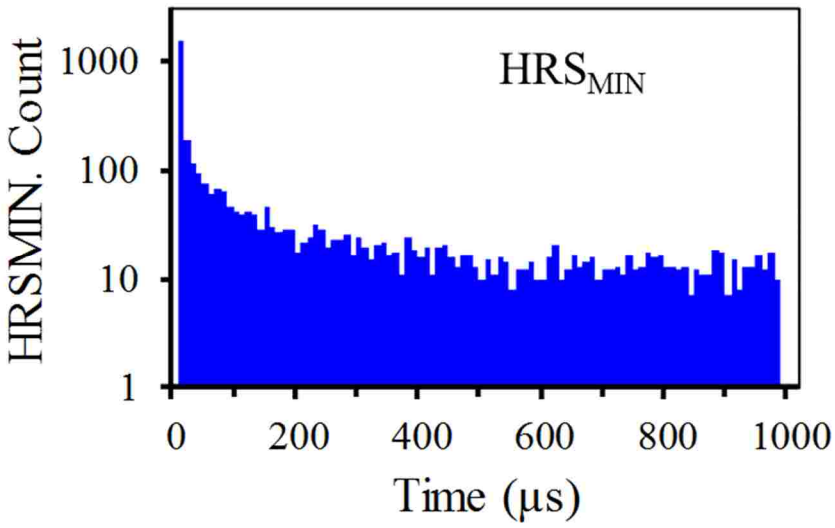

Figure 32. Distribution of time segments (10 µs long) in which the minimum resistance point (HRS_{MIN}) in HRS occurred. Similar to $\Delta R_{LARGEST}$, there is a higher probability of window closure occurring immediately after RESET pulse.

Immediately following the programming pulse, both distributions show conspicuously large counts in the first 10 µs. Note, the large count is likely an artifact of the chosen bin size. As stated previously, the bin size must be long enough to capture an entire fluctuation and not just

an edge; making segments smaller is not possible. In both $\Delta R_{LARGEST}$ and HRS_{MIN} distributions, the probability decays in time further away from the applied RESET pulse. It is tempting to assign this fluctuation decay behavior to a thermalizing process of the local area surrounding the filament, since the temperature of the CF is expected to increase due to Joule heating from the programming current. However, the use of ultra-short pulses for programming makes this unlikely.

A first-order estimation of this thermalization process can be constructed by assuming that i) the heat diffusion equation is valid for very short time scales and ii) the CF acts as a point heat source in a homogenous dielectric medium. At the end of the applied 100 ps (τ_0), heating terminates, and the heat diffuses over a spherical volume ($4\pi r^3/3$), where r is the radius of the sphere. The hot-sphere-volume radius expands, due to diffusion, as the square root of the product of the thermal diffusivity and time. The temperature drop is assumed to be proportional to the volume increase.

Figure 33 is constructed by i) normalizing the initial temperature after the programming pulse to 1 and ii) taking a ratio of temperature (volume) as a function of time delay after the initial pulse. While not exact, this heuristic estimation of thermal diffusion reveals a telling conclusion. Recall that Figure 30 and Figure 32, show that the fluctuation decay occurs on time scales of tens of microseconds (10^4 and 10^5 times longer than τ_0). This decay corresponds to a temperature difference of about 10^7 from the time after the pulse is terminated to microseconds later. Since the temperature of the CF cannot cool below ambient (experiments are performed at room temperature), this puts the initial temperature, immediately following the programming pulse, at improbably high values.

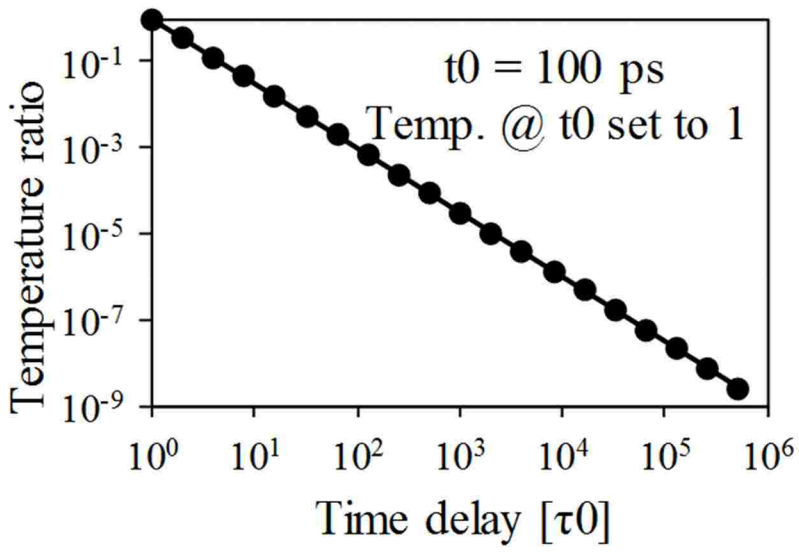

Figure 33. Calculated temperature drop ratio referenced to an initial temperature as a function of multiples of characteristic time τ0 (pulse duration).

CF temperature values reported in the literature using advanced methods and simulations are typically in the range of several hundred to thousands of degrees Kelvin regardless of a diffusive or non-diffusive thermal transport mechanism [27, 170-172]. It is conceivable that similar temperatures are reached in our case. However, considering that the ultra-short programming pulse heats up a small volume, the temperature drop due to heat diffusion will be extremely short. Based on Figure 33, the CF temperature reaches room temperature before our first read point, 100 ns later (oscilloscope sampling rate). Thus, we conclude that the tens-of-microseconds decay of the fluctuations is not a thermally driven process. This first-order thermal-diffusion estimation does not account for the conduction of heat away from the CF by the electrodes. Still, this underestimation cannot account for the many orders of magnitude discrepancy.

As previously stated, a constant read operation is utilized in this study to generate a complete record of the CF resistance evolution. Consequently, “read disturb” could be responsible for the observed behavior of the fluctuations [173]. During programming, the device cycles between high and low resistance levels. Since the target HRS and LRS levels differ by at least a factor of 10x, the read current in the low-resistance state should be equivalently at least 10x larger than the high-resistance state. The applied read voltage (-80mV), induces both an electrical field and a temperature that persist for 1 ms. The fluctuation decay behavior which occurs over tens of microseconds could be a result of the read operation.

To test the validity of this hypothesis, we repeat the fluctuation analysis with a 10x lower read voltage (-8 mV). This translates to a 100x decrease in power (approximately 100x decrease in temperature). We find no discernable difference in the fluctuation decay phenomena with this different read condition. A correlation coefficient = 0.93 was calculated; thus, the fluctuation behavior is likely not a result of read disturb. Our analysis of the thermal gradients during programing and read operations suggest a very limited role in the observed fluctuation dynamics. However, fluctuations in RRAM have also been demonstrated to be associated with the electrically driven charge trap/detrap mechanism. Although this may be the likely cause of fluctuations in our case, it does not fully explain the decay nature of the fluctuations.

So far, we have conclusively shown that fluctuations are responsible for the tail of the distribution in Figure 27. However, this does not exclude the presence of relaxation. In fact, the increased dispersion of the distribution with time is indicative of some relaxation. We investigate the presence and impact of relaxation on the programming error by revisiting the longer read time (1 s) data set. Figure 34 shows a time-lag representation of the resistance states taken 10 µs and 1 s post programming, accumulated for 13000 cycles. The black dashed line is a correlation

line drawn to show an ideal case where the post-programmed resistance state is stable for 1s. Therefore, resistance states above or below the line are an indication that the initial (10 µs) resistance value increased or decreased respectively, after 1 s. It is evident that both LRS and HRS are unstable causing the initial resistance to randomly increase or decrease over time. However, we find that 72.3% of HRS decrease while 65.4% of LRS increase leading to severe window closure. Although there is some instability in LRS, HRS dominates the window closure (see Figure 27) due to the larger magnitude of the deviations. The trend towards lower values for high resistance values is consistent with our observations in Figure 32.

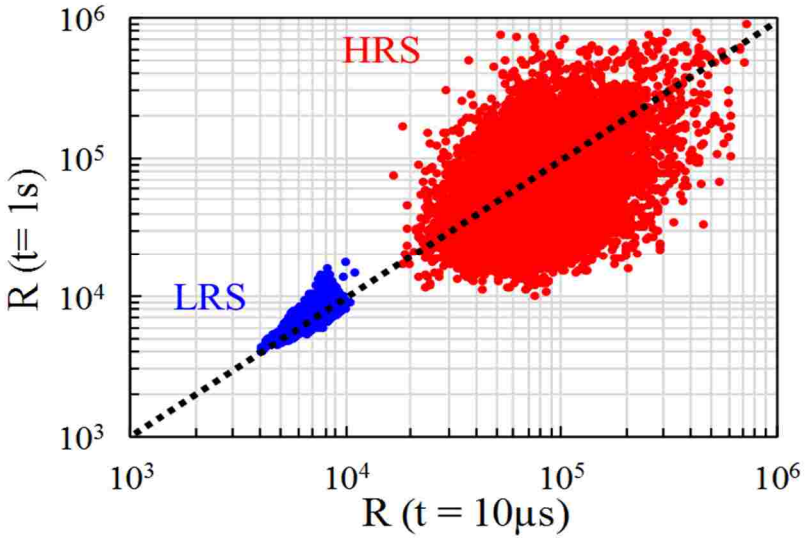

Figure 34. Time lag representation of HRS and LRS taken at 10 µs and 1s post-programming. The black line is drawn to show perfect correlation in an ideal case without fluctuations or relaxation. While LRS trends above the line, HRS trends below the line. HRS trend is not apparent in this representation.

Still, it is incorrect to assign the instability in Figure 34 to fluctuations only. In RRAMs, the resistance state is determined by the configuration of the conductive filament (CF). LRS and

HRS are distributed because this configuration changes from cycle to cycle. This is unavoidable as this randomness is inherent in the resistive switching process. For each cycle/state, there exists an energetically favorable configuration that may cause the resistance to gradually (= relaxation) or abruptly (= fluctuation) change. Therefore, it is premature to assign the instability to fluctuations as relaxation can produce a similar time lag plot. To differentiate between fluctuation and relaxation, we examine the intermediate resistance state behavior between the initial and final states.

The broad initial distribution of HRS is caused by the unstable resistance state verification. We artificially null out this effect by normalizing. This way, we track the evolution of each HRS, regardless of its initial value. Figure 35 (a) shows the resulting cumulative distribution in such an exercise, where we take a ratio of the resistance at some time, t, post programming, R(t), to the resistance measured at 10 µs post-programming, R(10 µs). The choice of R (10 µs) as the reference is arbitrary; selecting a different reference time does not change the result of the analysis. As such, at t = 10 µs, the ratio R(t)/R(10 µs) equals 1 and results in a vertical line located at 1. The distributions corresponding to ratios taken at 100 µs, 1 ms and 1 s post-programming extend above and below the reference vertical line (see Figure 35 (a)), indicating a shift of the initial resistance value towards higher and lower resistance values, respectively.

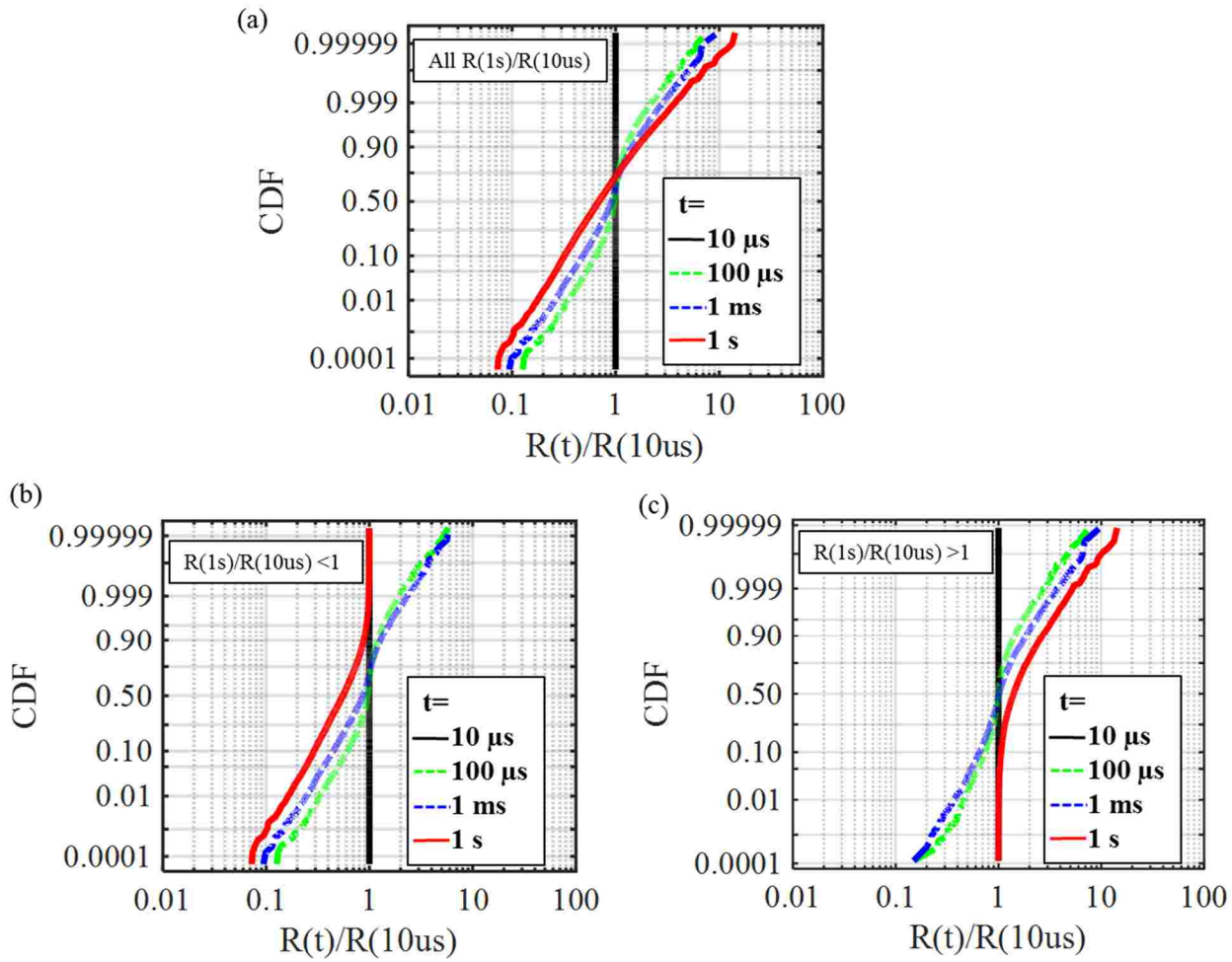

Figure 35. (a) Cumulative distribution of resistance ratio, R(t)/R(10 µs) for t = 10 µs , 100 µs , 1 ms and 1s for all cycles. The initial normalized values (t = 10 µs) widen with time above and below indicating a random shift. Distribution of HRS cycles which (b) decreased or (c) increased after 1 s. In both cases, the distributions at 100 µs and 1 ms show values above and below the references. Indicating that the resistance fluctuates randomly over time.

The ratios taken for t = 10 µs and 1 s represent the initial and final value of each cycle, while the ratios taken at t = 100 µs and 1 ms represent the intermediate states. In this work, we assume that in a relaxation process, if a state increases/decreases to its final value, the intermediate state should also increase/decrease to that final value. To examine this, we sort the distributions to select all R(1 s)/R(10 µs) (red curve Figure 10 (a)) that ended either below ((R(1

s)/R(10 μs)< 1) or above ((R(1 s)/R(10 μs) > 1) the initial value (black line). This tracks all cycles which either decreased or increased from initial to final value, respectively. These two subsets are illustrated in Figure 35 (b) and (c), respectively. Thus, Figure 35 (b) illustrates the distribution of all HRS cycles that decreased in resistance after 1s (notice distribution for R (1 s) falls below the reference line). Figure 35 (c) illustrates the distribution of all HRS cycles that increased in resistance after 1s (notice distribution for R (1 s) falls above the reference line). However, for both Figure 35 (b) and (c), the intermediate states are distributed above and below the reference line. If the instability were indeed due to a relaxation effect, the intermediate reads distribution should have trended towards the final distribution, which clearly is not the case. This data is notably consistent with a system dominated by random fluctuations and inconsistent with a relaxation effect.

From this analysis, we conclude that fluctuations are the major culprit of window closure in our case. The observed minor impact of relaxation compared to earlier works [153, 166] may be due to several factors. A difference in material stack has been shown to impact performance including retention [33]. The thermodynamics associated with the nature of the programming setup may also account for the observed behavior. As discussed, programming has a direct impact on retention and overall performance of RRAM. However, it is plausible that techniques that utilize pulse read are blind to these fluctuations and are, therefore, limited to observing only mean value shifts. For example, a pulse read at 10 μs and 1 s after programming cannot directly observe the filament behavior between the two reads.

To improve reliability in RRAM operations, the impact of fluctuations must be diminished significantly. We explored a few techniques that can possibly minimize the fluctuation impact on the RW. One such technique is guard banding. This technique does not try

to minimize fluctuations; instead, it tries to minimize their impacts. Figure 36 shows the HRS distributions for various target resistances. These distributions were generated by cycling the same device with different sets of LRS and HRS targets. Artificially shifting, or increasing the target HRS, shifts the whole HRS distribution (including the tail) and forces the "initial" resistance window to be larger. Others have noted [42, 168, 174], and as shown here in Figure 36, a general increase in resistance increases the fluctuation amplitude (i.e. smaller filaments, larger fluctuations). Thus, a simple increase of target HRS also causes an increase of the distribution tail. Therefore, the "sweet spot" for target resistance needs to be identified through careful characterization of the RRAM devices.

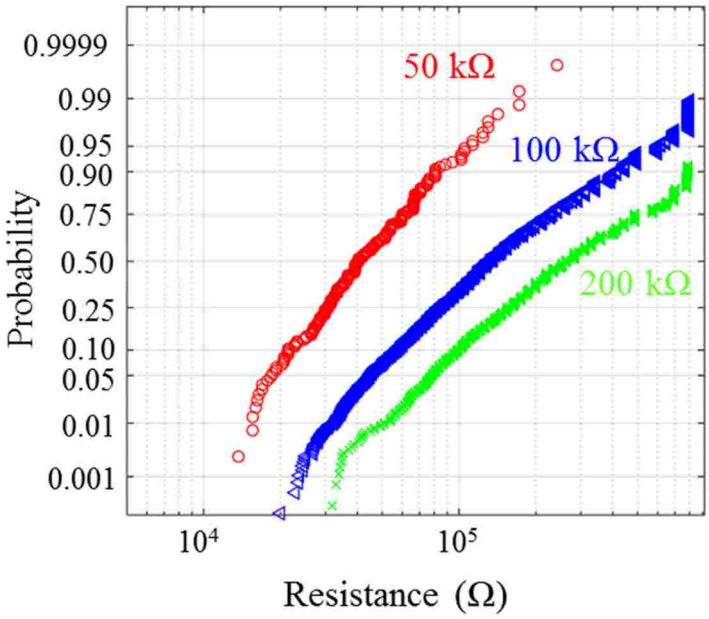

Figure 36. HRS distributions for different target resistances. Artificially shifting, or guard banding the HRS target to higher values shifts the entire HRS distribution but also increases dispersion suggesting limited benefits to guard banding.

Other techniques to minimize the impact of fluctuations require slowing down the operating time either by utilizing a read delay or a slower (longer integration) read operation. Our analysis of both the fluctuation amplitude and rate-of-occurrence showed that within a 1 ms long read, large fluctuations still exist, but their frequency of occurrence decreases with time. Figure 37 shows the distribution of HRS 300 µs and 1 s post-programming. A 200 µs integration time was used to obtain the mean values for the distribution and to calculate the maximum and minimum (gray crosses) within the integration time. We observe a statistical reduction in fluctuation amplitude for longer post-programming times; nevertheless, fluctuations still persist even after clearly reaching thermal equilibrium (1 s after programming). The reduction in the amplitude of fluctuations after 1 s does not conflict with our earlier analysis where we observe no decrease in fluctuation amplitude for 1 ms. Figure 37 does not capture the largest fluctuation during an entire 1 s read. The lower amplitudes of fluctuations are expected because the probability of encountering the largest fluctuation in the cycle decreases with time. Therefore, it is less likely to capture large fluctuation over such a short integration time, 1 s after programming.

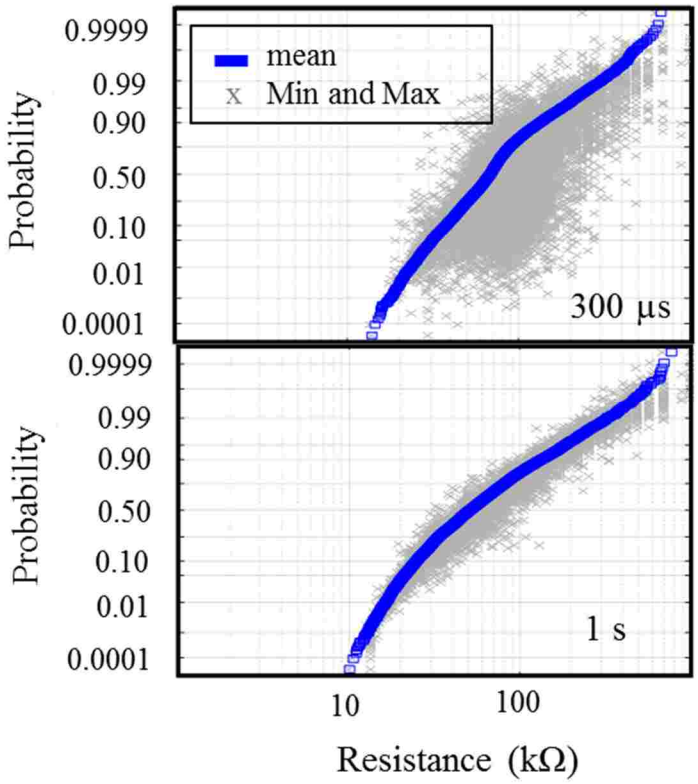

Figure 37. Cumulative distribution of HRS for 300 µs and 1 s post programming. Mean value (blue), maximum (gray) and minimum (gray) values are calculated in a 200 µs integration window. A significant reduction of fluctuation amplitudes is observed 1 s post-programming.

Slower read operations, which require longer integration times, have been shown to eliminate some fluctuations but expose slower fluctuations. In Figure 38, we examine the impact of integration on the fluctuation amplitude for 200 µs and 30 ns integration times. The apparently broad time distribution of the fluctuations limits the benefit of increased read times. Perhaps the best method to reduce dispersion is to increase the filament size, thereby minimizing fluctuations. An optimization across these approaches (guard-banding, increased read integration time, and larger filament sizes) may prove beneficial.

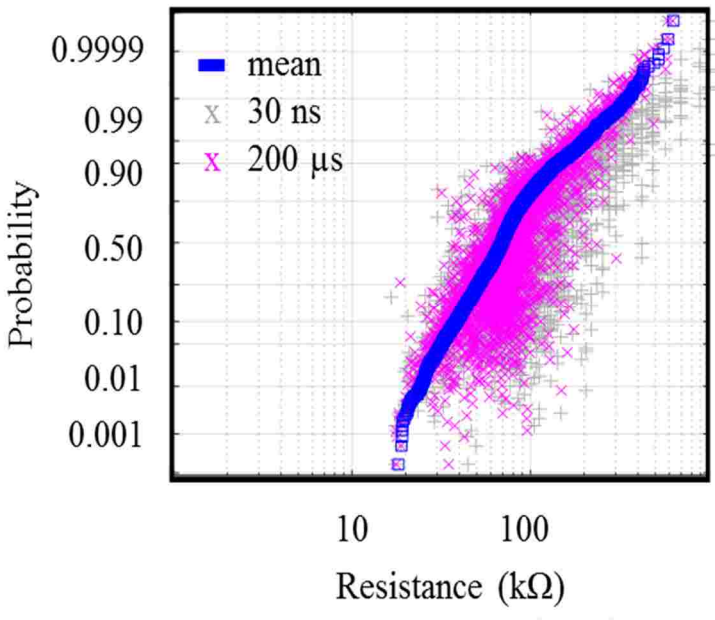

Figure 38. Increased read averaging of the resistance distributions for 200 µs (purple) vs. 30 ns (gray) reduces the amplitude of fluctuations, but the widely distributed temporal distribution of the fluctuations limits the usefulness of this tactic.

4.3 Summary

A recent study of the post-programmed, resistance-state stability revealed that relaxation processes cause resistance-window closure [153, 166]. This phenomenon defeats program-verify strategies since they are mainly employed to create a resistance window. With the enhanced read operation, we examined this short-term (microsecond) instability in HfO_2-based RRAMs using CUSPP. We determined that the stochastic resistance fluctuations, which are known to be present in oxide-based RRAMs [39, 167, 168], dominate the short-term instability of the device while relaxation played a lesser role. The program-verify algorithm records a successful SET/RESET operation when the resistance reaches LRS_{TARGET}/HRS_{TARGET}. However, once this happens, the resistance state can fluctuate randomly below or above the threshold level. Therefore, any subsequent read, at some time delay from this verify time, would produce values lower than the threshold targets. This mechanism closes the window and creates an impression of

a failure of program-verify. The true failure of program-verify is that it can obtain a resistance window but cannot maintain the resistance window due to fluctuations and relaxation.

The CUSPP control circuit utilizes a high-speed comparator with a propagation time of < 5ns. Therefore, it is conceivable that a slower comparator may avoid “false” triggering from fluctuations. However, a statistical study of the fluctuation behavior in the first millisecond after the trigger point (t = 0) proved otherwise. We observed that the probability of encountering a fluctuation decayed as time increased from the read operation t =0. However, this probability, which asymptotically approaches 0, never completely disappears. The fluctuation probability reaches a non-zero steady state after t = ~ 150 μs [43]. Interestingly, the magnitude of the fluctuations did not decrease significantly within the entirety of the 1 millisecond read. We also observed that though averaging may provide some headroom, slowly changing fluctuations limits this approach. These observations suggest that using a slower comparator or waiting longer to read (since probability decreases) does not avoid fluctuations, and it can significantly decrease programming time.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

Charge-based memory technology is fast approaching a fundamental scaling limit. As such, new solutions require devices that can meet current, memory-performance requirements and support further miniaturization. New memory devices (RRAM) based on a resistive-switching mechanism have emerged as a serious replacement. RRAM devices are highly scalable, they demonstrate switching performance that exceeds FLASH, and they are comparable to DRAM and SRAM. However, severe variability and instability issues remain at the forefront of RRAM reliability concerns [14, 37]. Variability of the resistance (memory) state is intrinsic and, therefore, unavoidable. Hence, programming solutions are required to minimize the impact of variability and to ensure a workable resistance window [86, 88, 154, 158].

In Chapter 2, we proposed a program-verify technique called Compliance-free Ultra-short Smart Pulse Programming (CUSPP). CUSPP relies on a novel, current-control scheme to achieve a low-programming energy operation that is desirable for guarding against programming errors. First, a 100 ps pulse generator constructed from several rf components was developed. To deliver the fast pulses for programming with minimal degradation and high signal integrity, transmission line design and impedance matching are employed throughout the measurement system. Second, a custom-built feedback control circuit was designed to sense the device resistance state and output appropriate logic to control the 100 ps pulse generator. The circuit houses high-speed amplifiers and comparators to enable fast response to RRAM switching transients. Programming protocol utilizes identical pulse trains to tune the device resistance to two distinct preset target levels. The functionality of the integrated ultra-short pulse and control circuit is demonstrated in programming measurements including FORMING and cycling (SET

and RESET) with state verification. In addition, a new read operation concept is demonstrated and later shown to be useful in short-term stability studies of RRAMs. We successfully demonstrated for the first time 100 ps write/erase speeds and 500 kHz cycling rate of HfO_2-based RRAM devices.

In Chapter 3, we investigated the cycling performance of RRAM using CUSPP. Devices were programmed in a long cycling experiment to achieve 10^8 cycles. We observed that the RESET operation required on average 8x more pulses and RESET dictated the cycling rate of the device. However, the cycling rate is not constant; rather, it is subject to variation due to the intrinsic variation of the resistive switching process. Occasionally, large excursions of the cycling rate occur, signifying switching failure. In our measurements, all failures were recoverable. This is attributed to the delicate and controlled nature of the measurement setup. Low programming energy minimizes probability of over-programming, which results in permanent irrecoverable failures. We showed that by increasing the programming pulse voltage, cycling could be recovered without degrading cycling rate, endurance or LRS/HRS distribution. Based on these observations, a programming protocol with recovery was proposed.

In Chapter 4, we investigated short-term stability, which is critical in programming RRAMs. We showed experimentally that stochastic resistance fluctuations are responsible for the rapid loss of the post-programmed resistance state. The fluctuations trigger the verification mechanism and result in resistance states below/above the HRS/LRS targets. Although relaxation effects are present and continue to widen the distribution, the initial distribution is caused primarily by the fluctuations. Analysis of the fluctuations with enhanced resolution and shorter times reveals new dynamics of the filament behavior immediately after RESET. We observe that the largest fluctuations during read occur more frequently at the start of the read operation, close

to the programming pulse, but decay to a non-zero value with time. However, the fluctuation amplitude does decrease within the same time duration. We rule out thermal effects as the source of this behavior both in the programming and read operations. Strategies to minimize the impact of fluctuations would require optimization of the target resistance states, read delay and slower read.

Metal-Organic Framework (MOF) based experimental Switching Devices

As part of future outlook, a novel non-traditional RRAM device based on a surface anchored metal organic framework (SURMOF) insulator layer has been fabricated and tested using the voltage sweep mode of a semiconductor parameter analyzer. HKUST-1 (Hong Kong Univeristy of Technology 1) is a typical, metal-organic framework composed of Cu2+ dimers connected by benzenetricarboxylate (BTC) units. This MOF forms a crystalline, 3D pore structure with cubic symmetry and a lattice constant of a = 26.343 Å. HKUST-1 MOF thin films (SURMOFs) can be grown on surface-functionalized Au-substrates by employing a layer-by-layer, quasi-epitaxial method. The resulting SURMOFs are monolithic and possess a high crystallinity with low defect densities, as well as excellent mechanical properties. In recent years, a number of works published by different groups have demonstrated interesting electrical transport properties of MOFs and MOF thin films. There are several ways to modify the electrical conductivity of MOFs. The first one involves the modification and optimization of the organic linkers used to assemble the framework materials, e.g. using transition metals bonded to organocyanide ligands such as tetracyano-quinodimethane (TCNQ) as well as dicyanoquinonediimine (DCNQI). A second possibility is to use the periodic array of pores connected by channels inside MOF materials to allow for the trapping and transport of ions inside the supramolecular network. Loading the pores with appropriate organic semiconductors

can change electrical properties (Ferrocene, TCNQ) which also allow ion transport in addition to electron transport. For the RRAM devices reported here, SURMOFs were grown on Au-substrates pre-functionalized by deposition of a MHDA (16-mercaptohexadecanoic acid, Sigma) SAM (self-assembled monolayer). On such modified substrates HKUST-1 SURMOFs were grown employing the LPE (Liquid Phase Epitaxy) spray method. In this case, a 1 mM ethanolic solution of copper acetate hydrate (metal source) 0.2mM solutions of the organic linker solution (1,3,5-benzenetricarboxylic acid (BTC)) were used in the spraying process. The SURMOF material used in this study is based on Hong Kong University of Science and Technology 1 (HKUST-1) films, with Cu^{2+} dimers connected by benzenetricarboxylate (BTC) units [175]. Furthermore, some of the MOF devices were doped with ferrocene to investigate the effect of loading on the sweep-mode switching performance and to benchmark against pristine unloaded MOF devices. In principle, ferrocene can be loaded inside the HKUST-1 SURMOFs either from the gas-phase or from the liquid-phase. For this study we used liquid phase ferrocene loading. Device fabrication was carried out by depositing top electrodes from either Cu or Au targets using a commercial e-beam deposition system. Figure 39 shows the completed MIM stack of two devices, device A (left) and B (right), used in this study. Device A consists of Au/ SURMOF (75 nm)/Cu/Au with top electrode Cu (200 nm)/Au (25 nm) patterned using a shadow mask of 200 µm in diameter. Device B consists of Au/SURMOF (10 nm: 20 nm : 50 nm)/Al /Au with top electrode Al (700 nm) /Au (100 nm) patterned with a shadow mask of 150 µm in diameter.

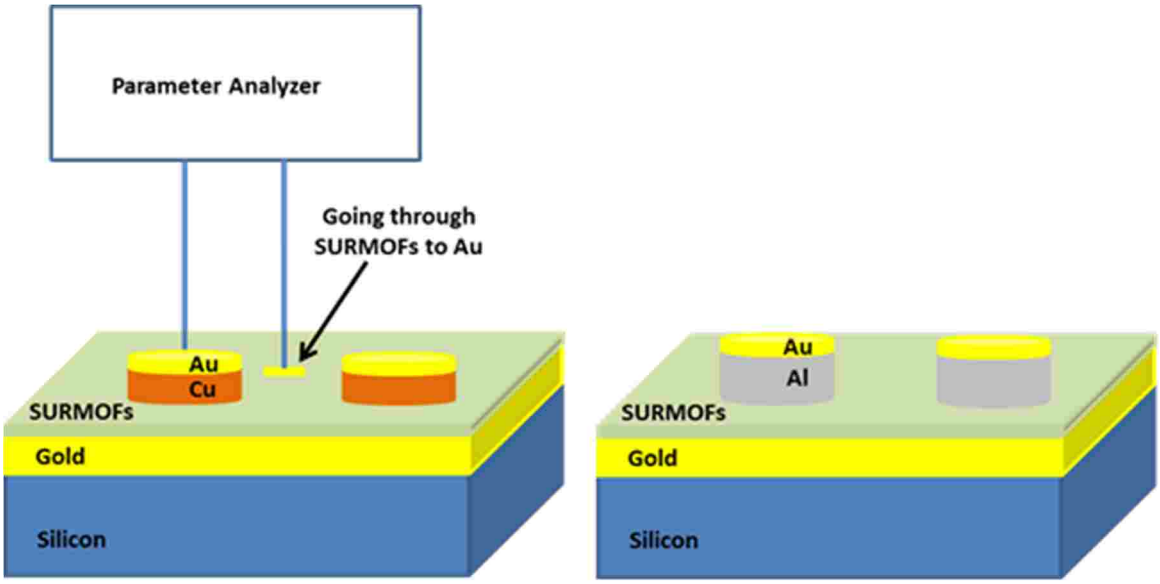

Figure 39. Illustration of SURMOF device test structure and measurement setup. Different electrode materials were used to investigate the contribution of the electrode to switching behavior.

Figure 40 shows the switching behavior of device A under voltage sweep mode. In all measurements, the voltage was applied to the top electrode (TE) with the bottom electrode (BE) grounded. The SURMOF devices did not require FORMING. A clear hysteresis pinched at zero is observed for loaded SURMOF devices. This I-V curve is indicative of resistive switching. The respective SET and RESET transition occur at negative bias and positive bias. This is not expected for a cell with an electrochemically active electrode. Normally, a positive bias on TE would result in a SET operation since Cu ions drift towards BE to create the conductive filamentary path. Observing the reverse behavior suggests that the SET is not a result of Cu ions from the TE. Therefore, the formation of a metal conductive filament can be ruled out.

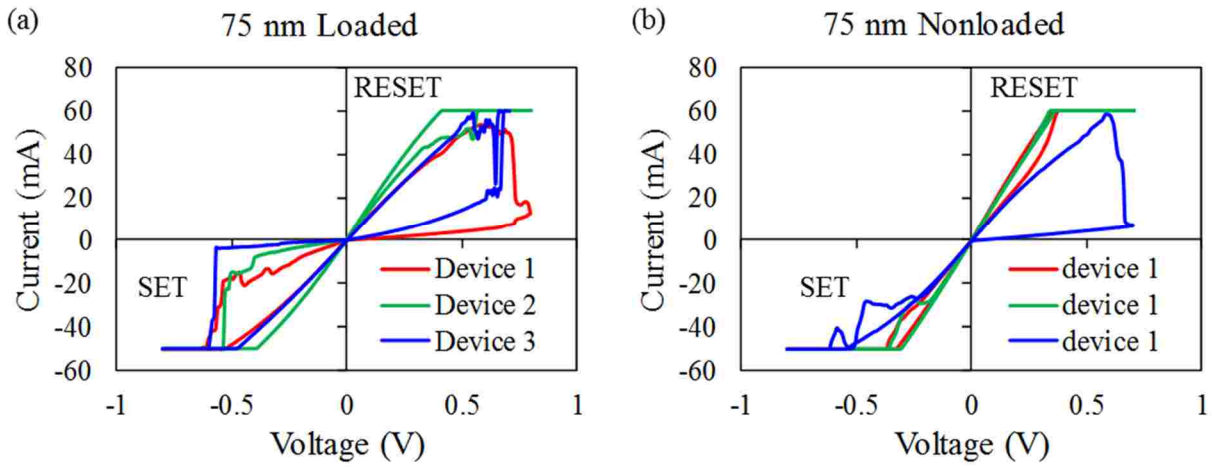

Figure 40. Measured I-V curve for SURMOF device A (a) loaded with ferrocene and (b) pristine nonloaded.

Figure 41 shows the switching behavior of Device B for all three SURMOF thicknesses (10 nm, 20 nm, 50 nm). All devices are loaded with ferrocene. Switching is observed for all investigated SURMOF film thicknesses. However, the ON state conductance appears to be increasing with thickness, which may be attributed to the increased number of ferrocene molecules in the switching layer. Two key observations are made here by comparing Device B to Device A. The resistance switching does not require an electrochemically active electrode since Device B has an Al/Au TE. Additionally, the switching direction is reversed in Device B; SET occurs on a positive sweep, and RESET occurs on a positive sweep. Switching direction has been shown to be dependent on the insulator type; n-type or p-type [14].

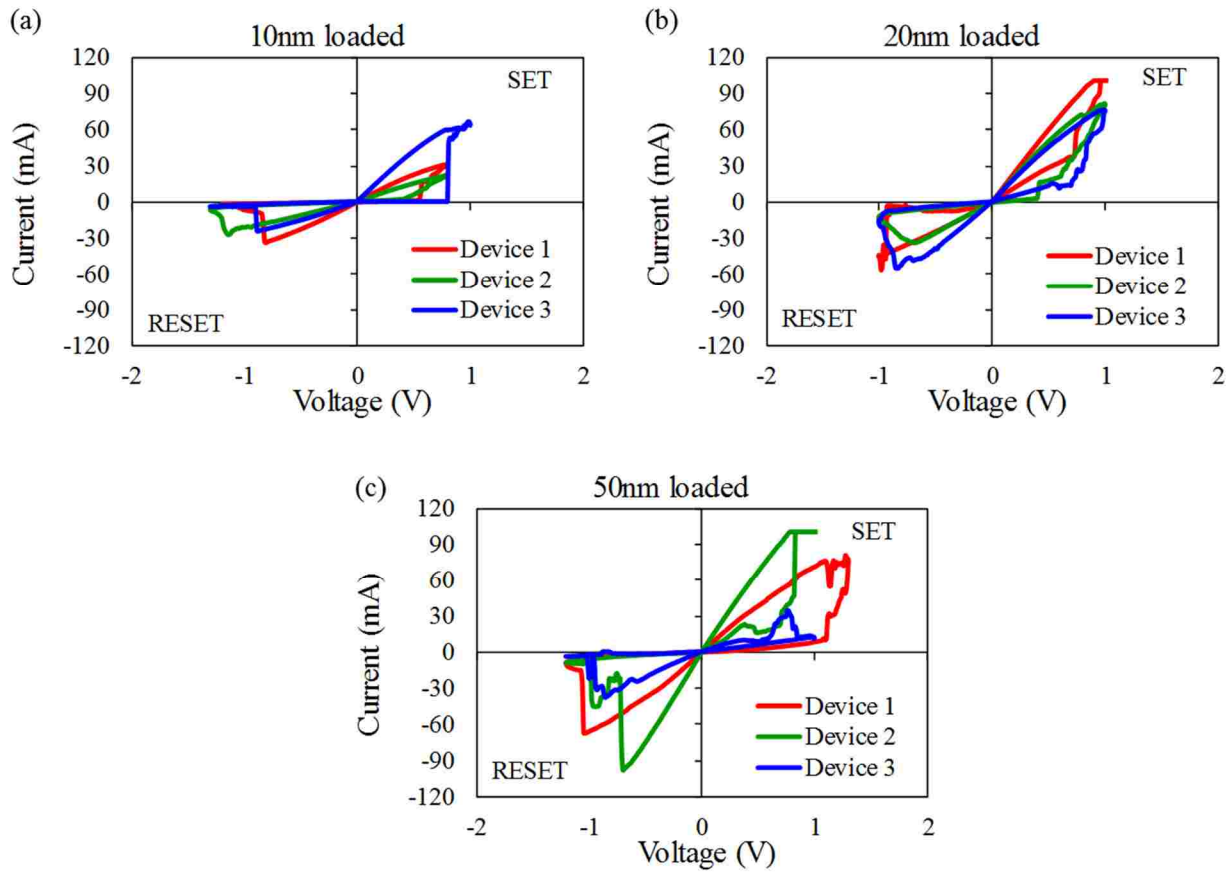

Figure 41. Measured I-V curve for device B with loaded SURMOF films of thickness (a)10 nm, (b) 20 nm and (c) 50 nm.

Figure 42 shows the On-state and OFF-state of loaded and unloaded 10nm SURMOF devices. Ferrocene loaded devices display an observable resistance window (separated ON and OFF), while pristine nonloaded SURMOF devices do not. It demonstrates the importance ferrocene loading plays in the resistive switching observed in these SURMOF devices.

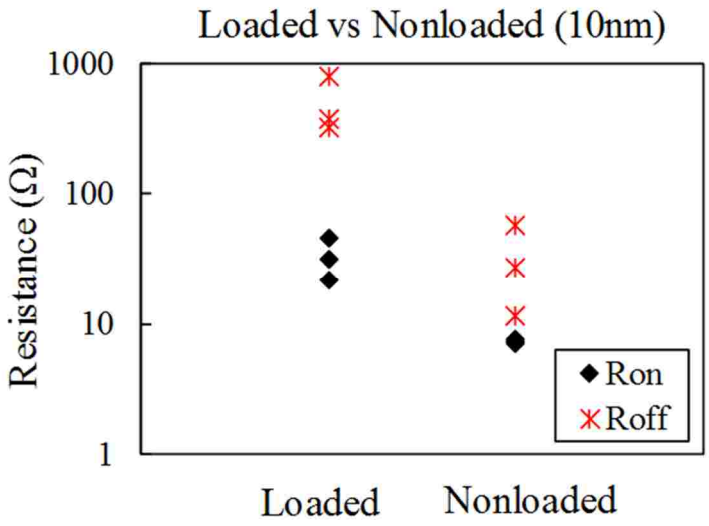

Figure 42. ON-state and OFF-state of loaded and pristine nonloaded 10 nm SURMOF devices. Loaded devices display a resistance window demonstrating the important of ferrocene for resistive switching.

This exploratory study demonstrates for the first time the resistive switching phenomena in KHUST-1 SURMOF devices. We demonstrated that resistive switching did not require an electrochemically active electrode and that ferrocene loading improved switching behavior. However, these results are preliminary and many further studies are needed to understand the possible mechanism responsible for RS and the statistics of switching behavior. The ON and OFF state conductance are very high, which is not desirable for low-power operation. Analysis of the switching behavior for each device has to be performed to study endurance and possible failure mechanism. In addition, the devices explored here are not suitable for pulse programming because of the large capacitance of the MIM structure. Smaller devices would have to be fabricated to test the high-speed performance of this novel RRAM device.

Future Work

In this work, several novel concepts were combined to study the high-speed performance of HfO_2-based RRAM devices in order to address the technological challenges, issues and serious barriers to RRAM memory cell commercialization. We demonstrated 1) a viable, current-control approach to achieving high endurance and 2) the usefulness of a DC read to study filament instability at microsecond delays from the programming instant. These demonstrations relied on an ultra-short, picosecond, programming pulse. The main novelty in this work is associated with programming RRAM devices without a series, current-limiting element. We postulated that this approach is advantageous since it limits the programming energy [73] – an approach that the majority of reported experimental setups do not use [58, 86, 159, 176]. However, we have yet to perform an extensive study on the impact of pulse amplitude and height in this compliance-free programming environment. Therefore, studies that involve a range of pulses – wide pulse (ns), short (ps), single-pulse (high V_{SET}/V_{RESET}), and multi-pulse operation (low V_{SET}/V_{RESET}) – are needed.

To gain a deeper understanding of the fluctuation and instability behavior of RRAMs, system modifications and different device stacks are essential. We will first discuss some of the future work related to system modifications. The CUSPP measurement tool uses 100 ps for the three programming operations: FORMING, SET and RESET. Since current control is provided by the pulse width, it is likely that even shorter pulse widths may result in finer resistance programming. For example, in some cases, device resistance changed by up to 9 orders of magnitude during FORMING. This large range suggests that even shorter pulses can minimize these large resistance transitions. Admittedly, fast pulse programming is limited by the size of

the memory array. However, these studies are critical to understanding the resistive switching mechanism.

The program-verify protocol in CUSPP uses identical positive/negative pulses (fixed amplitude and width) for SET/RESET. In Chapter 3, we proposed a recovery operation algorithm that relies on increasing the pulse amplitude in incremental steps. To accomplish this, modifications to the setup are required. First, a counter to track the programming time is essential to determine when to deploy the recovery protocol. Second, a voltage-controlled attenuator or amplifier would also be needed to incrementally adjust the programming pulse amplitude. With these inclusions, the proposed programming algorithm could be experimentally proven.

All measurements using CUSPP were performed on HfO_2-based devices called valance change memory (VCM) devices. The experimental setup, however, is not limited to HfO_2-based RRAM or VCM type cells. Recent studies of conductive-bridge RAM (CBRAM) suggest that these devices may be more stable compared to VCM [88]. There have also been reports that show that a Ta oxygen exchange layer (OEL) will create a more stable CF compared to the Ti OEL used in this work [120]. The stability of the filament is related to the diffusion of oxygen species and charge trap/detrap mechanisms, which are influenced by the composition of the CF. Therefore, it would be useful to explore the response of other RRAM devices using CUSPP. The short-term analysis and observation of fluctuation behavior can be used as a reliability metric to determine the appropriate optimized material choice for RRAMs.

REFERENCES

[1] T. W. Hickmott, "Low-Frequency Negative Resistance in Thin Anodic Oxide Films," *Journal of Applied Physics,* vol. 33, pp. 2669-2682, 1962.

[2] J. F. Gibbons and W. E. Beadle, "Switching properties of thin Nio films," *Solid-State Electronics,* vol. 7, pp. 785-790, 11// 1964.

[3] G. Dearnaley, A. M. Stoneham, and D. V. Morgan, "Electrical phenomena in amorphous oxide films," *Reports on Progress in Physics,* vol. 33, p. 1129, 1970.

[4] J. G. Simmons, "Conduction in thin dielectric films," *Journal of Physics D: Applied Physics,* vol. 4, p. 613, 1971.

[5] A. Asamitsu, Y. Tomioka, H. Kuwahara, and Y. Tokura, "Current switching of resistive states in magnetoresistive manganites," *Nature,* vol. 388, pp. 50-52, Jul 3 1997.

[6] A. Beck, J. G. Bednorz, C. Gerber, C. Rossel, and D. Widmer, "Reproducible switching effect in thin oxide films for memory applications," *Applied Physics Letters,* vol. 77, pp. 139-141, 2000.

[7] Y. Watanabe, "Electrical transport through Pb(Zr, Ti) O3 p-n and p-p heterstructures modulated by bound charges at a ferroelectric surface: ferroelectric p-n diode," *Physical Review B,* vol. 59, pp. 11257-11266, 05/01/ 1999.

[8] I. G. Baek, M. S. Lee, S. Seo, M. J. Lee, D. H. Seo, D. S. Suh*, et al.*, "Highly scalable nonvolatile resistive memory using simple binary oxide driven by asymmetric unipolar voltage pulses," in *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004.*, 2004, pp. 587-590.

[9] S. Seo, M. J. Lee, D. H. Seo, S. K. Choi, D.-S. Suh, Y. S. Joung*, et al.*, "Conductivity switching characteristics and reset currents in NiO films," *Applied Physics Letters,* vol. 86, p. 093509, 2005.

[10] C. Rohde, B. J. Choi, D. S. Jeong, S. Choi, J.-S. Zhao, and C. S. Hwang, "Identification of a determining parameter for resistive switching of TiO2 thin films," *Applied Physics Letters,* vol. 86, p. 262907, 2005.

[11] H. Y. Lee, P. S. Chen, T. Y. Wu, Y. S. Chen, C. C. Wang, P. J. Tzeng*, et al.*, "Low Power and High Speed Bipolar Switching with A Thin Reactive Ti Buffer Layer in Robust HfO2 Based RRAM," *Ieee International Electron Devices Meeting 2008, Technical Digest,* pp. 297-300, 2008.

[12] F. Miao, J. P. Strachan, J. J. Yang, M.-X. Zhang, I. Goldfarb, A. C. Torrezan*, et al.*, "Anatomy of a Nanoscale Conduction Channel Reveals the Mechanism of a High-Performance Memristor," *Advanced Materials,* vol. 23, pp. 5633-5640, 2011.

[13] H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, and P.-S. Chen, "Metal-oxide RRAM," *Proc IEEE.,* vol. 100, 2012.

[14] F. Pan, S. Gao, C. Chen, C. Song, and F. Zeng, "Recent progress in resistive random access memories: Materials, switching mechanisms, and performance," *Materials Science & Engineering R-Reports,* vol. 83, pp. 1-59, Sep 2014.

[15] S. Clima, K. Sankaran, Y. Y. Chen, A. Fantini, U. Celano, A. Belmonte*, et al.*, "RRAMs based on anionic and cationic switching: a short overview," *physica status solidi (RRL) – Rapid Research Letters,* vol. 8, pp. 501-511, 2014.

[16] R. Waser, R. Dittmann, M. Salinga, and M. Wuttig, "Function by defects at the atomic scale – New concepts for non-volatile memories," *Solid-State Electronics,* vol. 54, pp. 830-840, 9// 2010.

[17] I. Daniele, "Resistive switching memories based on metal oxides: mechanisms, reliability and scaling," *Semiconductor Science and Technology,* vol. 31, p. 063002, 2016.
[18] Z. Wei, Y. Kanzawa, K. Arita, Y. Katoh, K. Kawai, S. Muraoka*, et al.*, "Highly reliable TaOx ReRAM and direct evidence of redox reaction mechanism," in *2008 IEEE International Electron Devices Meeting*, 2008, pp. 1-4.
[19] J. P. Strachan, G. Medeiros-Ribeiro, J. J. Yang, M.-X. Zhang, F. Miao, I. Goldfarb*, et al.*, "Spectromicroscopy of tantalum oxide memristors," *Applied Physics Letters,* vol. 98, p. 242114, 2011.
[20] U. Celano, L. Goux, R. Degraeve, A. Fantini, O. Richard, H. Bender*, et al.*, "Imaging the Three-Dimensional Conductive Channel in Filamentary-Based Oxide Resistive Switching Memory," *Nano Letters,* vol. 15, pp. 7970-7975, 2015/12/09 2015.
[21] D.-H. Kwon, K. M. Kim, J. H. Jang, J. M. Jeon, M. H. Lee, G. H. Kim*, et al.*, "Atomic structure of conducting nanofilaments in TiO2 resistive switching memory," *Nat Nano,* vol. 5, pp. 148-153, 02//print 2010.
[22] X. Wu, D. Cha, M. Bosman, N. Raghavan, D. B. Migas, V. E. Borisenko*, et al.*, "Intrinsic nanofilamentation in resistive switching," *Journal of Applied Physics,* vol. 113, p. 114503, 2013.
[23] S. Kumar, Z. Wang, X. Huang, N. Kumari, N. Davila, J. P. Strachan*, et al.*, "Oxygen migration during resistance switching and failure of hafnium oxide memristors," *Applied Physics Letters,* vol. 110, p. 103503, 2017.
[24] S. Yu and H. S. P. Wong, "A Phenomenological Model for the Reset Mechanism of Metal Oxide RRAM," *IEEE Electron Device Letters,* vol. 31, pp. 1455-1457, 2010.
[25] P. Zhou, M. Yin, H. J. Wan, H. B. Lu, T. A. Tang, and Y. Y. Lin, "Role of TaON interface for CuxO resistive switching memory based on a combined model," *Applied Physics Letters,* vol. 94, p. 053510, 2009.
[26] D. S. Jeong, H. Schroeder, and R. Waser, "Coexistence of Bipolar and Unipolar Resistive Switching Behaviors in a Pt/TiO2/Pt Stack," *Electrochemical and Solid-State Letters,* vol. 10, pp. G51-G53, August 1, 2007 2007.
[27] G. Bersuker, D. C. Gilmer, D. Veksler, P. Kirsch, L. Vandelli, A. Padovani*, et al.*, "Metal oxide resistive memory switching mechanism based on conductive filament properties," *Journal of Applied Physics,* vol. 110, p. 124518, 2011.
[28] H. J. Wan, P. Zhou, L. Ye, Y. Y. Lin, T. A. Tang, H. M. Wu*, et al.*, "In Situ Observation of Compliance-Current Overshoot and Its Effect on Resistive Switching," *IEEE Electron Device Letters,* vol. 31, pp. 246-248, 2010.
[29] K. Kinoshita, K. Tsunoda, Y. Sato, H. Noshiro, S. Yagaki, M. Aoki*, et al.*, "Reduction in the reset current in a resistive random access memory consisting of NiOx brought about by reducing a parasitic capacitance," *Applied Physics Letters,* vol. 93, pp. -, 2008.
[30] Z. Wang, D. Nminibapiel, P. Shrestha, J. Liu, W. Guo, P. G. Weidler*, et al.*, "Resistive Switching Nanodevices Based on Metal–Organic Frameworks," *ChemNanoMat,* vol. 2, pp. 67-73, 2016.
[31] Y. Cong, Z. Chao, T. Tsung-Ming, C. Kuan-Chang, C. Min-Chen, C. Ting-Chang*, et al.*, "Low-power bipolar resistive switching TiN/HfO 2 /ITO memory with self-compliance current phenomenon," *Applied Physics Express,* vol. 7, p. 034101, 2014.
[32] K. Kentaro, T. Tetsuro, A. Masaki, S. Yoshihiro, and T. Hitoshi, "Lowering the Switching Current of Resistance Random Access Memory Using a Hetero Junction

Structure Consisting of Transition Metal Oxides," *Japanese Journal of Applied Physics,* vol. 45, p. L991, 2006.

[33] Y. Guo and J. Robertson, "Materials selection for oxide-based resistive random access memories," *Applied Physics Letters,* vol. 105, p. 223516, 2014.

[34] A. Belmonte, A. Fantini, A. Redolfi, M. Houssa, M. Jurczak, and L. Goux, "Excellent Roff/Ron ratio and short programming time in Cu/Al2O3-based conductive-bridging RAM under low-current (10 μA) operation," *physica status solidi (a),* vol. 213, pp. 302-305, 2016.

[35] B. Govoreanu, G. S. Kar, Y. Y. Chen, V. Paraschiv, S. Kubicek, A. Fantini*, et al.*, "10x10nm^2 Hf/HfOx Crossbar Resistive RAM with Excellent Performance, Reliability and Low-Energy Operation," *2011 Ieee International Electron Devices Meeting (Iedm),* 2011.

[36] Z. Zhang, Y. Wu, H. S. P. Wong, and S. S. Wong, "Nanometer-Scale HfOx RRAM," *IEEE Electron Device Letters,* vol. 34, pp. 1005-1007, 2013.

[37] R. Degraeve, A. Fantini, N. Raghavan, L. Goux, S. Clima, B. Govoreanu*, et al.*, "Causes and consequences of the stochastic aspect of filamentary RRAM," *Microelectronic Engineering,* vol. 147, pp. 171-175, Nov 1 2015.

[38] D. Ielmini, F. Nardi, and C. Cagli, "Resistance-dependent amplitude of random telegraph-signal noise in resistive switching memories," *Applied Physics Letters,* vol. 96, p. 053503, 2010.

[39] S. Balatti, S. Ambrogio, A. Cubeta, A. Calderoni, N. Ramaswamy, and D. Ielmini, "Voltage-dependent random telegraph noise (RTN) in HfOx resistive RAM," in *2014 IEEE International Reliability Physics Symposium*, 2014, pp. MY.4.1-MY.4.6.

[40] F. M. Puglisi, P. Pavan, A. Padovani, and L. Larcher, "A study on HfO2 RRAM in HRS based on I–V and RTN analysis," *Solid-State Electronics,* vol. 102, pp. 69-75, 12// 2014.

[41] F. M. Puglisi, P. Pavan, A. Padovani, L. Larcher, and G. Bersuker, "RTS noise characterization of HfOx RRAM in high resistive state," *Solid-State Electronics,* vol. 84, pp. 160-166, Jun 2013.

[42] S. Ambrogio, S. Balatti, A. Cubeta, A. Calderoni, N. Ramaswamy, and D. Ielmini, "Statistical Fluctuations in HfOx Resistive-Switching Memory: Part I - Set/Reset Variability," *IEEE Transactions on Electron Devices,* vol. 61, pp. 2912-2919, 2014.

[43] D. M. Nminibapiel, D. Veksler, P. R. Shrestha, J. H. Kim, J. P. Campbell, J. T. Ryan*, et al.*, "Characteristics of Resistive Memory Read Fluctuations in Endurance Cycling," *IEEE Electron Device Letters,* vol. 38, pp. 326-329, 2017.

[44] U. Russo, D. Kamalanathan, D. Ielmini, A. L. Lacaita, and M. N. Kozicki, "Study of Multilevel Programming in Programmable Metallization Cell (PMC) Memory," *IEEE Transactions on Electron Devices,* vol. 56, pp. 1040-1047, 2009.

[45] A. Prakash, J. Park, J. Song, J. Woo, E. J. Cha, and H. Hwang, "Demonstration of Low Power 3-bit Multilevel Cell Characteristics in a TaOx-Based RRAM by Stack Engineering," *IEEE Electron Device Letters,* vol. 36, pp. 32-34, 2015.

[46] J.-C. Liu, I.-T. Wang, C.-W. Hsu, W.-C. Luo, and T.-H. Hou, "Investigating MLC variation of filamentary and non-filamentary RRAM," *Symp VLSI Technol.,* vol. 1, 2014.

[47] C. Nauenheim, C. Kugeler, A. Rudiger, R. Waser, A. Flocke, and T. G. Noll, "Nano-Crossbar Arrays for Nonvolatile Resistive RAM (RRAM) Applications," in *2008 8th IEEE Conference on Nanotechnology*, 2008, pp. 464-467.

[48] S. H. Jo, K.-H. Kim, and W. Lu, "High-Density Crossbar Arrays Based on a Si Memristive System," *Nano Letters,* vol. 9, pp. 870-874, 2009/02/11 2009.

[49] T. Y. Liu, T. H. Yan, R. Scheuerlein, Y. Chen, J. K. Lee, G. Balakrishnan*, et al.*, "A 130.7-mm^(2) 2-layer 32-Gb ReRAM Memory Device in 24-nm Technology," *IEEE Journal of Solid-State Circuits,* vol. 49, pp. 140-153, 2014.

[50] A. Benoist, S. Blonkowski, S. Jeannot, S. Denorme, J. Damiens, J. Berger*, et al.*, "28nm advanced CMOS resistive RAM solution as embedded non-volatile memory," in *2014 IEEE International Reliability Physics Symposium*, 2014, pp. 2E.6.1-2E.6.5.

[51] A. Sawa, "Resistive switching in transition metal oxides," *Materials Today,* vol. 11, pp. 28-36, Jun 2008.

[52] S. Yu, H.-Y. Chen, B. Gao, J. Kang, and H. S. P. Wong, "HfOx-Based Vertical Resistive Switching Random Access Memory Suitable for Bit-Cost-Effective Three-Dimensional Cross-Point Architecture," *ACS Nano,* vol. 7, pp. 2320-2325, 2013/03/26 2013.

[53] D. S. Jeong, R. Thomas, R. S. Katiyar, J. F. Scott, H. Kohlstedt, A. Petraru*, et al.*, "Emerging memories: resistive switching mechanisms and current status," *Reports on Progress in Physics,* vol. 75, Jul 2012.

[54] E. Linn, R. Rosezin, C. Kugeler, and R. Waser, "Complementary resistive switches for passive nanocrossbar memories," *Nat Mater,* vol. 9, pp. 403-406, 05//print 2010.

[55] D. Walczyk, C. Walczyk, T. Schroeder, T. Bertaud, M. Sowińska, M. Lukosius*, et al.*, "Resistive switching characteristics of CMOS embedded HfO2-based 1T1R cells," *Microelectronic Engineering,* vol. 88, pp. 1133-1135, July 2011.

[56] Y. Li, H. Lv, Q. Liu, S. Long, M. Wang, H. Xie*, et al.*, "Bipolar one diode-one resistor integration for high-density resistive memory applications," *Nanoscale,* vol. 5, pp. 4785-4789, 2013.

[57] D. Niu, C. Xu, N. Muralimanohar, N. P. Jouppi, and Y. Xie, "Design of cross-point metal-oxide ReRAM emphasizing reliability and cost," in *2013 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2013, pp. 17-23.

[58] M. J. Lee, C. B. Lee, D. Lee, S. R. Lee, M. Chang, J. H. Hur*, et al.*, "A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta2O5-x/TaO2-x bilayer structures," *Nature Materials,* vol. 10, pp. 625-630, Aug 2011.

[59] H. Zhang, L. Liu, B. Gao, Y. Qiu, X. Liu, J. Lu*, et al.*, "Gd-doping effect on performance of HfO2 based resistive switching memory devices using implantation approach," *Applied Physics Letters,* vol. 98, p. 042105, 2011.

[60] Z. Wang, W. G. Zhu, A. Y. Du, L. Wu, Z. Fang, X. A. Tran*, et al.*, "Highly Uniform, Self-Compliance, and Forming-Free ALD HfO2-Based RRAM With Ge Doping," *IEEE Transactions on Electron Devices,* vol. 59, pp. 1203-1208, 2012.

[61] B. J. Choi, A. C. Torrezan, K. J. Norris, F. Miao, J. P. Strachan, M.-X. Zhang*, et al.*, "Electrical Performance and Scalability of Pt Dispersed SiO2 Nanometallic Resistance Switch," *Nano Letters,* vol. 13, pp. 3213-3217, 2013/07/10 2013.

[62] H. Y. Lee, Y. S. Chen, P. S. Chen, P. Y. Gu, Y. Y. Hsu, S. M. Wang*, et al.*, "Evidence and solution of over-RESET problem for HfOx based resistive memory with sub-ns switching speed and high endurance," in *2010 International Electron Devices Meeting*, 2010, pp. 19.7.1-19.7.4.

[63] Y.-S. Chen, P.-S. Chen, H.-Y. Lee, T.-Y. Wu, K.-H. Tsai, F. Chen*, et al.*, "Enhanced endurance reliability and low current operation for AlOx/HfOx based unipolar RRAM with Ni electrode," *Solid-State Electronics,* vol. 94, pp. 1-5, July 2014.

[64] H. Wu, X. Li, M. Wu, F. Huang, Z. Yu, and H. Qian, "Resistive Switching Performance Improvement of Ta2O5-x/TaOy Bilayer ReRAM Devices by Inserting AlO Barrier Layer," *IEEE Electron Device Letters,* vol. 35, pp. 39-41, 2014.

[65] B. Traoré, P. Blaise, E. Vianello, H. Grampeix, S. Jeannot, L. Perniola*, et al.*, "On the Origin of Low-Resistance State Retention Failure in HfO2-Based RRAM and Impact of Doping/Alloying," *IEEE Transactions on Electron Devices,* vol. 62, pp. 4029-4036, 2015.

[66] S. Ning, T. O. Iwasaki, and K. Takeuchi, "50 nm AlxOy ReRAM program 31% energy, 1.6× endurance, and 3.6× speed improvement by advanced cell condition adaptive verify-reset," *Solid-State Electronics,* vol. 103, pp. 64-72, Jan. 2015.

[67] H. Liu, H. Lv, B. Yang, X. Xu, R. Liu, Q. Liu*, et al.*, "Uniformity Improvement in 1T1R RRAM With Gate Voltage Ramp Programming," *IEEE Electron Device Letters,* vol. 35, pp. 1224-1226, 2014.

[68] Y. Y. Chen, B. Govoreanu, L. Goux, R. Degraeve, A. Fantini, G. S. Kar*, et al.*, "Balancing SET/RESET Pulse for > 10^10 endurance in HfO2/Hf 1T1R Bipolar RRAM," *IEEE Transactions on Electron Devices,* vol. 59, pp. 3243-3249, 2012.

[69] Y. S. Chen, H. Y. Lee, P. S. Chen, W. H. Liu, S. M. Wang, P. Y. Gu*, et al.*, "Robust High-Resistance State and Improved Endurance of HfOX Resistive Memory by Suppression of Current Overshoot," *Ieee Electron Device Letters,* vol. 32, pp. 1585-1587, Nov 2011.

[70] R. Waser, R. Dittmann, G. Staikov, and K. Szot, "Redox-Based Resistive Switching Memories - Nanoionic Mechanisms, Prospects, and Challenges," *Advanced Materials,* vol. 21, pp. 2632-+, Jul 13 2009.

[71] Y.-T. Su, K.-C. Chang, T.-C. Chang, T.-M. Tsai, R. Zhang, J. C. Lou*, et al.*, "Characteristics of hafnium oxide resistance random access memory with different setting compliance current," *Applied Physics Letters,* vol. 103, p. 163502, 2013.

[72] S. Koveshnikov, K. Matthews, K. Min, D. C. Gilmer, M. G. Sung, S. Deora*, et al.*, "Real-time study of switching kinetics in integrated 1T/ HfOx 1R RRAM: Intrinsic tunability of set/reset voltage and trade-off with switching time," in *2012 International Electron Devices Meeting*, 2012, pp. 20.4.1-20.4.3.

[73] P. Shrestha, D. M. Nminibapiel, J. H. Kim, H. Baumgart, K. P. Cheung, and J. P. Campbell, "(Invited) Compliance-Free Pulse Forming of Filamentary RRAM," *ECS Transactions,* vol. 75, pp. 81-92, August 25, 2016 2016.

[74] B. Chen, B. Gao, S. W. Sheng, L. F. Liu, X. Y. Liu, Y. S. Chen*, et al.*, "A Novel Operation Scheme for Oxide-Based Resistive-Switching Memory Devices to Achieve Controlled Switching Behaviors," *IEEE Electron Device Letters,* vol. 32, pp. 282-284, 2011.

[75] G. Wang, S. Long, Z. Yu, M. Zhang, Y. Li, D. Xu*, et al.*, "Impact of program/erase operation on the performances of oxide-based resistive switching memory," *Nanoscale Research Letters,* vol. 10, p. 39, 2015.

[76] P. R. Shrestha, D. Nminibapiel, J. H. Kim, J. P. Campbell, K. P. Cheung, S. Deora*, et al.*, "Energy control paradigm for compliance-free reliable operation of RRAM," in *2014 IEEE International Reliability Physics Symposium*, 2014, pp. MY.10.1-MY.10.4.

[77] X. Xu, H. Lv, H. Liu, T. Gong, G. Wang, M. Zhang*, et al.*, "Superior Retention of Low-Resistance State in Conductive Bridge Random Access Memory With Single Filament Formation," *IEEE Electron Device Letters,* vol. 36, pp. 129-131, 2015.

[78] B. Butcher, G. Bersuker, K. G. Young-Fisher, D. C. Gilmer, A. Kalantarian, Y. Nishi, *et al.*, "Hot Forming to Improve Memory Window and Uniformity of Low-Power HfOx-Based RRAMs," in *2012 4th IEEE International Memory Workshop*, 2012, pp. 1-4.

[79] G. Wang, S. Long, M. Zhang, Y. Li, X. Xu, H. Liu, *et al.*, "Operation methods of resistive random access memory," *Science China Technological Sciences,* vol. 57, pp. 2295-2304, 2014.

[80] B. Chen, Y. Lu, B. Gao, Y. H. Fu, F. F. Zhang, P. Huang, *et al.*, "Physical Mechanisms of Endurance Degradation in TMO-RRAM," *2011 Ieee International Electron Devices Meeting (Iedm),* 2011.

[81] Y. Lu, B. Chen, B. Gao, Z. Fang, Y. H. Fu, J. Q. Yang, *et al.*, "Improvement of endurance degradation for oxide based resistive switching memory devices correlated with oxygen vacancy accumulation effect," in *2012 IEEE International Reliability Physics Symposium (IRPS)*, 2012, pp. MY.4.1-MY.4.4.

[82] S. Balatti, S. Ambrogio, Z. Wang, S. Sills, A. Calderoni, N. Ramaswamy, *et al.*, "Voltage-Controlled Cycling Endurance of HfOx-Based Resistive-Switching Memory," *IEEE Transactions on Electron Devices,* vol. 62, pp. 3365-3372, 2015.

[83] R. Waser, S. Menzel, and V. Rana, "Recent Progress in Redox-Based Resistive Switching," *2012 Ieee International Symposium on Circuits and Systems (Iscas 2012),* pp. 1596-1599, 2012.

[84] S. R. Lee, Y. B. Kim, M. Chang, K. M. Kim, C. B. Lee, J. H. Hur, *et al.*, "Multi-level switching of triple-layered TaOx RRAM with excellent reliability for storage class memory," in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, pp. 71-72.

[85] G. M. Wang, S. B. Long, Z. A. Yu, M. Y. Zhang, T. C. Ye, Y. Li, *et al.*, "Improving resistance uniformity and endurance of resistive switching memory by accurately controlling the stress time of pulse program operation," *Applied Physics Letters,* vol. 106, Mar 2 2015.

[86] F. M. Puglisi, C. Wenger, and P. Pavan, "A Novel Program-Verify Algorithm for Multi-Bit Operation in HfO2 RRAM," *Ieee Electron Device Letters,* vol. 36, pp. 1030-1032, Oct 2015.

[87] Y. Meng, X. Y. Xue, Y. L. Song, J. G. Yang, B. A. Chen, Y. Y. Lin, *et al.*, "Fast step-down set algorithm of resistive switching memory with low programming energy and significant reliability improvement," in *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers*, 2014, pp. 1-2.

[88] A. Belmonte, A. Fantini, A. Redolfi, M. Houssa, M. Jurczak, and L. Goux, "Optimization of the write algorithm at low-current (10 uA) in Cu/Al2O3-based conductive-bridge RAM," in *2015 45th European Solid State Device Research Conference (ESSDERC)*, 2015, pp. 114-117.

[89] X. G. Chen, X. B. Ma, Y. B. Yang, L. P. Chen, G. C. Xiong, G. J. Lian, *et al.*, "Comprehensive study of the resistance switching in SrTiO3 and Nb-doped SrTiO3," *Applied Physics Letters,* vol. 98, p. 122102, 2011.

[90] Z. B. Yan and J. M. Liu, "Coexistence of high performance resistance and capacitance memory based on multilayered metal-oxide structures," *Scientific Reports,* vol. 3, p. 2482, 2013.

[91] R. Meyer, L. Schloss, J. Brewer, R. Lambertson, W. Kinney, J. Sanchez, *et al.*, "Oxide dual-layer memory element for scalable non-volatile cross-point memory technology," in *2008 9th Annual Non-Volatile Memory Technology Symposium (NVMTS)*, 2008, pp. 1-5.

[92] Z. Fang, H. Y. Yu, W. J. Fan, G. Ghibaudo, J. Buckley, B. DeSalvo, *et al.*, "Current Conduction Model for Oxide-Based Resistive Random Access Memory Verified by Low-Frequency Noise Analysis," *IEEE Transactions on Electron Devices,* vol. 60, pp. 1272-1275, 2013.

[93] B. J. Choi, D. S. Jeong, S. K. Kim, C. Rohde, S. Choi, J. H. Oh, *et al.*, "Resistive switching mechanism of TiO2 thin films grown by atomic-layer deposition," *Journal of Applied Physics,* vol. 98, p. 033715, 2005.

[94] J. Y. Son and Y.-H. Shin, "Direct observation of conducting filaments on resistive switching of NiO thin films," *Applied Physics Letters,* vol. 92, p. 222106, 2008.

[95] M.-J. Lee, S. Han, S. H. Jeon, B. H. Park, B. S. Kang, S.-E. Ahn, *et al.*, "Electrical Manipulation of Nanofilaments in Transition-Metal Oxides for Resistance-Based Memory," *Nano Letters,* vol. 9, pp. 1476-1481, 2009.

[96] J. P. Strachan, M. D. Pickett, J. J. Yang, S. Aloni, A. L. David Kilcoyne, G. Medeiros-Ribeiro, *et al.*, "Direct Identification of the Conducting Channels in a Functioning Memristive Device," *Advanced Materials,* vol. 22, pp. 3573-3577, 2010.

[97] K. Jung, H. Seo, Y. Kim, H. Im, J. Hong, J.-W. Park, *et al.*, "Temperature dependence of high- and low-resistance bistable states in polycrystalline NiO films," *Applied Physics Letters,* vol. 90, p. 052104, 2007.

[98] M.-J. Lee, S. I. Kim, C. B. Lee, H. Yin, S.-E. Ahn, B. S. Kang, *et al.*, "Low-Temperature-Grown Transition Metal Oxide Based Storage Materials and Oxide Transistors for High-Density Non-volatile Memory," *Advanced Functional Materials,* vol. 19, pp. 1587-1593, 2009.

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

[100] S. Gao, C. Song, C. Chen, F. Zeng, and F. Pan, "Formation process of conducting filament in planar organic resistive memory," *Applied Physics Letters,* vol. 102, p. 141606, 2013.

[101] Y. Yang, P. Gao, L. Li, X. Pan, S. Tappertzhofen, S. Choi, *et al.*, "Electrochemical dynamics of nanoscale metallic inclusions in dielectrics," *Nature Communications,* vol. 5, p. 4232, 2014.

[102] R. Waser and M. Aono, "Nanoionics-based resistive switching memories," *Nature Materials,* vol. 6, pp. 833-840, Nov 2007.

[103] C. Chen, C. Song, J. Yang, F. Zeng, and F. Pan, "Oxygen migration induced resistive switching effect and its thermal stability in W/TaOx/Pt structure," *Applied Physics Letters,* vol. 100, p. 253509, 2012.

[104] F. Kurnia, Hadiyawarman, C. U. Jung, R. Jung, and C. Liu, "Composition dependence of unipolar resistance switching in TaOx thin films," *physica status solidi (RRL) – Rapid Research Letters,* vol. 5, pp. 253-255, 2011.

[105] C. Chen, S. Gao, F. Zeng, G. S. Tang, S. Z. Li, C. Song, *et al.*, "Migration of interfacial oxygen ions modulated resistive switching in oxide-based memory devices," *Journal of Applied Physics,* vol. 114, p. 014502, 2013.

[106] D. Ielmini, R. Bruchhaus, and R. Waser, "Thermochemical resistive switching: materials, mechanisms, and scaling projections," *Phase Transitions,* vol. 84, pp. 570-602, 2011.

[107] J.-Y. Chen, C.-L. Hsin, C.-W. Huang, C.-H. Chiu, Y.-T. Huang, S.-J. Lin*, et al.*, "Dynamic Evolution of Conducting Nanofilament in Resistive Switching Memories," *Nano Letters,* vol. 13, pp. 3671-3677, 2013.

[108] G.-S. Park, X.-S. Li, D.-C. Kim, R.-J. Jung, M.-J. Lee, and S. Seo, "Observation of electric-field induced Ni filament channels in polycrystalline NiOx film," *Applied Physics Letters,* vol. 91, p. 222103, 2007.

[109] J. Kang and I. S. Park, "Asymmetric Current Behavior on Unipolar Resistive Switching in Pt/HfO2/Pt Resistor With Symmetric Electrodes," *IEEE Transactions on Electron Devices,* vol. 63, pp. 2380-2383, 2016.

[110] N. Xu, L. F. Liu, X. Sun, X. Y. Liu, D. D. Han, Y. Wang*, et al.*, "Characteristics and mechanism of conduction/set process in TiN/ZnO/Pt resistance switching random-access memories," *Applied Physics Letters,* vol. 92, Jun 9 2008.

[111] G.-H. Buh, I. Hwang, and B. H. Park, "Time-dependent electroforming in NiO resistive switching devices," *Applied Physics Letters,* vol. 95, p. 142101, 2009.

[112] R. Degraeve, G. Groeseneken, R. Bellens, J. L. Ogier, M. Depas, P. J. Roussel*, et al.*, "New insights in the relation between electron trap generation and the statistical properties of oxide breakdown," *IEEE Transactions on Electron Devices,* vol. 45, pp. 904-911, 1998.

[113] C. Schindler, G. Staikov, and R. Waser, "Electrode kinetics of Cu–SiO2-based resistive switching cells: Overcoming the voltage-time dilemma of electrochemical metallization memories," *Applied Physics Letters,* vol. 94, p. 072109, 2009.

[114] B. Chakrabarti, R. V. Galatage, and E. M. Vogel, "Multilevel Switching in Forming-Free Resistive Memory Devices With Atomic Layer Deposited HfTiOx Nanolaminate," *IEEE Electron Device Letters,* vol. 34, pp. 867-869, 2013.

[115] F. Kurnia, C. Liu, C. U. Jung, and B. W. Lee, "The evolution of conducting filaments in forming-free resistive switching Pt/TaOx/Pt structures," *Applied Physics Letters,* vol. 102, p. 152902, 2013.

[116] X. Cao, X. Li, X. Gao, W. Yu, X. Liu, Y. Zhang*, et al.*, "Forming-free colossal resistive switching effect in rare-earth-oxide Gd2O3 films for memristor applications," *Journal of Applied Physics,* vol. 106, p. 073723, 2009.

[117] U. Russo, D. Ielmini, C. Cagli, A. L. Lacaita, S. Spiga, C. Wiemer*, et al.*, "Conductive-filament switching analysis and self-accelerated thermal dissolution model for reset in NiO-based RRAM," in *2007 IEEE International Electron Devices Meeting*, 2007, pp. 775-778.

[118] B. Gao, S. Yu, N. Xu, L. F. Liu, B. Sun, X. Y. Liu*, et al.*, "Oxide-based RRAM switching mechanism: A new ion-transport-recombination model," in *2008 IEEE International Electron Devices Meeting*, 2008, pp. 1-4.

[119] B. Gao, J. F. Kang, Y. S. Chen, F. F. Zhang, B. Chen, P. Huang*, et al.*, "Oxide-based RRAM: Unified microscopic principle for both unipolar and bipolar switching," in *2011 International Electron Devices Meeting*, 2011, pp. 17.4.1-17.4.4.

[120] Y. Y. Chen, L. Goux, S. Clima, B. Govoreanu, R. Degraeve, G. S. Kar*, et al.*, "Endurance/Retention Trade-off on HfO2-metal Cap 1T1R Bipolar RRAM," *IEEE Transactions on Electron Devices,* vol. 60, pp. 1114-1121, 2013.

[121] J. J. Yang, M.-X. Zhang, J. P. Strachan, F. Miao, M. D. Pickett, R. D. Kelley*, et al.*, "High switching endurance in TaOx memristive devices," *Applied Physics Letters,* vol. 97, p. 232102, 2010.

[122] A. Chen, S. Haddad, Y. C. Wu, T. N. Fang, S. Kaza, and Z. Lan, "Erasing characteristics of Cu2O metal-insulator-metal resistive switching memory," *Applied Physics Letters,* vol. 92, p. 013503, 2008.

[123] E. A. Miranda, C. Walczyk, C. Wenger, and T. Schroeder, "Model for the Resistive Switching Effect in HfO2 MIM Structures Based on the Transmission Properties of Narrow Constrictions," *IEEE Electron Device Letters,* vol. 31, pp. 609-611, 2010.

[124] C. Li, B. Gao, Y. Yao, X. Guan, X. Shen, Y. Wang*, et al.*, "Direct Observations of Nanofilament Evolution in Switching Processes in HfO2-Based Resistive Random Access Memory by In Situ TEM Studies," *Advanced Materials,* vol. 29, pp. 1602976-n/a, 2017.

[125] F. De Stefano, M. Houssa, V. V. Afanas'ev, J. A. Kittl, M. Jurczak, and A. Stesmans, "Nature of the filament formed in HfO2-based resistive random access memory," *Thin Solid Films,* vol. 533, pp. 15-18, 4/30/ 2013.

[126] K. Frohlich, P. Jancovic, B. Hudec, J. Derer, A. Paskaleva, T. Bertaud*, et al.*, "Atomic layer deposition of thin oxide films for resistive switching," *Atomic Layer Deposition Applications 9,* vol. 58, pp. 163-170, 2013.

[127] Y.-M. Kim and J.-S. Lee, "Reproducible resistance switching characteristics of hafnium oxide-based nonvolatile memory devices," *Journal of Applied Physics,* vol. 104, p. 114115, 2008.

[128] C. Wenger, M. Lukosius, H.-J. Müssig, G. Ruhl, S. Pasko, and C. Lohe, "Influence of the electrode material on HfO2 metal-insulator-metal capacitors," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena,* vol. 27, pp. 286-289, 2009.

[129] M. Lanza, G. Bersuker, M. Porti, E. Miranda, M. Nafría, and X. Aymerich, "Resistive switching in hafnium dioxide layers: Local phenomenon at grain boundaries," *Applied Physics Letters,* vol. 101, p. 193502, 2012.

[130] I. K. Yoo, B. S. Kang, S. E. Ahn, C. B. Lee, M. J. Lee, G. S. Park*, et al.*, "Fractal Dimension of Conducting Paths in Nickel Oxide (NiO) Thin Films During Resistance Switching," *IEEE Transactions on Nanotechnology,* vol. 9, pp. 131-133, 2010.

[131] K. Kyung Min, J. Doo Seok, and H. Cheol Seong, "Nanofilamentary resistive switching in binary oxide system; a review on the present status and outlook," *Nanotechnology,* vol. 22, p. 254002, 2011.

[132] S. U. Sharath, T. Bertaud, J. Kurian, E. Hildebrandt, C. Walczyk, P. Calka*, et al.*, "Towards forming-free resistive switching in oxygen engineered HfO2−x," *Applied Physics Letters,* vol. 104, p. 063502, 2014.

[133] S. M. Yu, B. Gao, H. B. Dai, B. Sun, L. F. Liu, X. Y. Liu*, et al.*, "Improved Uniformity of Resistive Switching Behaviors in HfO2 Thin Films with Embedded Al Layers," *Electrochemical and Solid State Letters,* vol. 13, pp. H36-H38, 2010.

[134] L. Goux, X. P. Wang, Y. Y. Chen, L. Pantisano, N. Jossart, B. Govoreanu*, et al.*, "Roles and Effects of TiN and Pt Electrodes in Resistive-Switching HfO2 Systems," *Electrochemical and Solid-State Letters,* vol. 14, pp. H244-H246, Jun 2011.

[135] T. Bertaud, D. Walczyk, C. Walczyk, S. Kubotsch, M. Sowinska, T. Schroeder*, et al.*, "Resistive switching of HfO2-based Metal–Insulator–Metal diodes: Impact of the top electrode material," *Thin Solid Films,* vol. 520, pp. 4551-4555, May 2012.

[136] C. Cagli, J. Buckley, V. Jousseaume, T. Cabout, A. Salaun, H. Grampeix, *et al.*, "Experimental and theoretical study of electrode effects in HfO2 based RRAM," in *2011 International Electron Devices Meeting*, 2011, pp. 28.7.1-28.7.4.

[137] P. Y. Gu, Y. S. Chen, H. Y. Lee, P. S. Chen, W. H. Liu, W. S. Chen, *et al.*, "Scalability with silicon nitride encapsulation layer for Ti/HfOx pillar RRAM," in *Proceedings of 2010 International Symposium on VLSI Technology, System and Application*, 2010, pp. 146-147.

[138] Z. Fang, H. Y. Yu, X. Li, N. Singh, G. Q. Lo, and D. L. Kwong, "HfOx/TIOx/HfOx/TiOx Multilayer-Based Forming-Free RRAM Devices With Excellent Uniformity," *IEEE Electron Device Letters,* vol. 32, pp. 566-568, 2011.

[139] M. Sowinska, T. Bertaud, D. Walczyk, S. Thiess, M. A. Schubert, M. Lukosius, *et al.*, "Hard x-ray photoelectron spectroscopy study of the electroforming in Ti/HfO2-based resistive switching structures," *Applied Physics Letters,* vol. 100, p. 233509, 2012.

[140] Y. Y. Chen, L. Goux, S. Clima, B. Govoreanu, R. Degraeve, G. S. Kar, *et al.*, "Endurance/Retention Trade-off on HfO2/Metal Cap 1T1R Bipolar RRAM," *IEEE Transactions on Electron Devices,* vol. 60, pp. 1114-1121, 2013.

[141] Y. S. Chen, H. Y. Lee, P. S. Chen, P. Y. Gu, C. W. Chen, W. P. Lin, *et al.*, "Highly scalable hafnium oxide memory with improvements of resistive distribution and read disturb immunity," in *2009 IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1-4.

[142] T. Bertaud, M. Sowinska, D. Walczyk, S. Thiess, A. Gloskovskii, C. Walczyk, *et al.*, "In-operando and non-destructive analysis of the resistive switching in the Ti/HfO2/TiN-based system by hard x-ray photoelectron spectroscopy," *Applied Physics Letters,* vol. 101, p. 143501, 2012.

[143] S. Privitera, G. Bersuker, B. Butcher, A. Kalantarian, S. Lombardo, C. Bongiorno, *et al.*, "Microscopy study of the conductive filament in HfO2 resistive switching memory devices," *Microelectronic Engineering,* vol. 109, pp. 75-78, Sep 2013.

[144] R. Degraeve, A. Fantini, S. Clima, B. Govoreanu, L. Goux, Y. Y. Chen, *et al.*, "Dynamic 'hour glass' model for SET and RESET in HfO2 RRAM," in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, pp. 75-76.

[145] L. Zhao, H. Y. Chen, S. C. Wu, Z. Jiang, S. Yu, T. H. Hou, *et al.*, "Multi-level control of conductive nano-filament evolution in HfO2 ReRAM by pulse-train operations," *Nanoscale,* vol. 6, pp. 5698-5702, 2014.

[146] S. Yu, R. Jeyasingh, W. Yi, and H. S. P. Wong, "Understanding the conduction and switching mechanism of metal oxide RRAM through low frequency noise and AC conductance measurement and analysis," in *2011 International Electron Devices Meeting*, 2011, pp. 12.1.1-12.1.4.

[147] C. Walczyk, D. Walczyk, T. Schroeder, T. Bertaud, M. Sowinska, M. Lukosius, *et al.*, "Impact of Temperature on the Resistive Switching Behavior of Embedded HfO2-Based RRAM Devices," *IEEE Transactions on Electron Devices,* vol. 58, pp. 3124-3131, 2011.

[148] W. Yan, L. Qi, L. Shibing, W. Wei, W. Qin, Z. Manhong, *et al.*, "Investigation of resistive switching in Cu-doped HfO 2 thin film for multilevel non-volatile memory applications," *Nanotechnology,* vol. 21, p. 045202, 2010.

[149] U. Chand, C.-Y. Huang, J.-H. Jieng, W.-Y. Jang, C.-H. Lin, and T.-Y. Tseng, "Suppression of endurance degradation by utilizing oxygen plasma treatment in HfO2 resistive switching memory," *Applied Physics Letters,* vol. 106, p. 153502, 2015.

[150] S. S. Sheu, K. H. Cheng, M. F. Chang, P. C. Chiang, W. P. Lin, H. Y. Lee*, et al.*, "Fast-Write Resistive RAM (RRAM) for Embedded Applications," *IEEE Design & Test of Computers,* vol. 28, pp. 64-71, 2011.

[151] C. T. Antonio, S. John Paul, M.-R. Gilberto, and R. S. Williams, "Sub-nanosecond switching of a tantalum oxide memristor," *Nanotechnology,* vol. 22, p. 485203, 2011.

[152] P. R. Shrestha, D. Nminibapiel, J. H. Kim, J. P. Campbell, K. P. Cheung, S. Deora*, et al.*, "Energy control paradigm for compliance-free reliable operation of RRAM," in *Reliability Physics Symposium, 2014 IEEE International*, 2014, pp. MY.10.1-MY.10.4.

[153] A. Fantini, G. Gorine, R. Degraeve, L. Goux, C. Y. Chen, A. Redolfi*, et al.*, "Intrinsic program instability in HfO2 RRAM and consequences on program algorithms," in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 7.5.1-7.5.4.

[154] L. G. Gao, P. Y. Chen, and S. M. Yu, "Programming Protocol Optimization for Analog Weight Tuning in Resistive Memories," *Ieee Electron Device Letters,* vol. 36, pp. 1157-1159, Nov 2015.

[155] B. Gao, H. Zhang, B. Chen, L. Liu, X. Liu, R. Han*, et al.*, "Modeling of Retention Failure Behavior in Bipolar Oxide-Based Resistive Switching Memory," *IEEE Electron Device Letters,* vol. 32, pp. 276-278, 2011.

[156] P. Lorenzi, R. Rao, and F. Irrera, "Conductive filament evolution in HfO2 resistive RAM device during constant voltage stress," *Microelectronics Reliability,* vol. 55, pp. 1446-1449, Aug 2015.

[157] P. Shrestha, D. Nminibapiel, J. P. Campbell, J. H. Kim, C. Vaz, K. P. Cheung*, et al.*, "Accurate RRAM transient currents during forming," in *VLSI Technology, Systems and Application (VLSI-TSA), Proceedings of Technical Program - 2014 International Symposium on*, 2014, pp. 1-2.

[158] K. Higuchi, T. Iwasaki, and K. Takeuchi, "Investigation of Verify-Programming Methods to Achieve 10 Million Cycles for 50nm HfO2 ReRAM," in *2012 4th IEEE International Memory Workshop*, 2012, pp. 1-4.

[159] P. Huang, B. Chen, Y. J. Wang, F. F. Zhang, L. Shen, R. Liu*, et al.*, "Analytic Model of Endurance Degradation and Its Practical Applications for Operation Scheme Optimization in Metal Oxide Based RRAM," *2013 Ieee International Electron Devices Meeting (Iedm),* 2013.

[160] A. Kawahara, K. Kawai, Y. Ikeda, Y. Katoh, R. Azuma, Y. Yoshimoto*, et al.*, "Filament scaling forming technique and level-verify-write scheme with endurance over 107 cycles in ReRAM," in *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 2013, pp. 220-221.

[161] F. Alibart, L. G. Gao, B. D. Hoskins, and D. B. Strukov, "High precision tuning of state for memristive devices by adaptable variation-tolerant algorithm," *Nanotechnology,* vol. 23, Feb 24 2012.

[162] A. Chen, "Current overshoot during set and reset operations of resistive switching memories," in *2012 IEEE International Reliability Physics Symposium (IRPS)*, 2012, pp. MY.2.1-MY.2.4.

[163] B. Chen, J. F. Kang, B. Gao, Y. X. Deng, L. F. Liu, X. Y. Liu*, et al.*, "Endurance Degradation in Metal Oxide-Based Resistive Memory Induced by Oxygen Ion Loss Effect," *IEEE Electron Device Letters,* vol. 34, pp. 1292-1294, 2013.

[164] Z. Q. Wang, S. Ambrogio, S. Balatti, S. Sills, A. Calderoni, N. Ramaswamy*, et al.*, "Cycling-induced degradation of metal-oxide resistive switching memory (RRAM)," in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 7.6.1-7.6.4.

[165] C. Yu-Sheng, L. Wen-Hsing, L. Heng-Yuan, C. Pang-Shiu, W. Sum-Min, T. Chen-Han*, et al.*, "Impact of compliance current overshoot on high resistance state, memory performance, and device yield of HfOx based resistive memory and its solution," in *VLSI Technology, Systems and Applications (VLSI-TSA), 2011 International Symposium on*, ed, 2011, pp. 1-2.

[166] X. Li, H. Wu, B. Gao, N. Deng, and H. Qian, "Short Time High-Resistance State Instability of TaOx-Based RRAM Devices," *IEEE Electron Device Letters,* vol. 38, pp. 32-35, 2017.

[167] N. Raghavan, R. Degraeve, A. Fantini, L. Goux, D. J. Wouters, G. Groeseneken*, et al.*, "Modeling the Impact of Reset Depth on Vacancy-Induced Filament Perturbations in HfO2 RRAM," *IEEE Electron Device Letters,* vol. 34, pp. 614-616, 2013.

[168] F. M. Puglisi, L. Larcher, A. Padovani, and P. Pavan, "A Complete Statistical Investigation of RTN in HfO2-Based RRAM in High Resistive State," *IEEE Transactions on Electron Devices,* vol. 62, pp. 2606-2613, 2015.

[169] N. Raghavan, R. Degraeve, L. Goux, A. Fantini, D. J. Wouters, G. Groeseneken*, et al.*, "RTN insight to filamentary instability and disturb immunity in ultra-low power switching HfOx and AlOx RRAM," in *VLSI Technology (VLSIT), 2013 Symposium on*, 2013, pp. T164-T165.

[170] D. Ielmini, "Modeling the Universal Set/Reset Characteristics of Bipolar RRAM by Field- and Temperature-Driven Filament Growth," *IEEE Transactions on Electron Devices,* vol. 58, pp. 4309-4317, 2011.

[171] B. Govoreanu, S. Clima, I. P. Radu, Y. Y. Chen, D. J. Wouters, and M. Jurczak, "Complementary Role of Field and Temperature in Triggering ON/OFF Switching Mechanisms in Hf/HfO2 Resistive RAM Cells," *IEEE Transactions on Electron Devices,* vol. 60, pp. 2471-2478, 2013.

[172] K. T. Regner and J. A. Malen, "Nondiffusive Thermal Transport Increases Temperature Rise in RRAM Filaments," *IEEE Electron Device Letters,* vol. 37, pp. 572-575, 2016.

[173] P. Lorenzi, R. Rao, T. Prifti, and F. Irrera, "Impact of the forming conditions and electrode metals on read disturb in HfO2-based RRAM," *Microelectronics Reliability,* vol. 53, pp. 1203-1207, Sep 2013.

[174] D. Veksler, G. Bersuker, B. Chakrabarti, E. Vogel, S. Deora, K. Matthews*, et al.*, "Methodology for the statistical evaluation of the effect of random telegraph noise (RTN) on RRAM characteristics," in *Electron Devices Meeting (IEDM), 2012 IEEE International*, 2012, pp. 9.6.1-9.6.4.

[175] S. S.-Y. Chui, S. M.-F. Lo, J. P. H. Charmant, A. G. Orpen, and I. D. Williams, "A Chemically Functionalizable Nanoporous Material [Cu3(TMA)2(H2O3)n," *Science,* vol. 283, pp. 1148-1150, 1999.

[176] E. Perez, A. Grossi, C. Zambelli, P. Olivo, and C. Wenger, "Impact of the Incremental Programming Algorithm on the Filament Conduction in HfO2-Based RRAM Arrays," *Ieee Journal of the Electron Devices Society,* vol. 5, pp. 64-68, Jan 2017.

VITA

David Malien Nminibapiel
Department of Electrical and Computer Engineering
Old Dominion University (ODU)
Norfolk, Virginia 23529

CONTACT

mbapiel@gmail.com

EDUCATION

Ph.D. Electrical and Computer Engineering, ODU, August 2017
M.S. Electrical and Computer Engineering, ODU, August 2012
B.S. Electrical and Computer Engineering, ODU, December 2010

AWARDS

Best Student Award in Nonvolatile Memory Symposium at Electrochemical Society (2015)
Ph.D. Researcher Award at ODU ECE Department (2015)
First Prize, Mid-Atlantic Section, American Vacuum Society Research Poster Winner (2013)
Third, Mid-Atlantic Section, American Vacuum Society Undergraduate Poster Winner (2011)

SELECT PUBLICATIONS (9 Refereed Journal Publications, and 10 Conference Proceedings)

D. M. Nminibapiel, D. Veksler, P. R. Shrestha, J. P. Campbell, J. T. Ryan, H. Baumgart and K. P. Cheung, IEEE Electron Device Lett. (EDL), **38**, 736, (2017)

Q. Smets, A. Verhulst, J.-H. Kim, J. P. Campbell, D. Nminibapiel, D. Veksler, et. al., IEEE Trans. Electron Devices, **64**, 1489, (2017)

D. M. Nminibapiel, D. Veksler, P. R. Shrestha, J-H. Kim, J. P. Campbell, J. T. Ryan, H. Baumgart and K. P. Cheung, IEEE Electron Device Lett. (EDL), **38**, 326, (2017)

J-H. Kim, P. R. Shrestha, J. P. Campbell, J. T. Ryan, D. M. Nminibapiel, J. J. Kopanski and K. P. Cheung, IEEE Trans. Electron Devices, **63**, 3851, 2016.

Z. Wang, D. Nminibapiel, P. Shrestha, J. Liu, W. Guo, P.G. Weidler, H. Baumgart, C. Wöll and E. Redel, in ChemNanoMat, **2**, 67, (2015)

P. Shrestha, D. M. Nminibapiel, J.-H. Kim, H. Baumgart, K. P. Cheung, *Electrochemical Society Transactions*, **75** (13), 81, (2016)

D. Nminibapiel, K. Zhang, et al., in ECS Journal of Solid State Science and Technology, **3** (4), 95, (2014)

K. Zhang, A. D. R. Pillai, M. Tangirala, D. Nminibapiel, K. Bollenbach, W. Cao, H. Baumgart, V. S. K. Chakravadhanula, Christian Kübel and V. Kochergin, Phys. Status Solidi (a), **211** (6), 1329, (2014)

M. Tangirala, K. Zhang, D. Nminibapiel, V. Pallem, C. Dussarrat, W. Cao, T. N. Adam, C. S. Johnson, H. E. Elsayed-Ali, H. Baumgart, ECS Journal of Solid State Science and Technology, **3** (6), N89, (2014)

P. Shrestha, D. M. Nminibapiel, J.P. Campbell, J-H. Kim, C. Vaz, K.P. Cheung and H. Baumgart, *VLSI Technology, Systems and Application (VLSI-TSA),* pp. 1-2, (2014)

P. R. Shrestha, D. M. Nminibapiel, J-H. Kim, J. P. Campbell, K. P. Cheung, S. Deora, G. Bersuker and H. Baumgart, *IEEE International Reliability Physics Symposium (IRPS)*, pp. MY. 10.1 – MY. 10.4, 2013.

K. Zhang, M. Tangirala, D. Nminibapiel, W. Cao, V. Pallem, C. Dussarrat, H. Baumgart, *Electrochemical Society Transactions*, pp. 175 – 182, (2013)

D. M. Nminibapiel has authored 4 oral and 5 poster presentations at national and international conferences