# Multilayer Thin Film Oxides for Resistive Switching

by

Zheng Jie Tan

Submitted to the Department of Materials Science and Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019

© Massachusetts Institute of Technology 2019. All rights reserved.

Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Department of Materials Science and Engineering
May 17, 2019

Certified by. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Nicholas Fang
Professor
Thesis Supervisor

Certified by. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Juejun Hu
Associate Professor
Co-advisor

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Donald R. Sadoway
Chairman, Departmental Committee on Graduate Studies

# Multilayer Thin Film Oxides for Resistive Switching

by

Zheng Jie Tan

Submitted to the Department of Materials Science and Engineering
on May 17, 2019, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

## Abstract

Resistive switching devices are hotly being pursued for use as the fundamental units in next-generation hardware deep-learning or neuromorphic systems. However, these devices are still tricky both to fabricate and to operate with consistency.

We present strategies which guarantees that switching devices are functional post-fabrication, and with switching cycles that are consistent both from cycle-to-cycle and device-to-device. The resistance of all observed high and low resistance states (HRS/LRS) spanned just 0.23 and 0.19 decades on the logarithm scale across all devices, with both states spanning 0.05 within single devices and all SET transitions falling within a 0.3V span in our multilayer FIB-processed device. DFT simulations suggest that Au atoms from the top metal electrode implanted deeper in the device by FIB would serve as bridging atoms for oxygen vacancies filament by promoting the formation of these vacancies. In addition, multilayer thin oxide films reduce the stochasticity of filament formation and further improves the switching consistency.

This strategy for high consistency resistive switching devices was subsequently exploited for a few purposes. Firstly, multi-bit switching was demonstrated to yield approximately 7 distinguishable states with a single blind set, i.e, without having to program current compliances or use iterative schemes. Secondly, further insights into the SET and RESET mechanisms using pulsed measurements could be obtained since switching stochasticity no longer obscures subtle trends in experimental data.

The implications of this study goes beyond the demonstration of a single high consistency device. Future understandings in resistive switching devices shall be achieved more easily since causality between processing parameters and device behaviors can now be quickly established under the significant reduction in switching stochasticity. The new degrees of freedom introduced here in designing resistive switching devices will also hasten the search for an optimal device, bringing forward the realization of large scale resistive RAM arrays for machine learning or hardware neuromorphic computing applications towards a nearer future.

Thesis Supervisor: Nicholas Fang
Title: Professor

# Acknowledgments

This thesis work is a culmination of many aspects of my time spent at MIT and I am very fortunate to have crossed paths with many individuals whom have helped to make this work possible.

First and foremost, I sincerely thank Prof. Nicholas Fang for his patience and guidance, while also providing plentiful resources and a lot of space for scientific exploration. This has provided an optimal environment with which I was able to work and learn without barriers.

I thank my group seniors Dafei Jin, Sang Hoon Nam, Qing Hu and Eunnie Lee. They have been very patient at the early time when I was not, and taught me not just technical knowledge, but also values which saw me through the rest of my PhD. I also appreciate groupmates Chu Ma and Xinhao Li just for being around, though of course they were more than that.

I am fortunate to be able to work with Prof. Bildge Yildiz and Vrindaa Somjit. Our collaboration had been productive and also much of a pleasure. I also thank my thesis committee members, Prof. Yang Shao-Horn, Prof. Rafael Jaramillo and Prof. Juejun Hu for being very encouraging throughout this journey.

Almost all of the experimental work presented here can be traced back to some technician or instructor who have taught me everything I wished to learn. Kurt Broderick and Michael (Mike) Tarkanian both deserve special mention for patiently training me in so many tools in microelectronics fabrications and mechanical engineering starting from absolute zero. I am also fortunate to be able to learn from or work with Ken Stone, Hayami Arakawa, Coby Unger at the MIT Hobby Shop, Donal Jamieson, Eric Lim, Bob Bicchieri at the Microsystems Technology Laboratories (MTL), Steve Kooi, William (Bill) DiNatale at the Institute for Soldier Nanotechnologies (ISN), Shiahn Chen at the Electron Microscopy lab of MRL, and James Hunter at DMSE. I appreciate their mentorship and accommodation to all my requests or questions and making research work actually fun.

I also thank many others, including Jingzhi An, Nicha Apichitsopa, Nanxi Li, Chenyang Yuan, Liu Yun, Alvin Tan, and Vincent Tjeng, each an expert in their respective fields, for technical discussions from different perspective related to this project, and other friends with whom I had have many interesting discussions on other things while at MIT. Yi-Min Lin was very helpful in navigating through the many DMSE requirements, especially near the end.

I thank A*STAR and Singapore taxpayers for the financial support that has allowed me to plan far without short term economic worries. I appreciate Philip Yeo's trust in scholarship candidates to remember his advice for later when we had been too young to understand his words.

Finally, I thank my family for providing enduring support of an unconditional nature.

# Contents

**1 Introduction** **11**

1.1 Motivation of work . . . 11

1.2 Current understanding and literature review . . . 14

1.2.1 Classification of mechanism . . . 18

1.2.2 Direct imaging attempts . . . 19

1.2.3 Limitations in inconsistencies . . . 20

1.2.4 Modeling attempts . . . 22

1.2.5 Proposed work . . . 23

**2 Quantification of consistency** **25**

2.1 Introduction . . . 25

2.2 Method . . . 27

2.2.1 Evaluating the log coefficient of variation, $C_{lv}$ . . . 27

2.2.2 Invariant qualities of $C_{lv}$ . . . 29

2.3 Summary . . . 33

**3 Design of semi-automated probe station** **35**

3.1 Guiding principles for design of characterization station . . . 36

3.2 Modular description of probe station . . . 38

3.3 Summary . . . 42

**4 Strategies for improved consistency in resistive switching** **45**

4.1 Method . . . 46

4.2 Experimental observations . . . 47

4.3 Interpretation of results . . . 50

4.3.1 Multibit operation . . . . . . . . . . . . . . . . . . . . . . . . 52
4.4 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 58

**5 Further investigations** **59**
5.1 Heating and aging measurements . . . . . . . . . . . . . . . . . . . 60
5.2 Bipolar nature of triple-layer oxide devices . . . . . . . . . . . . . . 61
5.3 Transient as-fabrication features . . . . . . . . . . . . . . . . . . . 63
5.4 SRIM ion-implantation simulation . . . . . . . . . . . . . . . . . . 65
5.5 Network-level simulation of filaments . . . . . . . . . . . . . . . . . 69

**6 Pulsed measurements** **75**
6.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 75
6.2 Experimental method . . . . . . . . . . . . . . . . . . . . . . . . . 76
6.3 Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 79
6.3.1 Pulsed operation behavior . . . . . . . . . . . . . . . . . . . . 79
6.3.2 Device endurance . . . . . . . . . . . . . . . . . . . . . . . . . 84
6.4 Further discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . 86
6.5 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 87

**7 Future directions** **89**
7.1 Analysis of second order effects in I-V curves . . . . . . . . . . . . . 89
7.2 Atomic probe tomography . . . . . . . . . . . . . . . . . . . . . . . 91
7.3 New degrees of freedom in stack design . . . . . . . . . . . . . . . . 92
7.4 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 94

**A Microfabrication methods** **97**
A.1 ALD deposition of WN/Al2O3 stack . . . . . . . . . . . . . . . . . . 97
A.2 Au deposition and FIB ion milling . . . . . . . . . . . . . . . . . . . 97
A.3 Reactively sputtered deposition . . . . . . . . . . . . . . . . . . . . 98

**B DFT method** **99**

**C Standard electrical test set** **101**

# List of Figures


1-1 Bare layout of simplest RS device . . . . . . . . . . . . . . . . . . . . . 15
1-2 Definition of unipolar and bipolar operation . . . . . . . . . . . . . . . 16
1-3 Reproduced HRTEM image of RS filament. . . . . . . . . . . . . . . . . 20

2-1 Description and simulation of CDFs . . . . . . . . . . . . . . . . . . . . 28
2-2 Invariance of lg CDF with device I or R . . . . . . . . . . . . . . . . . . 30
2-3 Invariance of $C_{lv}$ on non-linear device . . . . . . . . . . . . . . . . . . . 31

3-1 Modular breakdown of probe station . . . . . . . . . . . . . . . . . . . 39
3-2 Probing detail of custom-built probe station . . . . . . . . . . . . . . . 40
3-3 Physical probing assembly . . . . . . . . . . . . . . . . . . . . . . . . . 41

4-1 Device material stack for 1,2,3-layer of oxides . . . . . . . . . . . . . . 47
4-2 Switching of single and multilayer devices . . . . . . . . . . . . . . . . . 48
4-3 I-V of 15 devices of 1-layer oxide . . . . . . . . . . . . . . . . . . . . . 49
4-4 I-V of 15 devices of 3-layer oxide . . . . . . . . . . . . . . . . . . . . . 50
4-5 Schematic for probing of devices . . . . . . . . . . . . . . . . . . . . . . 51
4-6 Switching with different top metal contact . . . . . . . . . . . . . . . . 53
4-7 Multibit operation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55
4-8 Distinguishable states count for multibit device . . . . . . . . . . . . . 55
4-9 Variation with FIB milling length . . . . . . . . . . . . . . . . . . . . . 56

5-1 Effects of temperature and time aging . . . . . . . . . . . . . . . . . . . 61
5-2 Bipolar behavior in single layer structures . . . . . . . . . . . . . . . . 62
5-3 Bipolar behavior in multilayer structures . . . . . . . . . . . . . . . . . 63

5-4 Transient effects from as-fabricated device . . . . . . . . . . . . . . . . 65
5-5 SRIM schematic and ion implantation depth . . . . . . . . . . . . . . . 66
5-6 Effects of FIB scan direction on device properties . . . . . . . . . . . 68
5-7 Network-level simulation schematic . . . . . . . . . . . . . . . . . . . . 70
5-8 Network-level simulation findings . . . . . . . . . . . . . . . . . . . . . 72

6-1 Schematic for pulsed mode measurement . . . . . . . . . . . . . . . . . 78
6-2 Pulsed measurement width and voltage successes on different devices 80
6-3 Endurance results from pulsed measurements . . . . . . . . . . . . . . . 85

7-1 Changes to device with heavy cycling . . . . . . . . . . . . . . . . . . . 90
7-2 Changes to device with heavy dosage . . . . . . . . . . . . . . . . . . . 91

# Chapter 1

# Introduction

## 1.1 Motivation of work

The field of resistive switching or memristors was popularized in 2008 by two simultaneous proclamation in Nature[1, 2] that the fourth fundamental circuit element, predicted earlier in 1971[3] and as the symmetric complement to the resistor, inductor and capacitor, has been found. This circuit element had been previously named the "memristor", and the 2008 claim was that some devices displaying non-volatile resistive switching due to nanoionic migration effects fit the description of the 1971 prediction. This claim soon became untenable in face of opponents who published ferocious rebuttal[4, 5] arguing that such resistive switching devices do not qualify as the ideal memristive system because a memristor has to demonstrate memory from magnetic effects and memory from analogue mechanisms are not acceptable. It is now generally agreed, or perhaps just to avoid argument, that the ideal memristor has actually not been found. The publicity in the last decade on the topic of memristors has been of a mixed bag due to the debacle at the start, but the field of non-volatile resistive switching has held water and even flourished under all the attention indirectly focused on it. Nowadays, such resistive switching devices have been incorporated into large memory arrays and touted for use as next-generation electronics enabling compute function and memory storage on the same unit. Such units will allow for faster and more energy efficient training of deep neural networks which in turn powers

a wide variety of algorithms driving our information rich society in modern times.

There has been an explosion of growth in machine learning in recent years. The total annual external investment in artificial intelligence is estimated to be between $8 to $12billion in 2016, with machine learning accounting for 60% of this investment[6]. The growth in this field is being fueled by the availability of large data sets and easily accessible computational resources. Many services are slowly being migrated away from rule-based systems to A.I.-powered systems, such as voice and image recognition and language translation. Even customer chat service and job interviews are using A.I. systems, for good or for bad. Nonetheless, the increased proliferation of A.I. in human society is largely positive especially in increasingly burdened areas such as health care where doctors and specialists are simply too short in numbers to keep up with a growing population.

There exists many models for machine learning but most of them currently require huge computational resources to be properly trained before they produce usable results. Such training efforts are not energy efficient and not as fast as they could be because of limitations in computer architecture nowadays. This stems from the modularization of the computer into separate units to provide computational and memory functions. This modularization has served the computing community well for more than 8 decades but is now unable to keep up with the esoteric requirements of machine learning. This is because the fundamental operation in machine learning, take for example the training of neural networks, require the evaluation of large data sets using an array of weights and subsequently adjusting these weights based on the error of the evaluation as compared to some reference output. This is non-optimal on traditional computer architectures because it requires shifting large amount of data back and forth between the memory and the computational unit of the computer, costing time and energy in the process. Next-generation circuits are in development to provide both functions within the same array, and the best candidate of the fundamental element in this circuits are resistive switching devices, capable of taking over the function of both the traditional transistor (for computation) and flash (for memory).

There is a large body of existing literature on resistive switching devices. However,

there is no clear consensus on how an ideal RS device can be built because there are too many material systems being investigated and too little work to derive generalized principles from all the reported results. Part of the abundance of reported results is because a large variety of materials can be switched. Basically, any material which can take a soft breakdown and be recovered can loosely be called a resistive switching device. For example, besides the more commonly accepted oxides, organic materials such as gelatin[7], egg albumen[8], and chemical compounds such as eumelanin[9, 10, 11] and chitosan[12] can all be switched. To add fuel to the fire, the growth in this field has no doubt encouraged researchers to take a second look when their devices have electrically broken down and would normally have been discarded, by attempting to revive their devices using for example a reverse voltage. Even device such as tantalum capacitors are known to heal with a reverse voltage after a soft breakdown[13]. The inclusion of all these different switchable systems under the huge umbrella of resistive switching devices have made the analysis of results and extraction of general principles quite challenging.

We can make the analysis of RS systems a bit more tractable by ignoring chemical and biological systems and confining ourselves to common materials seen in micro-fabricated on-chip devices as most other researchers in this field do. However, meta-analysis of all the reported results is still a deep challenge. The simplest RS device will need to have a bottom electrode, the switchable layer, and a top layer. The identity of the two electrodes have been found to affect whether the RS device operates in a unipolar or bipolar mode and so are critical parameters. The identity of the switchable layer is of course important, but crucially also depends on factors which are difficult to control, such as impurity contents , deposition method or crystalline structure. These properties are usually tedious to quantify and are almost always not reported alongside the RS electrical measurements. This means that any dependence on the RS character on such properties are lost from literature. Variation in RS behaviors from papers to papers, even using the same material system, thus cannot be properly explained in the absence of complete information.

The limitations in reproducibility of RS devices have hindered attempts to improve

two of their particularly problematic aspects. One aspect is in the low device yield. The current electroforming process uses a large voltage to activate as-fabricated devices for switching, but this also has a chance of causing the device to suffer a permanent breakdown. This is however not too critical a matter because a controller circuit of a RS array can easily detect and ignore failed RS devices. The second aspect is the inconsistency in switching and is harder to work around. There is currently too much variation in switching, both from cycle-to-cycle and device-to-device. This is problematic because an overlapping range of resistance values for the low or high resistance state (LRS/HRS) means that some devices will have an ambiguous state, implying that the data stored as the state of the device is lost. RS devices will need to have predictable behaviors across wafers, devices and cycles in order for them to be usable by circuit designers.

Some groups propose to tackle the inconsistencies in switching by calling for tighter deposition requirements[14]. This probably works, but will disqualify many research groups who are limited in budget and do not have access to dedicated fabrication systems. The long-term goal should be to identify all the critical parameters which affect the switching behavior of an RS device so that future deposition requirements can focus on controlling these parameters without expending effort in controlling other non-critical parameters. We have made used of this principle to considerably simplify some of our deposition steps.

## 1.2 Current understanding and literature review

The popularization of memristors in 2008 and the search for next generation hardware neuromorphic circuits has spawned a plethora of research in the field of resistive switching. In this chapter, we review the trends and methods currently used by literature.

The simplest resistive device consists of 3 basic elements(see Fig. 1-1). These are the top electrode, the bottom electrode, and the switchable material, typically an oxide. The fundamental characteristics of resistive devices, as evident in its name, is

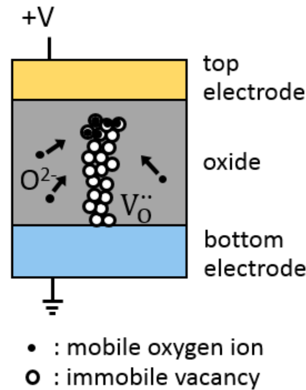



Figure 1-1: The simplest resistive switching device is made up of 3 parts, which are the top electrode, the switchable oxide and the bottom electrode. Switching between a high and low resistance state is due to the formation and destruction (dissipation) of conductive filaments in the otherwise high resistance oxide. In our device, the filaments are largely formed by oxygen vacancies as pictures, but resistive switching can also be formed by metallic conductive atoms such as migrating silver or copper atoms.

in its ability to be cycled between high and low resistance states repeatedly with the application of sufficient large SET or RESET voltage. The SET voltage is defined as the voltage required to cause the device to transit from the high resistance state (HRS) to the low resistance state (LRS). The SET voltage is so named because the main mechanism in which this state change happens is because dispersed material in the oxide, which are typically oxygen vacancies or trace metal atoms, migrate under the applied voltage to form a conductive bridge. This conductive bridge causes the resistance of the device to drop, thus the transition from a HRS to a LRS. The RESET is the opposite of SET. When the RESET voltage is applied, the filament breaks, thus putting the device back into the HRS. Filament breakage is more commonly known as filament dissipation in literature and the latter term shall be used in this text. Nonetheless, other mechanisms of switching are possible. For example, there could be an interfacial switching mechanism [15, 16] where the entire interface between the oxide and electrode is responsible for the switching due to a large concentration of oxygen vacancies near the interface. Webs of filament have also been hypothesized under the scenario that there is a large concentration of oxygen vacancies [17]. However,

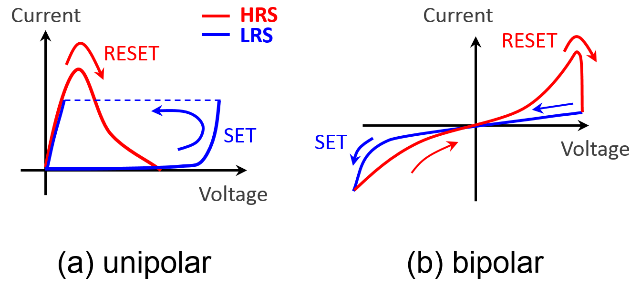



Figure 1-2: In unipolar operation, both RESET and SET transitions can occur with voltages of the same polarity. The resistive switching (RS) device will cycle between the high and low resistance states (HRS & LRS) with each voltage ramp. In bipolar operation, RESET and SET are controlled by going to a sufficiently large positive and negative voltage respectively (or vice versa).

it is currently difficult to rule out one mechanism over another due to lack of direct imaging techniques that can work at the filament level.

Both SET and RESET can be either an abrupt or a gentle transition. If SET and RESET happens with opposite voltage polarity, the device is then known as a bipolar device. If SET and RESET happens with the same voltage polarity, the device is then known as a unipolar device (see Fig. 1-2). A bipolar device can be cycled by repeatedly ramping from positive to negative voltages and vice versa. On the other hand, a unipolar device can be switched between HRS and LRS simply by ramping from 0V to either a positive or negative voltage and a state change will happen as long as the magnitude of the voltage is sufficiently large. The unipolar behavior is present mostly in systems with a symmetric stack where both the top and bottom electrode are of the same material. Thus, filament forming by the migration of materials has no preferred polarity. Literature suggests that filament breakage is due to Joule heating effect, which inherently has no preferred polarity as both negative and positive currents will have the same heating effect. Thus, for unipolar devices where there is no preferred direction of migration for filamentary elements, filament dissipation can also take place either at a positive or negative voltage. On the other hand, bipolar devices have a built-in asymmetry for filamentary migration, so RESET becomes asymmetric even though its underlying mechanism, like Joule heating, need not be. In other words, RESET is asymmetric in a bipolar device because any net

behavior will be a combination of the Joule-heating induced filamentary dissipation superimposed over an asymmetric filamentary migration behavior. Unipolar devices are typically not favored for actual use because it is trickier to control when the voltage polarity is not available as a degree for manipulation. The on-off ratio refers to the ratio of resistances between the HRS and LRS and can span from anything from 1.5 to $10^3$ across different material systems.

In literature, RS devices are typically characterized by I-V measurements using slowly varying DC voltages or so-called voltage ramps. The other type of electrical measurements are measurements in a pulsed mode. The main purpose of pulsed mode measurements are to obtain performance metric for operating speed and to demonstrate multibit operation. Multibit operation refers to the capability of setting an RS device not just to LRS and HRS, but to different intermediate states between the two extreme resistances. An RS device with a 3-bit capacity means that is can be set to a total of $2^3 = 8$ distinguishable states. Devices with large variability in the LRS or HRS resistances are unsuitable for multi-bit operation because it will be difficult to differentiate an intermediate resistance state when each state have a large resistance uncertainty. These intermediate states are typically obtained by applying voltage pulses to incremently change the conductance of a filament, or by inducing SET with an external current compliance programmed on the benchtop measurement device. When an external device enforces a current compliance, the growth of filament formed in the RS device will be limited to a certain diameter and thus have a corresponding resistance associated with it. This is a commonly used strategy in demonstrating multibit operation in RS devices.

Multibit operation has two major applications down the road. When RS devices are arranged into large memory arrays solely to be used for storage functions, multibit operation serves to increase the storage capacity per unit area in the footprint of the memory array. This is important because the primary performance indicator for memory cells is its storage capacity while operating speed is secondary. Conversely, the primary performance indicator for transistors is its operating speed while power consumption, though important, is still secondary. These rules would of course not

hold for special use cases, such as for memory cards requiring fast transfer speeds such as when recording HD videos, or for transistors in micro-controller units to be used in long life battery-operated devices, but these indicators are the primary criteria with which a semiconductor foundry will evaluate its manufactured products. The second major application of multibit operation is when RS devices are arranged in cross-bar arrays to be used for hardware machine learning. Such an architecture is the hardware realization of neural networks commonly used in machine learning libraries nowadays. The principle of transferring such circuits from software to hardware is so that the neural network can be realized on a smaller overall footprint. This comes as a result of the possibility of using a single device as a hardware-implemented neural network node to replace the footprint spanned by a software-implemented neural network which has to be emulated with a large cluster of transistors. Multibit operation would benefit hardware neural networks because each device needs to have weights associated with it, and a more accurate neural network will benefit from having more than 2 discrete weights at each device.

### 1.2.1 Classification of mechanism

If conduction occurs through the filament by direct ohmic conduction through metal atoms, the device can then be classified as a conductive-bridge random access memory (CBRAM). If conduction occurs by Poole-Frenkel hopping through oxygen vacancies defect states, then the device can be classified as a resistive random random access memory (RRAM or ReRAM). Resistive switching (RS) devices refers to the superset of devices with different conduction mechanism and is the term which we shall use to refer to our device. The term "memristor" is not used to avoid controversies on how model memristors oughts to be memory units switching using electromagnetic principles and not through some other physical constructs, such as through forming and dissipation of physical filaments. We will later show in our experimental results that our device can be better classified as a RRAM device. CBRAM devices are typically form with Cu or Ag filaments since these two elements are somehow more mobile and these metal atoms would migrate more easily under an applied potential. CBRAM devices

are however time-dependent because the formed filaments are unstable and dissipate by itself. This is unacceptable for memory storage applications, but has found use for emulating time-dependent functions such as the spike-timing-dependent plasticity effect observed in biological neuronal circuities.

The earliest reported resistive switching character was in 1967 by Lamb[18] who reported that it was possible to cycle SiO2 films between high and low resistance states. In that paper, he also managed to identify that the filamentary are not metallic in nature but have some activation energy for conduction. This was supported by determining the I-V dependence of the low resistance state to be non-linear and by showing that the temperature of coefficient of this resistance is negative. These checks have remained as the method of choice in determining the class of conduction of RS filaments, to differentiate between metallic and non-metallic forms of conduction.

### 1.2.2 Direct imaging attempts

Since then, and especially from 2008 onwards[1, 2], there has an explosion in the number of reported systems displaying RS effects. In spite of all the advanced analytical tools now available to typical research groups, efforts in the characterization these devices have only progressed at a moderate pace due to the dimensions of these filaments and stochastic nature for the formation of these filaments. Some groups have undertaken heroic efforts and successfully used HRTEM imaging to directly observe these hypothesized filaments (see reproduced TEM in Fig. 1-3[19]). Their results have not only provided a strong experimental evidence for the core hypothesis of filament formation for other groups to build upon, but also conveys the immense challenge in such direct imaging experiments. Kwon and coworkers deposited a Pt-TiO2-Pt stack and triggered a SET in their unipolar RS device with a relatively large current compliance of 20mA. This causes part of the top electrode to be blown off as an indication of the position where the formed filament is possibly the strongest/thickest. A lamella was then extracted from this region for TEM imaging. With the correct orientation in the lamella good for observing the Magneli phase of TiO2, a bridging Magneli filament can be identified, with this Magneli phase filament

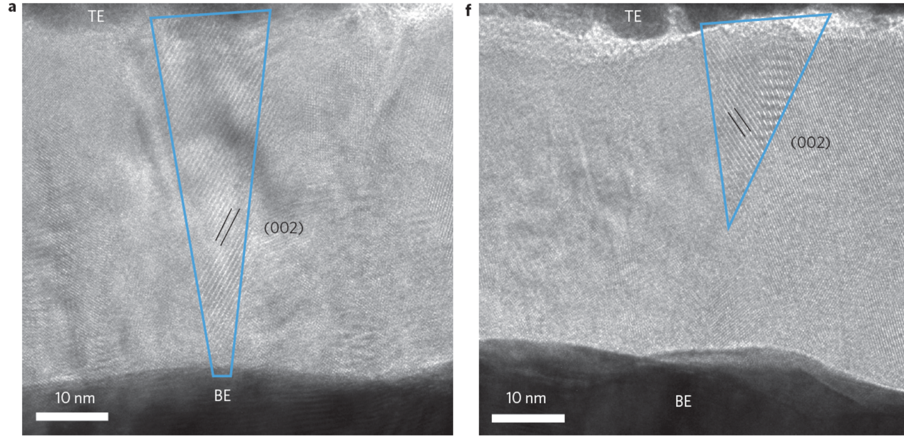



Figure 1-3: The direct image of the resistive switching filament is hard even under ideal circumstances. In (a), the filament can be electrically measured in-site be in the conductive low-resistance state (LRS). In (b), the filament is in the HRS. The two TiO2 structures only differ by crystallographic phase , where conduction is mediated in (a) with a $Ti_nO_{2n-1}$ Magneli phase confirmed through a separate electron beam diffraction analysis not shown here. Figures reproduced from Kwon's[19] Fig. 2a and 2f with permissions.

disappearing after an in-situ RESET transition. The dimension of this filament spanned approximately 10nm in width and differed only from neighboring material by a different crystallographic structure confirmed by a TEM diffraction pattern. The difficulty in imaging such filaments is also raised by the low probability of capturing a filament during lamella preparation. The effort to search for these nanoscopic filament has been commendable, but such technical capabilities are beyond the reach of most research groups to replicate anyhow. Thus, there has only been a handful of replicate TEM studies so far[20, 21, 22, 23] but they are so far all in agreement with Kwon et. al.'s findings.

### 1.2.3 Limitations in inconsistencies

When resistive switching devices are used for hardware machine learning applications, such as in crossbar arrays to implement edges of a deep neural network, the operational characteristics of the resistive switching devices become important in determining the limits of the training. Novice implementations of deep software neural networks using open-source libraries (Tensorflow, Pytorch, etc...) with a few lines of code can

easily reach test accuracies of 98% on the MNIST database[24], whereas state-of-the-art software-implemented neural networks has climbed in test accuracies to a mind-boggling 99.8% in 2015. However, simulations of hardware neural networks implemented by RS devices have found that test accuracies will be significantly limited when RS devices have lackluster switching performance[25]. Gokmen (2016)[25] simulated device-to-device switching inconsistencies and found that a 300% statistical variation in resistance step changes in response to some pulse will drop test accuracies from 99.8% to a limit of 90%. In typical resistive switching devices, the resistance ranges of the HRS and LRS can span several orders of magnitude (1000% statistical variation per order of magnitude), but Gokmen found that a variation of 1 order of magnitude would be sufficient to drop the test accuracy to 0%.

On the other hand, such simulations of hardware neural network are helpful in telling us which targets would be less important to pursue. Gokmen (2016)[25] shows that an on/off ratio of 10 to 50 is sufficient to minimize noise and not limit test accuracy. This stands in contrast to reports of devices which boasts on/off ratios of $10^4$ to $10^6$ by saying that such large values are not necessary. The knowledge of the relative significance of each parameter in determining the limits of hardware implemented neural network is important because there is often a trade-off when trying to optimize each character of an RS device, so insignificant parameters should not be optimized if doing so will jeopardize other significant aspects of the device.

Such simulations are also useful in studying mitigation strategies for what are commonly perceived to be limitations in RS devices. For example, Burr (2015)[26] find that asymmetries in the switching character from responses to positive versus negative voltage pulses can be mitigated by simply operating the neural network normally with one polarity and using the other voltage polarity occasionally to reset the network. Gokmen (2016)[25] further states that these response asymmetries are not an issue because the control circuit for such hardware-implemented neural networks can have compensating circuits built in to correct for this asymmetry, as long as such asymmetries are common to each device so that one compensating circuit can serve the entire array of RS devices instead of requiring one compensating circuit for each

device. Similar analysis made in the same paper further stresses the importance of device-to-device consistency for RS devices, where inconsistencies would lead to either test accuracy limitations or larger effective footprints.

### 1.2.4 Modeling attempts

Another aspect of development in the research of RS devices is in the modeling of such devices. However, this still a developing field due to the lack of breadth to the data available. As mention, directly imaging of RS filaments is difficult to conduct. In-situ observations of switching on open-face transverse RS device have also been reported, but there are criticisms in whether such observations are good proxies for how switching actually occurs in a bulk oxide in typical device because migration behaviors would be vastly different at open surfaces than in the bulk.

There are three standing models for the microscopic origin for the migration of filamentary material[27]. The first model is movement of oxygen vacancies due to the electromigration effect. This is simply due to the interaction of the positive charges with the electric field generated by an applied voltage. This model can also be coupled with Joule heating effects where local temperature rises can provide additional kinetic energy for oxygen vacancies to more successfully hop over to an adjacent site under the electromigration model. The second model is by the Soret force which develops as a result of Joule heating. In this model, local temperature rises creates a temperature gradient $\nabla T$ where an oxygen ion migrates down the temperature gradient, conversely implying that oxygen vacancies will migrate towards at local temperature hotspots. The three model is by Fick's diffusion, which can be accelerated by Joule heating. Fick's diffusion trivially describes the migration of oxygen vacancies away from high concentration regions, and can be easily applied to heated filaments to explain the RESET process where constituent materials in the filament break up to exhibit the LRS to HRS transition.

There is flexibility in mixing, interchanging and tweaking the mechanisms described above order to fit to experimental results and a variety of composite models have been proposed by many papers. However, there is not yet clear methods to differentiate

one model from another because this requires nanoscopic probing arrangement to detect temperature, electric fields, or ionic content, and is generally a challenging effort. Thus, the current bottleneck in understanding is not in the generation of more theoretical models, but in the designing of experimental procedures to expand the multitude of tools for differentiating one mechanism from another.

### 1.2.5 Proposed work

We contribute to this arsenal of tools by demonstrating a highly consistent RS stack whose behavior can be modified by incremental changes in order to correlate specific process parameters with changes in the observable switching character. This is an improvement over existing devices which may have excellent switching character but whose characters are highly sensitive with small changes in the device make-up or even just going from device-to-device. The ability to correlate individual process parameters with specific RS characters will enable the RS mechanism to be broken down into smaller steps to be made sense of and can help to break through the impediment in this field that each reported device have its own unique set of switching character and can hardly to compared to another reported device. In a future effort to increase the ease at which different devices reported from different groups can be compared, we will also present a figure of merit which can quantify switching consistency using a single variable. This will collapse 2D plots used to illustrate the consistency of each measured device into a single variable can then be consolidated and plotted against other process parameters in the hope of identifying a link between them. Such an identified link will then provide guidance for future design of RS devices with high switching consistency.

# Chapter 2

# Quantification of consistency

Resistive switching (RS) or RRAM devices typical have a large variability in cycle-to-cycle and device-to-device switching. It is a major aspect of ongoing research to address this device consistency, but there has so far been a lack of proper method to quantify this performance. We introduce a quantitative measure which can be easily extracted from a CDF plot commonly used to illustrate device consistency. We show how this measure remains constant with regards to the scale of the device, the measurement voltage used, whether the device is Ohmic or non-linear, and regardless whether this measure is extracted from the CDF for the device resistances or currents. This will allow future studies on device consistencies to be better quantified, and will also enable meta-analysis of results already published in literature to be conducted with a potential systematic review.

## 2.1 Introduction

One recurring theme in research on RS devices is on the need to improve the switching consistency. Switching consistency is needed for RS devices to be part of a larger system such as a RRAM array since a control circuit will need to be programmed in advance to know how to apply a waveform to toggle the RS devices between states during a write cycle, and to know what probe current corresponds to what state during a read cycle. The switching consistency is most easily observed with I-V plots,

where the most consistent devices will have I-V sweeps from multiple cycles neatly superposed on one another, whereas erratic devices have will highly variable transition voltages and resistance states.

However, there are some limitations in relying solely on I-V plots for determining the consistency of devices. Firstly, it is hard to obtain a quantifiable metric from the scatter of sweeps in I-V plots. The largely qualitative nature of this metric means that it will be hard to apply statistical methods to find correlations in this consistency with device fabrication parameters. This will be a handicap on the use of big data analysis in this field. Secondly, I-V plots between different devices, material stacks and groups will be hard to compare with one another. A perfectly consistent RS device display different characters, such as being unipolar or bipolar, have transitions that can span from a gentle one to abrupt. These second-order descriptors come above the primary descriptor of resistance values and transition voltages, but are present in I-V plots and would complicate the analysis of consistency. The solution to obtain an analyzable quantity to evaluate device consistency is to extract properties of the RS device at small test voltages for each sweep cycle for further analysis. This has been done in most papers where the cumulative probability or cumulative distribution function (CDF) plots[28, 29, 30, 31, 32, 33] are shown for properties such as device current, resistance, or transition voltages. Some papers [34, 35, 36, 37, 38] have brought this a step further to extract a single numeric value for quantify device consistency by using the coefficient of variation,

$$C_v = \frac{\sigma}{\mu}$$

, where $\sigma$ is the standard deviation and $\mu$ is the mean, evaluated on the spread of device current, resistance, or transition voltages. The coefficient of variation is a well-known statistical measure, but its use has not yet been carefully evaluated in the context of RS devices. To the best of our knowledge, there are no other quantitative measures of consistencies used by literature on RS devices. Here, we show a shortcoming in $C_v$, but also show how it can be extended to what we call the logarithmic coefficient of variation, $C_{lv}$. The new statistical measure presents several advantages for universal

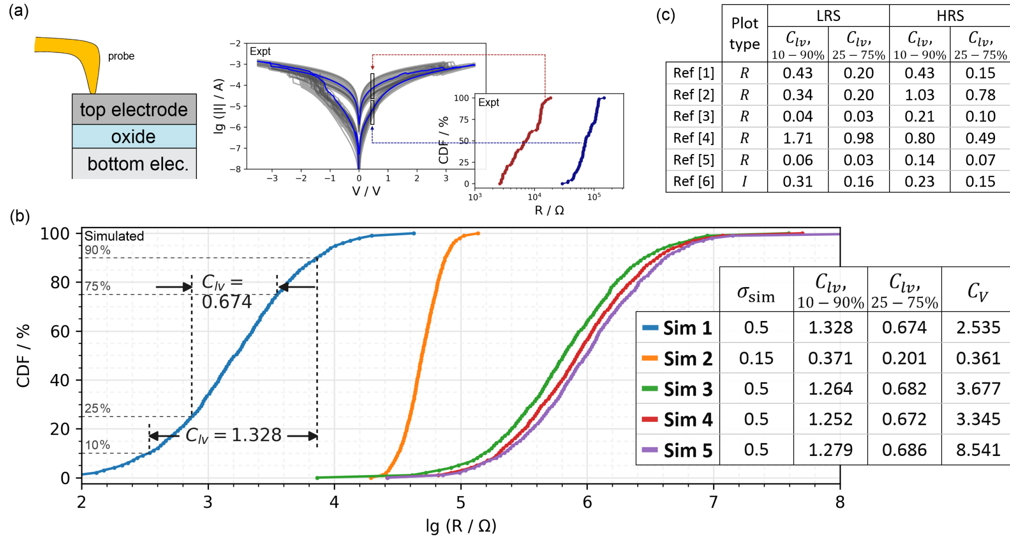
adoption as a figure of merit for consistency. We will show how this metric is invariant on the range of measurement or fabrication parameters. Specifically, we will show how this metric remains constant when evaluated on the various types of published CDFs, regardless whether device current or resistance is plotted, and showing insensitivity to device area scaling, polling voltage, etc. Such a quantity would thus be ideal for evaluating the consistency of RS devices.

## 2.2 Method

### 2.2.1 Evaluating the log coefficient of variation, $C_{lv}$

We will first describe the steps of extracting the log coefficient of variation, $C_{lv}$, for a RS device and leave the analysis of the invariant nature of this metric to the next section.

The first step to evaluate $C_{lv}$ is to collect values across RS sweep cycles and/or devices. These values can be the device resistances or currents near some small polling voltage. The magnitude of the test voltage should be sufficiently small so as not to trigger SET or RESET, but large enough to provide a measurable current. Frequent values seen in literature vary from 100mV to 1V. The logarithm of the device resistances (or currents) is then tabulated and plotted on a cumulative distribution function (CDF) plot (see Fig. 2-1). We define $C_{lv}$ to be the variability of these logarithmic values, extracted about the median value at some predefined interval, for example, the interquartile range of 25%-75%, the wider interdecile range of 10%-90%, or some other symmetric range about the median as preferred. Strictly speaking, this variability is not the usual standard deviation (given by $\sigma = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}}$) because it is evaluated about the median rather than the mean and is more accurately known as the median absolute deviation. Loosely defined, $C_{lv}$ can also be defined as the spread in the logarithmic values of device resistances (or currents). In literature, the coefficient of variation, $C_v = \sigma/\mu$, is typically used to quantify consistency as it is a conventional statistical measure but this quantity is inadequate in describing




| | Plot type | LRS $C_{lv}$, 10 – 90% | LRS $C_{lv}$, 25 – 75% | HRS $C_{lv}$, 10 – 90% | HRS $C_{lv}$, 25 – 75% |
|---|---|---|---|---|---|
| Ref [1] | $R$ | 0.43 | 0.20 | 0.43 | 0.15 |
| Ref [2] | $R$ | 0.34 | 0.20 | 1.03 | 0.78 |
| Ref [3] | $R$ | 0.04 | 0.03 | 0.21 | 0.10 |
| Ref [4] | $R$ | 1.71 | 0.98 | 0.80 | 0.49 |
| Ref [5] | $R$ | 0.06 | 0.03 | 0.14 | 0.07 |
| Ref [6] | $I$ | 0.31 | 0.16 | 0.23 | 0.15 |

| | $\sigma_{\text{sim}}$ | $C_{lv}$, 10 – 90% | $C_{lv}$, 25 – 75% | $C_V$ |
|---|---|---|---|---|
| Sim 1 | 0.5 | 1.328 | 0.674 | 2.535 |
| Sim 2 | 0.15 | 0.371 | 0.201 | 0.361 |
| Sim 3 | 0.5 | 1.264 | 0.682 | 3.677 |
| Sim 4 | 0.5 | 1.252 | 0.672 | 3.345 |
| Sim 5 | 0.5 | 1.279 | 0.686 | 8.541 |

Figure 2-1: (a) Typical schematic for measurement on a resistive switching device, showing the device stack, repeated cycles of I-V and the cumulative distributions for the two resistance states. (b) Simulated random drawings for device resistances were carried out to demonstrate how $C_{lv}$ varies (ignoring HRS or LRS assignment). For each of 5 sets of drawings, the log of these values were drawn from a normal distribution with a standard deviation of $\sigma_{\text{sim}}$ and an arbitrary mean. The cumulative probability of the drawings are shown. The annotations on first blue plot shows how our metric for consistency $C_{lv}$ can be easily extracted. This metric can be evaluated for the 25-75% or 10-90% interval as preferred. Sim 1,3,4,5 were drawn using the same $\sigma_{\text{sim}}$ value and subsequently have similar $C_{lv}$. On the other hand, the commonly used coefficient of variation, $C_v$, unreliably varies with minor changes in the random drawings. (c) Metric can be applied on literature data, also suitable for use regardless whether the CDF for resistance or current is plotted.

highly variable RS devices whose resistance states can often vary over multiple orders of magnitude. When large variability exists, $C_v$ is not a good measure because it is normalized by the highly variable mean value $\mu$ of the device property. $C_v$ thus requires a large sample population to converge and will be much less reliable as a metric than the $C_{lv}$ we propose. A simple simulation was carried out to demonstrate the variable nature of $C_v$ (see table inset of Fig. 2-1). On the other hand, $C_v$ remains suitable to quantify parameters which does not vary over orders of magnitude such as the onset voltage, but the simpler standard deviation $\sigma$ would suffice there too.

As intense research into RS devices continue, their switching consistencies will improve across the board, while this in turn causes the criteria for declaring devices to

be “consistent" to become stricter over time. $C_{lv}$ can adapt to the evolving demands by expanding the interval used, say from 25-75%, to 10-90% or perhaps to a strict 0.1-99.9% at some distant future when the consensus becomes that no more than 0.2% of devices shall have outlier performances. The use of a strict 0.1-99.9% interval will magnify the value of $C_{lv}$ and emphasis the effects of outliers. We shall use a 10%-90% interdecile interval in the following analysis on $C_{lv}$ and in subsequent chapters.

### 2.2.2 Invariant qualities of $C_{lv}$

A metric used to quantify switching consistency should ideally reflect the variability of the device and not be dependent on other factors which vary widely in literature. For example, reported RS device can span from 10nm×10nm[39] to 0.1mm×0.1mm or beyond, so the metric should of course be invariant on device area. This metric should also not depend on the polling voltage so that it remains valid for comparison between published reports where devices are evaluated at different poll voltages as freely determined by different research groups. Also, there is no consensus in literature on whether the distribution of device currents or resistances should be plotted to illustrate variability, so an ideal metric should have the flexibility to be evaluated on either plot. We will verify that our proposed metric has all the above properties.

Device resistance at some poll voltage, $R|_{V=V_p} = V_p/I$, is determined by applying some poll voltage $V_p$ and measuring the current through the device. The resistance of a non-linear device will vary with the poll voltage, but we leave its treatment to a later paragraph. Since device resistance is inversely related to this measured current with

$$\log \left(R|_{V=V_p}\right) = \log V_p - \log I, \tag{2.1}$$

the CDF for the logarithm of device current and resistances is related by a 180° rotation (see Fig. 2-2). Importantly, the horizontal spread of the CDF plot will not change. This means that $C_{lv}$ with a symmetric interval will have the same value when evaluated on either the device resistances or currents.

From Ohm’s law (Eq. 2.1) for a simple resistive device, we can also see how the

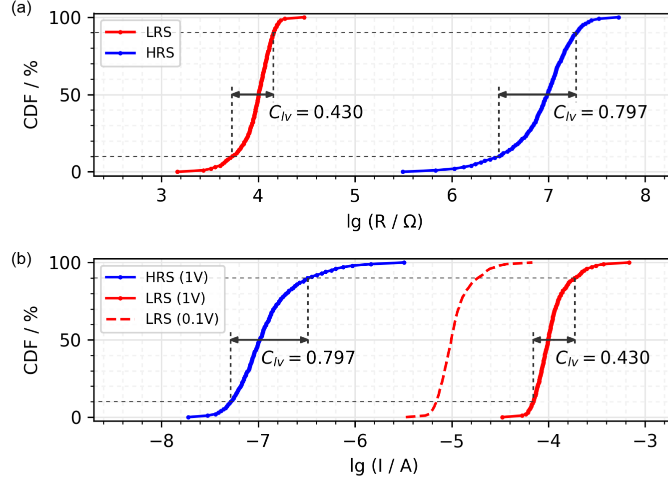



Figure 2-2: A random drawing of device parameters for the high- and low-resistance states (HRS & LRS) was made. The CDF for the logarithm of the device currents and resistances are rotational copies of each other. The red dotted line in (b) shows the effect of measuring device current using a different poll voltage. The consistency metric $C_{lv}$ remains constant throughout all these changes. The random distribution is deliberately made asymmetric with a narrower distribution of more resistive states above the median as a visual marker to help observe the 180° rotation.

polling voltage does not affect $C_{lv}$ because a change in the polling voltage from $V_{p1}$ to $V_{p2}$ corresponds to a horizontal shift of $\log\left(\frac{V_{p2}}{V_{p1}}\right)$ in the CDF for currents and no shift for the CDF for resistances, leaving no scale changes to either CDFs. RS devices in the low resistance states (LRS) usually obeys Ohm's law[40, 41, 42] because current can pass through the RS filaments either by direct conduction through metal dopants for CBRAM devices, or via defect states of oxygen vacancies chain with negligible activation energies[43].

We extend the invariance analysis of $C_{lv}$ for HRS states where devices typically have a dynamic impedance, such as that with a Poole-Frenkel response[44, 45, 46, 47],

$$I \propto E \exp\left(\frac{-q\left(\phi_B - \sqrt{qE/\pi\varepsilon}\right)}{k_B T}\right)$$

, where the applied electric field $E$ is proportional to the applied voltage, $\phi_B$ is the hopping barrier height, and $\varepsilon$ is the dielectric permittivity. This equation can be

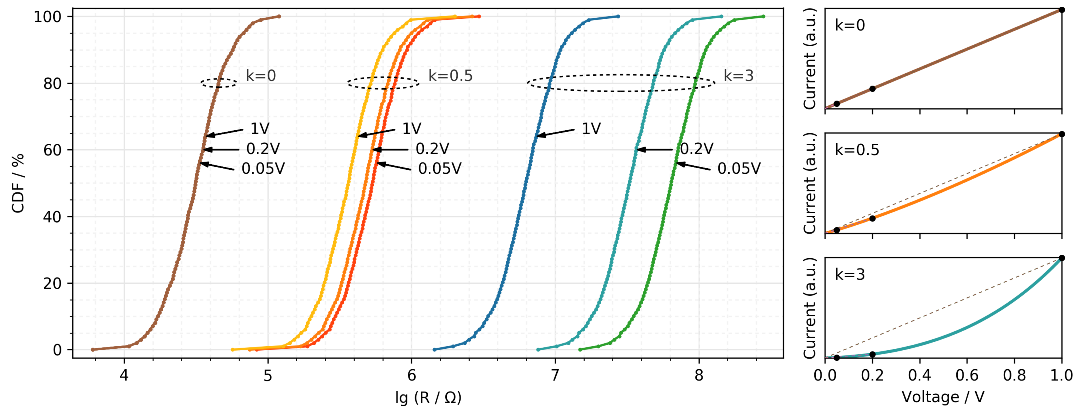



Figure 2-3: Three sets of simulations, each with a different $k$, were performed to get random distributions of Poole-Frenkel devices. For each of these distributions, the CDF of the resistances seen at 3 different voltages are shown. The variable $k$ describes non-linear behavior in devices obeying Poole-Frenkel conduction. An ohmic device has $k = 0$ and a measured resistance which does not depend on poll voltage. A larger $k$ value implies greater non-linearity, as seen by comparing (b),(c),(d). A non-linear I-V behavior means that the measured device resistance depends on the poll voltage. For example, a device with $k = 3$ in (d) will appear to have a smaller resistance when polled at 1V compared to when polled at 0.05V. However, this change in $R$ corresponds only to a horizontal translation of the CDF and so $C_{lv}$ remains invariant on the choice of poll voltage.

simplified to focus on experimentally determinable values,

$$\ln\left(\frac{I}{V}\right) = kV^{1/2} + \ln(A) \tag{2.2}$$

, where $k = (q^3/\pi\varepsilon d)^{1/2}$ is approximately a constant function of physical constants, dielectric thickness $d$ and electric permittivity $\varepsilon$, and $A$ is a catch-all parameter which varies from cycle to cycle. $A$ determines the switching consistency since it is directly dependent on the physical properties on the filaments formed, such as their combined cross-sectional area or conductance. At some poll voltage $V_p$ with a certain device current $I$, the resistance is given by

$$R|_{V=V_p} = \frac{V_p}{I} = \frac{1}{A} \cdot \exp\left(-kV^{1/2}\right) \ .$$

This means that the experimental choice to evaluate the resistance at a different poll voltage $V_p'$ will cause the determined resistance to see only a scaled change,

$$\frac{R|_{V=V_{p'}}}{R|_{V=V_p}} = \frac{\exp(kV_p^{1/2})}{\exp(kV_{p'}^{1/2})},$$

implying that ln $R_{V_p}$ only sees a shift and no scaling changes. This shows that $C_{lv}$ remains invariant under arbitrary choices of the polling voltage even for a non-linear Poole-Frenkel dependence of device currents on voltage (see Fig. 2-3). In this analysis, it is assumed that the dielectric permittivity $k$ remains constant between switching cycles. This is a plausible assumption since $k$ is a scale-invariant material constant and is unlikely to depend on the specific location or physical dimensions of filament formed during a particular switching cycle. Similar analysis of other conduction mechanisms commonly fitted to RS devices (see Table 2.1) shows that $C_{lv}$ will also remain invariant regardless of $V_p$.

| Type | Form | lg$R$ with different $V_p$ | $C_{lv}$ |
|---|---|---|---|
| Ohmic | $I = V/R$ | no change | invariant |
| space-charge limited current[48] | $J = \frac{9}{8}\mu\varepsilon_o\varepsilon_r\frac{V^2}{L^3}$ | shifts by $\ln\left(\frac{V_p}{V_{p'}}\right)$ | |
| hopping | $J = qn\lambda f \exp\left(\frac{q\lambda E - E_a}{k_B T}\right)$ | shifts (complicated form) | |
| Poole-Frenkel[49] | $I \propto E \exp\left(\frac{-q\left(\phi_B - \sqrt{qE/\pi\varepsilon}\right)}{k_B T}\right)$ | shifts by $\frac{\exp\left(kV_p^{1/2}\right)}{\exp\left(kV_{p'}^{1/2}\right)}$ | |

Table 2.1: Effects of changes in the CDF plot of lg $R$ and proposed metric $C_{lv}$ for various types of conduction mechanisms.

## 2.3 Summary

In summary, we have introduced the log coefficient of variation $C_{lv}$ as a metric to quantify the switching consistency of RS devices as an improvement over the metric used in existing literature. We showed that this metric is independent on the measurement choice of device current or resistance, on device area, on poll voltages and extending well to non-Ohmic devices. Because of these merits, $C_{lv}$ is valuable not just going forward, but will also be usable on the trove of existing CDF plots in literature results in contributing towards future review studies of RS device consistencies. This will hopefully uncover new correlations between process parameters and device performance and contribute to improvements in switching consistency in future RS devices.

# Chapter 3

# Design of semi-automated probe station

Given the importance of electrical measurements, it is paramount for any research involving RS devices to build up capabilities on performing such measurements. This is especially important in the field of RS because of the inconsistencies of devices from cycle-to-cycle and device-to-device. Other microfabricated devices may suffice from manual measurements of a few measurements taken in a single measurement session, but for RS devices, there are useful information to be learnt from the scatter of measurements as well.

This chapter describes the guiding principles used in designing a custom probe station for electrical measurements and how the various modules were put together to support semi-automatic DC or pulsed electrical measurements supporting sample translation, probe touch-down and with sample heating of up to 250°C. Given the limited time of this project and the vast amount of potential work in this field, not all the available options provided by this probe station were used to their fullest extents, but some preliminary work had been carried out for each available option (heating and pulsed measurements) and the results will be useful to guide future work.

## 3.1 Guiding principles for design of characterization station

The precision of I-V measurements increases with sampling time. This is a combination of two factors, one from the trivial improvement in the sourcemeter's analog-to-digital conversion accuracy with a larger time-averaging window, and the second is from the averaging of each measurements over each AC power line cycle at 60Hz (for US consumer supply). To characterize bipolar devices, each I-V cycle should also provide the device under test with a voltage excursion from 0V to beyond the SET voltage, then to beyond the RESET voltage going through 0V, and then back again to 0V. This sets the bare lower limit for a reasonable measurement. The very bare measurement would span from -4V to 3.5V with a minimum voltage step of 100mV and with each electrical current measurements spanning at least 5 power line cycles. This sets the lower limit of 6s for each I-V measurement, but in practice, data is collected with a higher accuracy with finer voltage steps, some finite soaking time of 100ms at each measurement point to avoid observing capacitance charging effects and using a longer time averaging period. For a fresh device, multiple I-V sweeps would be collected with incrementally increased voltage ranges in order not to accidentally cause a permanent breakdown with large voltages unexpected by a device. This step later became redundant when fabrication guaranteed the production of devices with predictable electrical characteristics but these incremental measurement programs were kept as it in because there was no good reason to use riskier measurement programs and lose potential readings when devices breakdown prematurely.

A freshly fabricated device is subjected to a series of test which starts with small probe voltage range of -0.5 to 0.5V intended to extract inital resistance states without causing any SET or RESET. The automatic program then subjects the device to incrementally wide voltage ranges to capture its SET and RESET points, repeats that, then does 20 cycles of voltage sweeps from -3.5V to 3.5V to capture cycle-to-cycle consistency performances. This basic test set takes 1200s to execute for each device. These preliminary data is then reviewed before continuing with more intensive

measurements depending on the aim of investigation for that particular device.

Some statistics were extracted from all the data collected from the probe station to provide further justification of the effort which went into building it. These data spanned from 16 Nov 2017 to 17 Feb 2019. Earlier data are ignored in this paragraph because they came from a probe station at a physically different location on campus. In the 458 days of this survey, 23016 I-V sweeps were carried out on 433 devices across 61 unique chips. These measurements took 148h combined. This is clearly not practical with a manual probe station where the experimentalist is expected to touch down a probe at each device and sit through with the device while the measurement runs. Data collected at the scale done for this project requires automation, both in terms of the hardware required to do probe positioning and software required to interpret the response to make decisions on subsequent steps, including how to recover from unresponsive device measurements or lose of probe contact. This is ideally done on a probe station dedicated for the project in contrast to one in a shared facility due to the long measurement times taken up.

The step in automation is also in line with the trend in using big data analysis in scientific research. However, this project is different in scope in that it focuses on the collection step of "big data", an aspect in the pipeline of big data analysis which is often neglected when research groups choose to focus more on applying algorithms on existing data sets. All these aspects in the pipeline of big data analysis, spanning from collection, to storage, and to analysis, are all essential components in the overall system or project. Fortunately, all these aspects have been made easier with the availability of libraries, tutorials or commercial services which have provided this probe station with capabilities with perhaps 5% of the effort than if it was built a decade ago. Libraries for hardware interfacing or data analysis are now open-source and friend to use. Tutorials now come from beyond books and journals, such as online forums and Youtube. Such new medium is expected to grow as an increasingly useful resource for an experimentalist-in-training as a complement to the traditional mentor when human time is increasingly more scarce as a resource. Finally, commercial services such as PCB manufacturing now available on the cheap, and provision of maker and machining

spaces on campus to support research, have all greatly improved the speed in which custom hardware can be created to tackle needs which not satisfied by off-the-shelves solutions. Going forward a decade, new tools will surface and render the following technical aspect irrelevant akin to how C++ subtleties in data analysis or bitstream aspects of device communication from a decade ago no longer demand attention. However, guiding principles in system design described in the above paragraphs will still hold.

## 3.2 Modular description of probe station

A systems view of the probe station is provided in Fig. 3-1. On the hardware side, provisions are included to allow electromechanical driving in 3 axis. Two of these axis are present as the X-Y axis on the stage holding the chip under test. The third is in the vertical axis of the arm holding the gold wire probe. Both the stage holding the chip and the arm holding in probe have other manual degrees of freedom (6 total) capable of fine positioning to 10$\mu$m, but coupling just 3 of these degree of freedoms to electromechanical actuators is sufficient to allow software-driven probe of arbitrary locations on a chip. Software driven stepper motors (17HM08-1204S, 0.9° steps, 1.2A, 15.6oz-in torque, NEMA17 face) control these 3 axis by pulling on positioning knobs through timing belts with a gear ratio that theoretically provide 0.6$\mu$m/step. The stepper motors can be driven either using an open-source GUI controller (GitHub gerritv/Grbl-Panel) or using a wireless gaming controller through a custom library. In practice, this positioning resolution is worsened to about 1 to 2$\mu$/step due to system backlash and other motion non-idealities, but more than sufficient for this project to contact the smallest 50$\mu$m square pads with ⌀25$\mu$m gold wire probes.

Gold wire probes were used since it is pliable and can be over-driven into a device without scratching it (see Fig. 3-2). This increases contact stability and minimizes the risk of vibrations coupling into the set-up in causing mechanical damage or loss of electrical contact. The gold wire probe also never oxidises under normal use, though it has to be trimmed or replaced completely after testing a device to destruction.

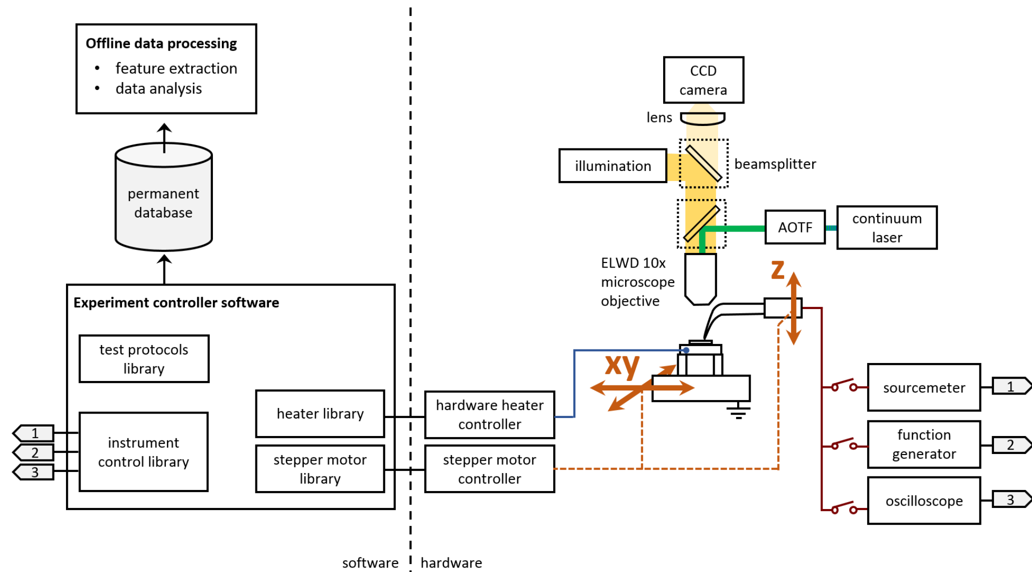



Figure 3-1: Breakdown of probe station showing software modules (left side) and hardware modules (right side). The software modules consist of a data collection, data storage and data analysis portion. The hardware modules consist of several modules as follows. The motion controllers (orange connections) drives the sample around in the X-Y direction and allows the probe to driven in the Z direction. The electrical test suite consists of a commercial off-the-shelf sourcemeter, function generator and oscilloscope. The optical modules consists of image capturing with a CCD camera, optical illumination with a white light LED and optical excitation with an AOTF-filtered laser from a supercontinuum laser. The heater module consists of a thermocouple probe and heater attached to the auxiliary stage, with PID control implemented by both a hardware controller and a software monitor. The details of the probing arrangement is in Fig. 3-2.

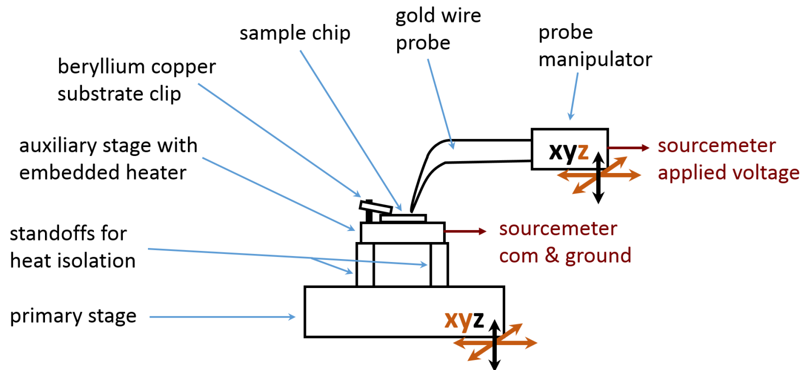



Figure 3-2: Zoomed-in detail of the probing arrangement. The probe manipulator hold a wire holder (not shown) which then holds the gold wire. The ⌀25$\mu$m gold wire is snipped at the end to produce a clean probing tip. A conductive beryllium copper clamp holds the substrate down and provides an electrical path to the sourcemeter common (connected to ground at sourcemeter). The sample is held on a heatable auxiliary stage with standoffs from the primary stage so as to minimize thermal losses into the primary stage. The primary stage and the probe manipulator each have 3 high resolution degree of freedom positionable to 10$\mu$m. 3 of these degree of freedoms were motorized (xy of stage and z of probe manipulator, with axes shown in brown) to enable software control while the other 3 remained as manual controls (axes shown in black).

The destructive testing causes debris to be coated to the tip of the gold wire and hinders good electrical contact with subsequent devices. Conventional tungsten probes (Signatone SE-T and SE-TB) are used in two special scenario. The first is in verification tests to confirm that measurements are not influenced by the choice of probing needle. The second is in electrical tests on devices with a top metal contact such as Ti, Cu and Al which oxidizes and thus have to be probed with a stiffer wire to break through this surface oxide. On the other hand, probing of Au, Pt and Ag surfaces can be done easily with the gold wire probe without hindrance from a surface oxide. Our measurement result is not affected by the material choice of the probing needle because it already has a top electrode. Other literature which perform conductive atomic force microscopy (CAFM) measurements directly on the switchable oxide would have to choose their wire probe material with more attention. The gold wire or tungsten probe are connected to a probe manipulator through various couplings (see Fig. 3-3).

Heating capabilities are provided by a 10Ω ceramic heater slab embedded in an

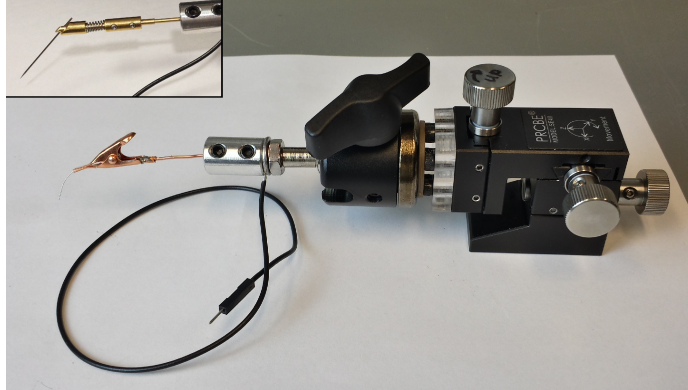


Figure 3-3: The probing assembly used is shown. In this picture, the $\varnothing 25\mu$m gold wire normally used is replaced with an coarse aluminum wire for greater visibility. When a tungsten probe is required, the front end of the probe assembly is replaced with an alternate probe holder (in inset).

aluminium channel, controlled via a 12V power supply toggled by a custom-built PID controller with temperature feedback via a K-type thermocouple sensing the surface temperature of the aluminum channel. This allows software control of heating to any user-specified temperature of temperatures less than 250°C. Higher temperatures can be obtained using a larger power supply. Commercial PID controllers with serial communication can also be used but the custom-built PID controller was available as surplus from some other work. The heater is elevated from the stage positioner to reduce heat loss.

Sample positioning is facilitated by a 1/2" color CMOS camera with white light illumination provided through a 50:50 beamsplitter (Thorlabs EBS1). The white light illumination was designed to be toggled on or off via software but the device was found to be insensitive to ambient lighting conditions. This sensitivity test was conducted by coupling a laser source through another beamsplitter (Thorlabs 70R:30T BST10R). The laser source is a white super-continuum laser (Fianium WhiteLase micro) filtered through an acousto-optical tunable filter (Fianium AOTF-DUAL). This optical excitation on the probe station was included to support another project but not further used in the characterisation of our RS devices.

Benchtop measurement equipment are connected to the device under test through the wire probe and with the return path through the substrate. A copper-beryllium

clip pinning the chip down provides this substrate contact. All substrates used were chips from n-type degenerately doped Si wafers for maximum substrate conductivity. Typically, only a sourcemeter (Keithley 2450 Sourcemeter) is hooked up to the device under test, but an arbitrary function generator (Tektronix AFG3052C) and oscilloscope (Agilent MSO-X 2024A) can be be connected in parallel to the device for pulsed mode measurements. All these benchtop equipment are exposed through the software controller program via USB (through the VISA protocol).

The software controller stack is built with Python, starting from libraries implementing communication with devices mentioned previously (stepper motors, CMOS camera, benchtop equipment) and higher-level libraries implementing the experimental procedure. Error handling code deals with unresponsive equipment by triggering 110AC power relays to reset the culprit equipment without human intervention. All measurements are tagged with sample, date time and equipment settings metadata and stored in a database for later deeper analysis, though a graphical summary is also automatically generated in real-time for immediate feedback.

A separate analysis library was written to simplify the simultaneous analysis of all the data collected. This is necessary to identify small trends hidden behind measurement noise, or to revisit data from old measurements from another perspective guided by fresh insights. Occasionally, measurements are wrongly tagged with erroneous sample information by human error. Such mistakes are immediately fixed by a traceable patch code so that the database is kept clean and accurate for ease of scientific analysis.

## 3.3 Summary

In summary, a custom probe station with semi-automated features was built from scratch with various different hardware and software modules put together as described above. This required great effort as opposed to using a generic probe station which would already be available at most shared microelectronics facilities. However, such customization and automation is necessary when dealing with poorly studied devices

whose behaviors are either unknown or not easily captured at a conventional probe station. Nonetheless, future researchers might be spared this effort in the future when all-in-one probe stations could become commercially affordable.

# Chapter 4

# Strategies for improved consistency in resistive switching

Poor device yields and low switching consistency in resistive switching devices can be attributed to two main causes which we will demonstrate how to tackle. Low device yields are largely due to the stochasticity in initial filament formation, a.k.a, electroforming[50]. We tackle this by using metal atom implantation to form a seeding site for filament formation with high yields. Earlier studies to test the effects of embedded nanocrystals in oxides were promising in improving device yields but had other performance limitations[51, 52] or were confined as empirical studies[53, 54]. Poor device-to-device and cycle-to-cycle switching consistency is due to large variances in the physical form of the filament from cycle to cycle[50, 55, 56].

We tackle this by splitting up single thick films into multiple thin films separated by conductive thin spacer films. In additional to the better performance, our strategy is also more scalable than the use of bilayer oxides now popular[57, 58, 59] for improving switching consistency because our oxide-conductor bilayer can be made thinner and repeated with a lower tendency for the thin films to interdiffuse due to better phase segregation characteristics between oxides and conductive nitrides[60]. From experimental data, resistive switching appears to be significantly more consistent for a series of multiple short oxide gaps in contrast to one large oxide gap. Each gap has to be spanned by some filament, web of filament, or interfacial conduction for the

gap have a drop in resistance. The two above approaches were used in conjunction to show that fabrication yields, device-to-device and cycle-to-cycle switching consistencies can all improve significantly with modified approaches of fabrication.

## 4.1 Method

We first describe a basic set of deposition procedures and the characteristic methods used. Counter experiments and variations were also performed to identify both critical and non-critical steps and are described in later paragraphs and in chapter 5.

A n-type degenerately-doped Si wafer was used as the substrate. The wafers were first dipped in HF to remove native oxide. Next, the wafers were loaded into a plasma-assisted atomic layer deposition (PEALD) equipment for deposition of a multilayer thin film stack. All devices have a stack which starts with a layer of 7.5nm tungsten nitride (WN) to serve as a passive conductive common layer to interface between the Si substrate and the subsequent active layers of Al2O3 and WN. A well-performing active stack further consists of three 2.0nm Al2O3 films separated by two 3.0nm WN films (as in Fig. 4-1). The wafer is then taken out of the PEALD chamber, cleaved into chips, and sputtered with 30nm Au. At this point, all the deposited films are still planar and unpatterned. About 6 to 15 counts of 50$\mu$m square devices are then defined on the chip by Ga-ion FIB to mill away the added materials and past the native Si surface so that a mesa remains. The top Au film of the mesa is contacted with a gold wire probe under a custom-built probe station. The common (COM) connection was made by a beryllium copper spring clip contacting the substrate and connected to ground. The sign of voltages described in this paper is defined as applied to the top probe with reference to a grounded Si substrate. A sourcemeter was used to perform voltage sweeps to switch the devices for study.

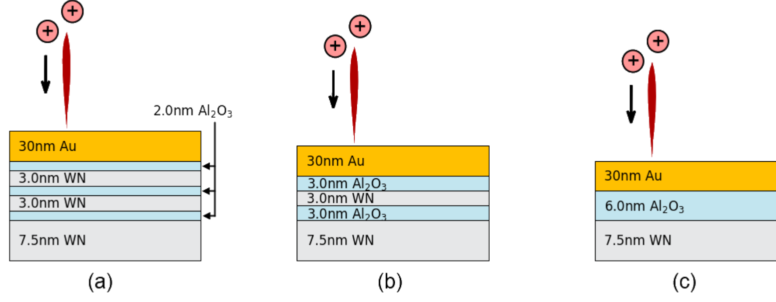



Figure 4-1: The properties of using multilayer oxides are investigated using the material stack as shown. The total thickness of the oxide was kept constant, because single layer oxides with a thickness of either above or below 6nm tends to breakdown permanently, and changing the total thickness of the oxides changes the LRS and HRS by several orders of magnitudes.

## 4.2 Experimental observations

The first observation with the electrical measurements are that all the devices are in the low resistance state upon fabrication with no need for electroforming. The devices exhibit resistive switching with an abrupt RESET voltage at about 3V and a gradual SET voltage spanning -1V to -4V (see Fig. 4-2d). Here, we will use a new measure of switching consistency as described in chapter 2 to evaluate the merits of our device. With the increase in emphasis in this field on demonstrating device switching uniformity, it is now necessary in this field to show a cumulative distribution function (CDF) of device characters. Even though there is no consensus on whether to use the device resistance or the electrical currents for this plot and on what probing voltage to use, the spread of this line is largely invariant on these choices and thus serves as a useful figure of comparison for quantifying how consistently a device switches. As an example, a 10% to 90% spread of 1.5 and 0.91 on the log scale can be extracted from the published LRS CDF plots from Qi[31] and Tsigkourakos[61] respectively (see Fig. 4-2c), indicating that the latter device has a generally more consistently switching behaviour. We denote this as the log coefficient of variation $C_{lv}$, an extension to the textbook statistical measure of the coefficient of variation $C_v$ better suited for other variables which do not vary orders of magnitude as is commonly seen in this field.

Many published resistive switching devices consist of only one single deposition

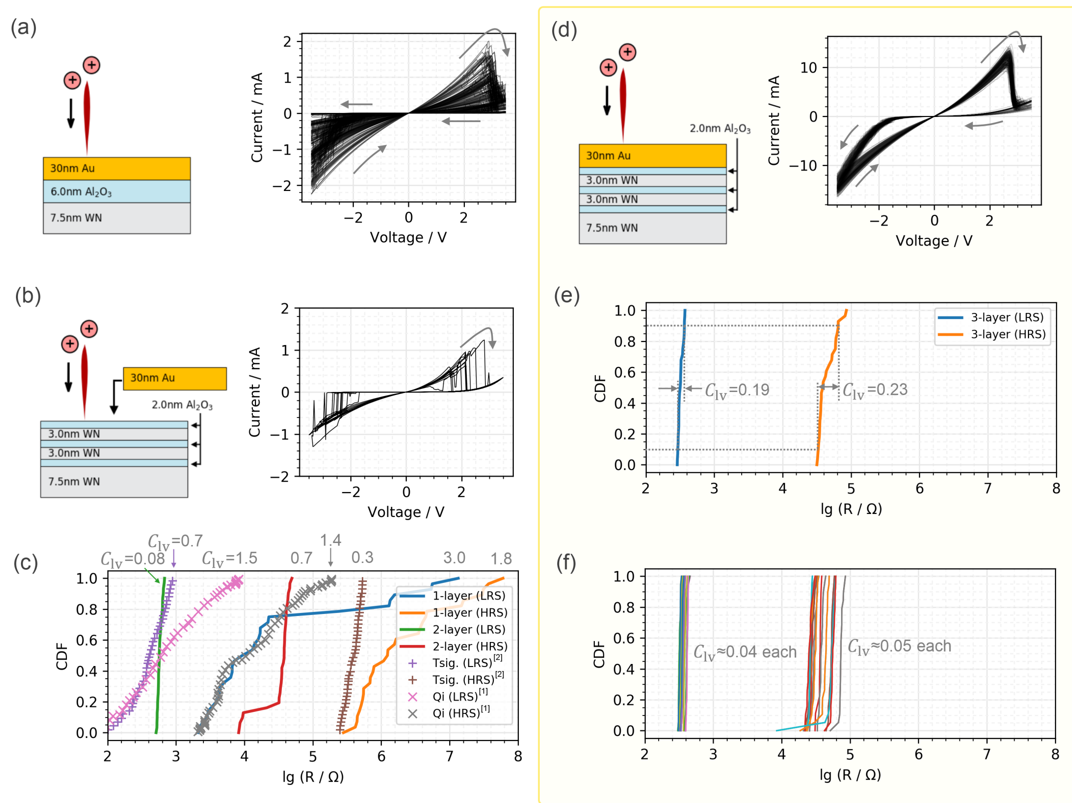



Figure 4-2: FIB-processing and multilayer oxide films are 2 key ingredients for our device to demonstrate consistent switching. (a) A single Al2O3 oxide layer can switch, but only erratically. (b) FIB-processing has limited effect on device switching consistency if performed on the multilayer stack before deposition of Au as the top metal electrode. Here, FIB-processing has poor effect because there are no Au atoms to be implanted deeper into the device. (c) Cumulative distribution function (CDF) plots show the typical large variation in the performance of resistive switching devices. The LRS and HRS resistances can span anywhere from between 0.08 to 3 orders of magnitude. We define the span to be between the 10th and the 90th percentile of the resistances observed. (d) With multiple Al2O3 oxide layers and FIB-processing on top of a deposited Au layer, switching immediately becomes consistent. The I-V plot shows 300 switching cycles superimposed on top of one another, taken from 15 consecutively fabricated devices and 20 consecutive cycles per device. (e) The consistency improvement using our stategy can be seen from the short span of resistances observed over the 300 switching cycles, covering only 0.19 and 0.23 decades for the LRS and HRS respectively. (f) The CDF in (e) can be broken down for the 15 devices. The cycle-to-cycle resistance span is now much smaller at less than 0.05, indicating that most of the variation seen in (e) comes from device-to-device differences rather than cycle-to-cycle variation.

stack. This limits understanding because these large variability of these devices means that published devices from different groups cannot be compared, but then a single device stack is not useful for identifying any links between certain fabrication parameters and the resultant switching property. We first identify all the critical steps in achieving a consistently switching device, and then show acceptable deposition variations.

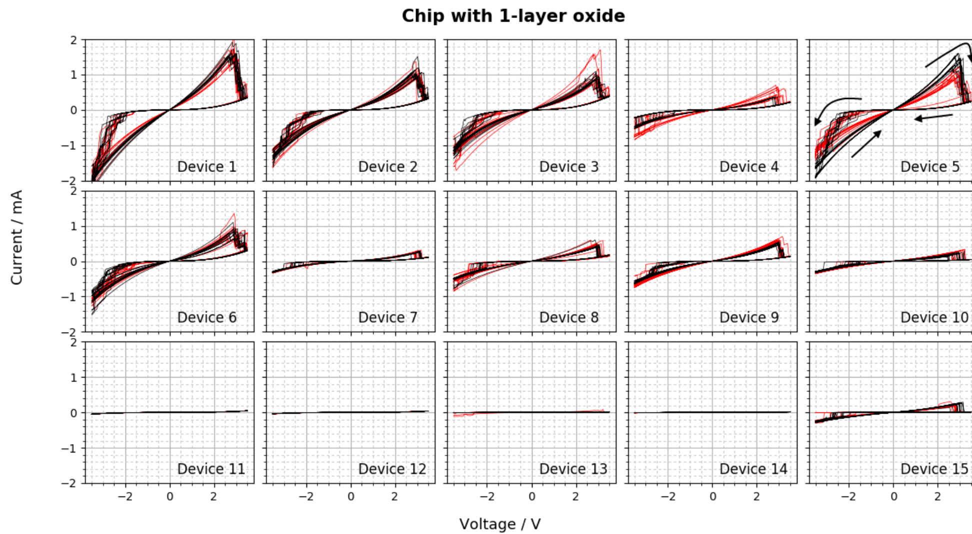



Figure 4-3: I-V sweeps for 15 consecutive devices with only a 1-layer oxide. For each device, 20 I-V sweep cycles are superimposed onto one another, with the 1st 10 cycles shown in black and the next 10 cycles shown in red. This geometry gives a poorly switching device with high variability from cycle-to-cycle and device-to-device.

Switching consistency is considerably improved going from one to two to three layers (shown in Fig 4-2a,c,d,e). This deduction was obtained by analyzing 300 sweep cycles for each one of the 1,2 and 3-layer geometries, using 15 devices investigated per geometry (see Fig. 4-3 and 4-4). We keep the total thickness of Al2O3 constant so that the expected LRS and HRS resistances stays approximately constant in these variants. The CDF plots for the LRS and HRS resistance can be broken up across device and cycles to identify the source of variable (see Fig. 4-2f). The breakdown shows that with a 3-layer oxide device, the switching variance comes mainly from device-to-device variation, since the cycle-to-cycle variation only has a small spread of about 0.04 and 0.05 on average for the LRS and HRS states respectively. This implies

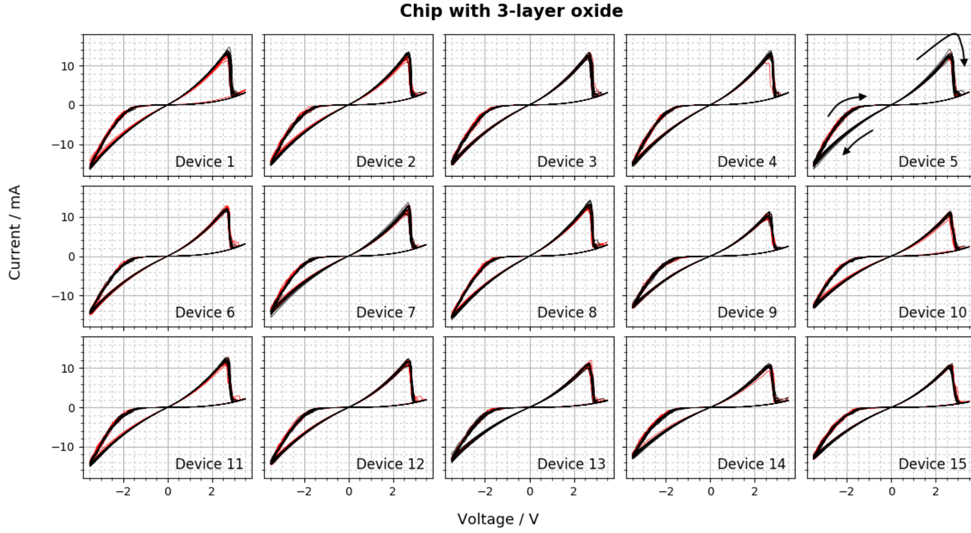



Figure 4-4: I-V sweeps for 15 consecutive devices with a 3-layer oxide. For each device, 20 I-V sweep cycles are superimposed onto one another, with the 1st 10 cycles shown in black and the next 10 cycles shown in red. This geometry gives a good RS device with high switching consistency across cycles and across devices. This 3-layer oxide configuration is our proposed geometry for achieving high consistency RS devices.

that fabrication variation accounts for most of the variation seen in the combined CDF plots of Fig 4-2e.

## 4.3 Interpretation of results

We propose that the mechanism for the activating effect from FIB is that Au atoms at the surface are being implanted into the Al2O3 layers when the surrounding of the mesa is being milled away (see Fig. 4-5). The implanted Au atoms then serve as stepping stones to bridge migrating oxygen vacancies to more easily form a filament which can span across each Al2O3 layer. To verify this hypothesis, we swapped the order of the final 2 deposition steps by FIB milling a line onto an ALD deposited repeated WN/Al2O3 stack before depositing a Au contact pad over the milled line. With this change, the device switching consistency degrades, in support of the hypothesis (see Fig 4-2b). Since this result suggests that Au plays an important role in the switching process, we substituted the Au top metal contact for either Pt, Pd, Cu, Al or Ti to

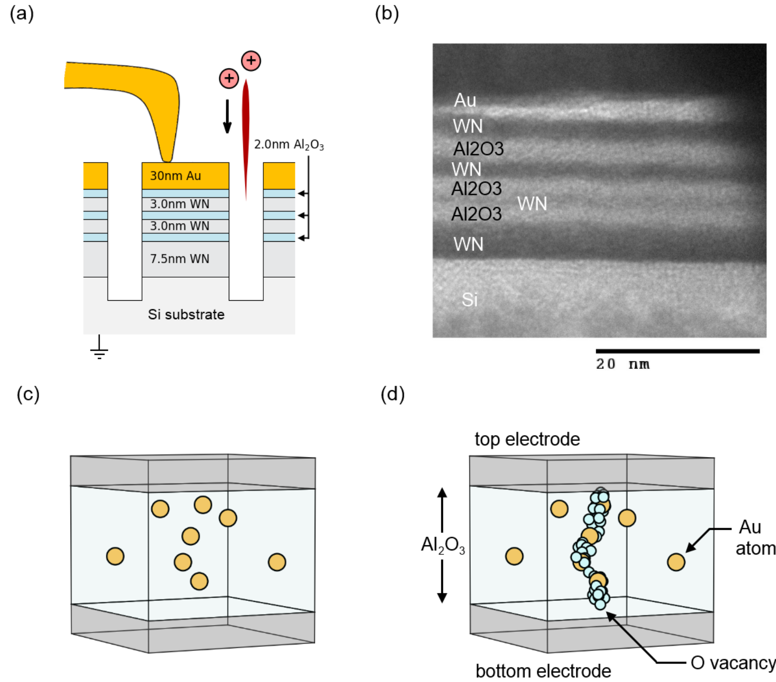



Figure 4-5: (a) Schematic for experimental probing of fabricated devices. Thin film WN and Al2O3 layers are deposited by plasma-enhanced ALD (PEALD) and is followed by Au deposition. FIB milling is used to isolate a mesa to define a single device. A gold wire is then used to contact the top gold electrode, with probe voltages defined with respect to this top contact.
(b) TEM image showing the films deposited by PEALD. The film thicknesses are for a different stack than show in Fig (a).
(c) The FIB milling is thought to implant Au atoms into the Al2O3 layer.
(d) Our DFT simulations and experiments suggest that oxygen vacancies are migrating to form or break filaments around Au dopant atoms. The DFT simulations suggest that these vacancies have a great ease of migration because of its lower formation energy at sites around Au dopant atoms. Only one filament is shown but the actual mechanism could involve a web of interconnected filaments.

observe the changes in switching behavior. As seen in Fig. 7.3, noble metals such as Pt, Au, and Pd gave devices with consistent switching behaviors, whereas the more reactive metals have erratic switching cycles. These dopant atoms are thought to be implanted by the action of the 30kV Ga FIB ion beam. SRIM calculations revealed that for this FIB parameter, surface Au atoms can be implanted to a depth of 20±10nm into Al2O3.

Resistive switching in many systems have been shown to be due to migrating oxygen vacancies. In order to understand how metal dopant atoms affect switching, we performed DFT studies to calculate the formation energy of oxygen vacancies in the vicinity of dopant atoms. Dopant atoms were placed at the interstitial sites as opposed to the less energetically favorable substitutional sites. In the vicinity of these dopants, we find that oxygen vacancy formation is vastly reduced for noble metals as opposed to the more reactive metals, spanning the range of 0.468eV for Pt to 4.383eV for Ti. This result is sensible because oxygen atoms does not have a tendency to form bonds with noble metal atoms and so oxygen vacancies are expected to form more easily near noble metal atoms rather than reactive metal atoms. There is a good general agreement with the calculated oxygen vacancy formation energy in the presence of each particular metal dopant with respect to the switching consistency of a device made with that material as the top metal contact. Higher order effects will have to be invoked to explain the differences in the switching behavior between atoms of similar reactivity such as between Pt and Au, or between Al and Ti. Such an analysis is likely to require the analysis of possible stable oxides and hydroxides identities for each dopant and is beyond our current abilities.

### 4.3.1 Multibit operation

The switching consistency of our device can be exemplified by multibit (also known as analog) resistive switching trials too. Since this device have a gradual SET transition, the strength of the conducting filament within a device can be variably set by using a SET voltage of some value between -0.5 to -4V. A more negative SET voltage puts the device in a more conductive LRS state (see Fig. 4-7c), and subsequently also has

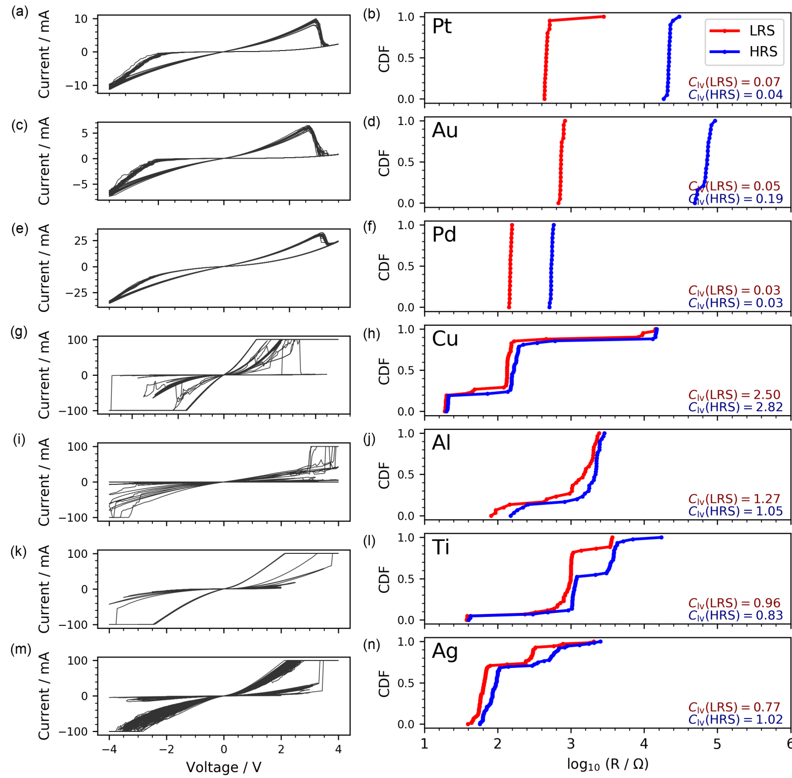



Figure 4-6: The use of a variety of metals, Pt, Au, Pd, Cu, Al, Ti, Ag (corresponding to (a),(c),(e),(g),(i),(k),(m)) was studied. A 100nm thickness of this material was deposited on the usual triply repeated WN/Al2O3 stack. Each metal gave rise to its own distinct switching behavior. The cumulative distribution function for the LRS and HRS resistances are plotted in (b),(d),(f),(h),(j),(l),(n). Pt, Au, Pd have CDF plots with narrower widths (tabulated in Table I), indicating that devices made with these metals tend to have consistent switching behaviors. Ag has a known behavior in this field of being highly susceptible to electromigration, explaining its behavior to be stuck permanently in the LRS probably by formation of a large filament.

| Dopant | Simulations $E_{f,\text{vac}}$ / eV |
|---|---|
| Pt | 0.468 |
| Au | 0.644 |
| Pd | 1.913 |
| Cu | 3.617 |
| Al | 3.632 |
| Ti | 4.383 |

Table 4.1: Varying the identity of the top deposited metal film before FIB milling has the effect of altering the dopant identity implanted into Al2O3. Devices implanted with noble metals such that Pt, Au, Pd has a significantly more consistent switching behaviour than more reactive metals such as Cu, Al, Ti. We hypothesized that the metal dopant atoms serve as bridges for facilitating oxygen vacancy migration. We found from DFT simulations (further described in Appendix B) that the formation energy of oxygen vacancies in Al2O3 is significantly lowered by the presence of noble metal dopants such as Pt and Au. There is a fair agreement between the oxygen vacancy formation energy and the switching consistency seen in Fig. 7.3.

a higher abrupt RESET transition. It is important to point out that this multibit switching is demonstrated using a blind strategy, i.e, without any feedback control to read back the resistance state and make adjustments, and also, without the use of external electronics to enforce a SET current compliance. The lack of external control circuitry will be useful in reducing the effective footprint of each cell for future RRAM arrays. From a quick analysis, there are at least 7 distinguishable states for this multibit switching (see Fig. 4-8). The definition for this bit count is not yet standardized and is a subjective variable because the amount of tolerable error, or equivalently, the clearance between the distribution windows of adjacent states, can be arbitrary defined. Nonetheless, a conservative estimate for our device is that it offers at least 2.8 bits of storage, corresponding to the 7 identifiable state. A change in our experimental procedure should also yield a higher bit count because our multibit switching tests was conducted using finely-stepped SET voltages, whereas the use of just $N \approx 8$ to 10 voltage points would yield distributions with a larger clearance between one another.

A few other deposition requirements exist in addition to the need of multiple oxide layers for consistent resistive switching. The use of FIB in defining the device was necessary. As a counter example, a chip deposited with the same repeating stack of

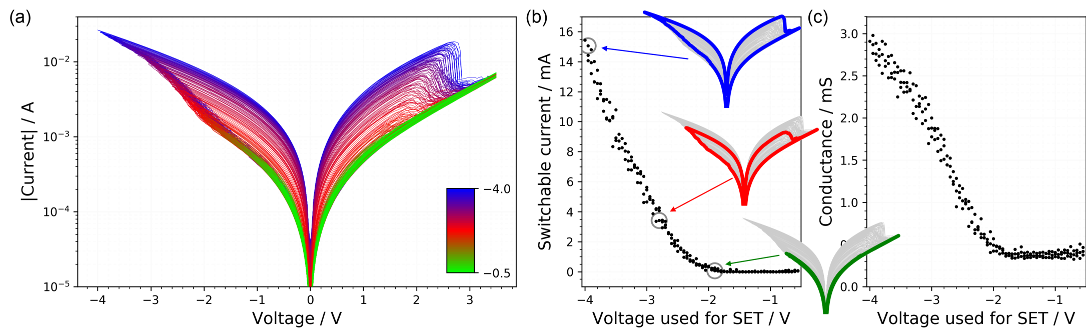



Figure 4-7: The gradual SET process between -0.5 to -4V is convenient for demonstrating multi-bit switching. In (a), multiple switching cycles were performed with some predefined minimum voltage of between -0.5 to -4V and a fixed maximum voltage at 3.5V. Each switching cycle is color-coded according to that minimum voltage used. For example, a voltage sweep from -4V to 3.5V is shown in blue, has the most hysteresis, and the sharpest RESET onset at about 2.8V. On the other hand, a voltage sweep from -1V to 3.5V is shown in green and has barely any hysteresis or switching. Subplot (b) and (c) shows that the simple use of a chosen SET voltage is able to put the device into a predictable state as characterized by the switchable current in (b) and the device conductance in (c). The switchable current is defined as the difference in the device current as set and the lowest current measured in the OFF (highest resistance) state. The green, red and blue overlays in (b) show the nature of one I-V sweep when the device is set by -2V, -2.8V and -3.9V respectively. No current compliance was needed to be programmed in the sourcemeter used for these measurements.

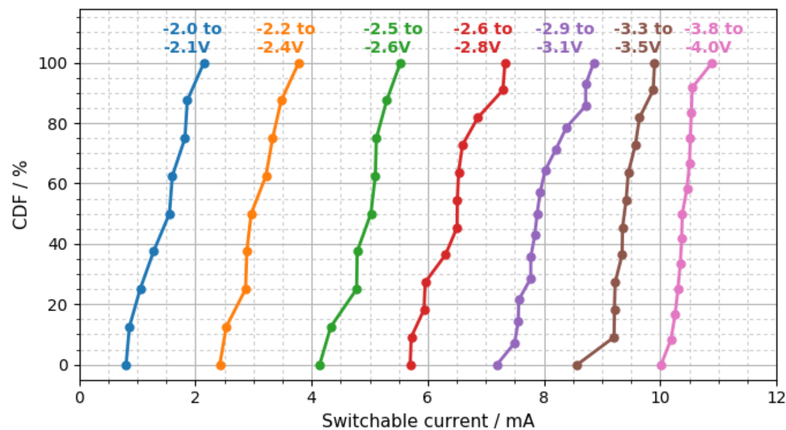



Figure 4-8: Approximately 7 states are distinguishable by choosing to use different SET voltages. The states are distinguishable by the non-overlapping window in their resultant switchable currents. The switchable current is the difference in the device current in the LRS versus that in the HRS, and is an indication of the strength of the filament formed in the device.

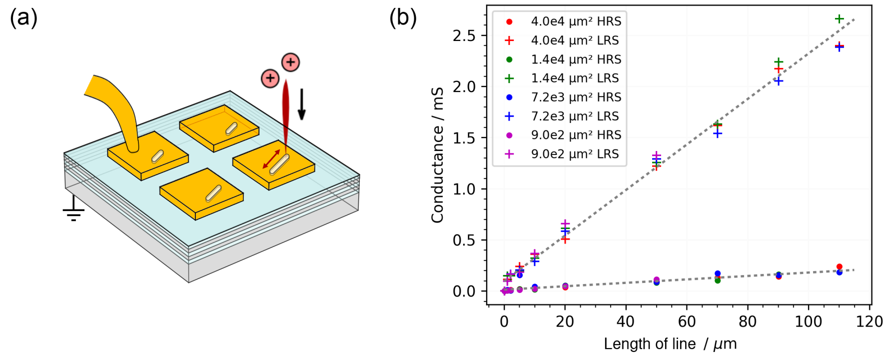



Figure 4-9: (a) A 30nm thick Au contact pad of variable size was deposited with e-beam deposition through a shadowmask onto the usual triply repeated WN/Al2O3 stack. A line with variable length was then milled into this Au contact pad with an FIB machine. (b) There is a linear relation of the obtained device conductances with the FIB milled lengths but no dependence on the contact pad area. This shows the device character is directly determined by the FIB milling and implies that the device area can be scaled down arbitrary.

WN and Al2O3 in the PEALD chamber was patterned with Au contact pads through a shadowmask using e-beam deposition. The devices were measured under a probe station and found to require large electroforming voltages of >10V, had wildly variable characteristics from device to device, and often failed by being permanently stuck in either the LRS or HRS. These devices with e-beam deposited contact pads can be made to demonstrate consistent switching by FIB milling lines into them. Multiple devices with e-beam deposited contact pads of various sizes and milled lines of various lengths were created. The conductance of both of the LRS and HRS was found to vary linearly with respect to the length of the milled line while being independent on the device area (see Fig. 4-9). We did not have enough resources to explore devices of sub-micron length scales, but the direct dependence on the resistive switching behaviour and the FIB milling and null dependence with device area indicates that this device can be scaled down significantly.

A robust resistive switching system is one which maintains its performance and switching character under varying conditions as long as these conditions do not disrupt the hypothesised mechanism of switching. The identification of unimportant process parameter ought to be as critical as the identification of critical process parameter so

that the fabrication process is not needlessly strict. The deposition method of the top Au layer is inconsequential as consistent switching is observed regardless whether the Au deposition is performed using e-beam deposition, DC or RF sputtering, or even using a tabletop Au sputterer. PEALD was used for the deposition of the repeated WN/Al2O3 stacks as the machine is an industrial grade tool in a strictly controlled process environment and can be reliably expected to produce repeatable depositions. We have recently migrated the recipe for deposition of the same stack with reactive sputtering in a RF/DC sputterer with success. As a side note, thermal ALD of the WN/Al2O3 stack did not produce a switchable device. This turns out to be simply due to the bis(tert-butylimino)bis(dimethylamino)-tungsten precursor for WN not adhering on preceding Al2O3 layers but is potentially solvable by some novel modification of the thermal ALD recipe.

Lastly, we also have initial success with substituting the FIB milling treatment with co-sputtering of Au with O2 reactive sputtering of Al to replicate the effects of Au implantation. The HRS and LRS of the Au co-doped device has a higher resistance than is typical of the usual FIB milled processed devices, but both these resistances can be lowered by increasing the doping concentration of Au. This modification will eliminate the need to do FIB milling processing or Au ion implantation into our device to activate these device for resistive switching. The use of co-doping to incorporate Au dopants into the Al2O3 layer will be a useful and independent parameter which can be independently used to adjust the HRS or LRS resistance as needed. In contrast, the magnitudes of the HRS and LRS resistances can be tuned by varying the thickness of the oxide in switching films but this would be an undesirable method because the switching behaviour is largely dependent on the oxide thickness. Furthermore, the division of a thick oxide film for switching into multiple thin oxides is beneficial not just from its contribution to switching consistency, but also the potential in pushing LRS and HRS resistance states higher than routinely possible without the erratic behavior that comes from the use of just one or two stacked oxide layers.

## 4.4 Conclusion

We demonstrated a device with highly consistent resistive switching behavior and shows that these properties are highly repeatable, not just across devices on the same chip, but also across different deposition methods, some different materials and also potentially with co-doping instead of ion implantation.

The key requirements for our resistive switching device to have a favorable switching behavior is that Au needs to be implanted within the Al2O3 layer as dopants, which we believe to act as bridging sites which facilitates the formation of oxygen vacancies around it. The replacement of a single thick layer of oxide into multiple thinner layers is also essential for minimizing switching stochasticity. The fabrication strategies reported here will directly benefit efforts in realizing RRAM circuits with higher yields and predictable performance, etc.., and move multibit switching another step closer for practical use. However, an equally important outcome to is provide a device platform for researchers to continue the study of resistive switching devices. With a robust and highly consistent resistive switching device, we are able to derive new understandings about the switching mechanism by linking changes in the switching behaviour to fabrication changes since any behavior changes will not simply obscured by the variability of a typical resistive switching device.

# Chapter 5

# Further investigations

As detailed in subsection 1.2.2, the pursuit of direct imaging evidence of RS filaments is challenging. Nonetheless, we can rely on electrical measurements remain promising as a tool to uncover deeper insights into the RS mechanism to support our hypothesis. Developments in electrical characterization have allowed benchtop instruments to reach measurement precisions of femtoamps. As a numerical example, 100fA is the equivalent of the passing of 625 electrons across a point in 1ms and can be measured by a benchtop ammeter. On the other hand, direct imaging of 625 atoms requires a HRTEM, and certainly no less than an AFM. There is thus a strong incentive to find out how to use electrical measurements to supplement other types of investigation routes because these methods will be the most accessible to the community and can be easily reproduced.

In this chapter, we show how subtle characterizations can be extracted from electrical measurements now that the noise from I-V data due to switching inconsistencies have been largely eliminated. We also pursue investigations from other approaches, such as through simulations, to further support our hypothesis that the existence of implanted Au dopant atoms are responsible for the highly consistency switching behavior in our devices.

## 5.1 Heating and aging measurements

The concurrent presence of two filamentary constituents in our devices, namely the Au dopant atoms and the oxygen vacancies, can be inferred from aging results as well. After a device was annealed at 300°C for 3 hours, the resistance of HRS increased by about 8x (see Fig. 5-1a). This is expected from previous studies and is thought to be due to filamentary material migrating out of the oxide layer[62]. More interesting, the effects of a two months room-temperature aging differs depending on whether the device was left to age after conducting I-V measurements on it, or before any measurement (i.e, untouched since the FIB-processing). When the device was aged between I-V measurements, the aged device shows the same character of a gradual SET and abrupt RESET albeit with both HRS and LRS becoming more resistive as expected (see Fig. 5-1b). However, if the device was aged after FIB-processing without first an initial I-V measurement, the aged device then shows abrupt SET and RESET transitions uncharacteristic of our typical device (see Fig. 5-1c).

This can be understood in the context of the DFT studies described above where the Au dopant atoms can act as reservoirs for oxygen vacancies. When the device are switched before letting to age, the Au dopants atoms are stabilized and pinned in place by the surrounding oxygen vacancies, and so will retain a similar switching character with time. On the other hand, a device left to age straight after the FIB-processing will see the implanted Au atoms phase segregating out of the oxide and the subsequent switching character loses the identifiable gradual SET character distinct in our Au-implanted device.

There are a few implications from these heating and aging measurements. First, these results provides good support for our hypothesis that there are two filamentary materials present in our system, and that migration of oxygen vacancies bridged by Au dopant atoms are the primary filamentary formation mechanism in our device. Secondly, these results point out the limits of our devices with operating in high temperature environment or over prolonged periods by revealing shortcomings of FIB-processed devices under those conditions. A device intended for manufacturing

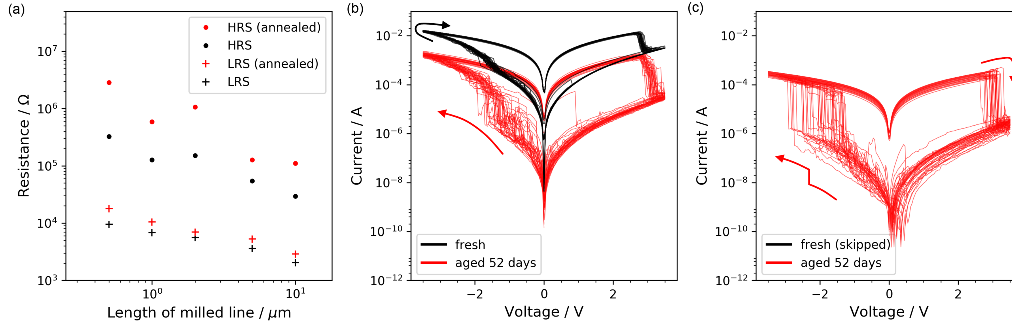



Figure 5-1: (a) Devices annealed at 300°C for 3h become more resistive both in the HRS and LRS. These devices were FIB-processed by milling a line of variable length to activate the devices for switching. (b) A device was characterised after FIB-processing and also after room temperature aging. The resistances of both HRS and LRS as expected, but the gradual SET character at V<0 is retained. (c) When a device is aged without first conducting I-V sweeps on it, it ages differently and now loses its gradual SET character. The differences in aging in (b) and (c) show that multiple filamentary constituents are likely to be present.

must thus address these issues of high temperature and long term stability. The proximity of Au atoms near FIB-milled exposed surfaces of the device is probably the cause of long term instability as dopant Au atoms have fast surface diffusions and have a high likelihood of phase segregating out of the Al2O3. This can be addressed in the future for example by new methods of doping immobile Au islands into the Al2O3. This is further described in the chapter on future directions in section 7.3, but we already see positive results from preliminary trials.

## 5.2 Bipolar nature of triple-layer oxide devices

RS devices can be generally classified into either the unipolar or bipolar class. There is currently no definitive understanding of how either class of switching would arise in a fabricated device, but the empirical experience is that RS devices with asymmetric materials for the top and bottom electrodes tend to exhibit bipolar behavior, whereas RS devices with similar materials for the top and bottom electrodes tend to exhibit unipolar behavior[99].

For a single layer oxide RS device, a dissimilar top and bottom electrode could

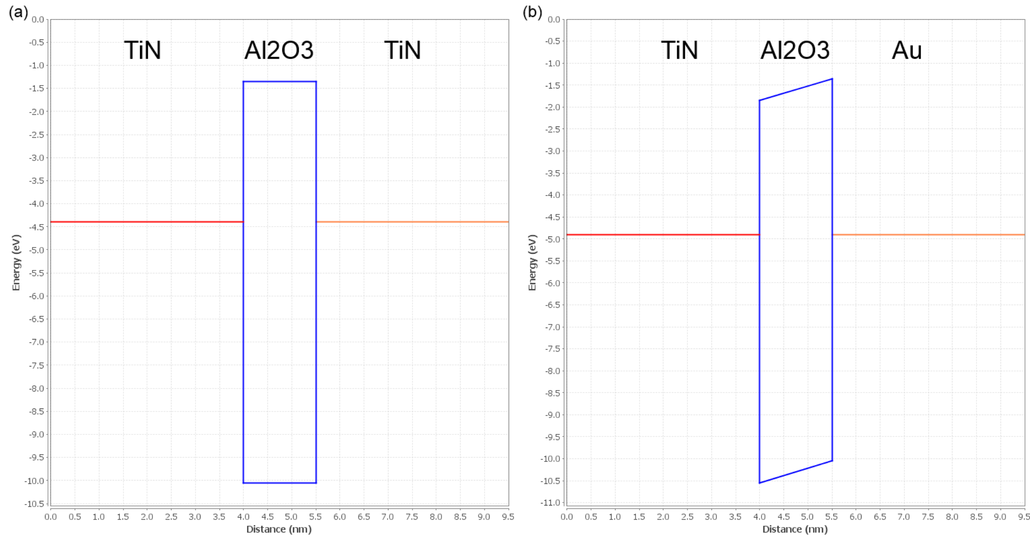



Figure 5-2: Energy diagram of a single-layer oxide RS stack with (a) similar top and bottom electrodes, and (b) dissimilar top and bottom electrodes, both shown at zero applied voltage. In subplot (a), there is no built-in electric field in the oxide and the RS device created with such a stack is predicted to be unipolar. In subplot (b), there is a built-in electric field shown by the tilted band edges in the oxide and the behavior of the device is predicted to be bipolar.

lead to a bipolar device because of dissimilar work functions in the two electrodes such that there is some built-in electric field at zero applied voltage. This is shown in Fig. 5-2 performed using the Band Diagram Program v3.1.6[100, 101]. The WN normally used in our device was replaced with TiN in the simulation due to a limitation of simulation choices. The workfunction used for TiN is 4.4eV and the workfunction used for Au is 4.9eV.

Our typical devices consist of an Au top electrode, a WN bottom electrode, and WN conductive spacers. This satisfies the criteria to obtain a bipolar device even though the conductive spacers are of the same material as the bottom electrode. This is because a built-in electric field will exist at all three $Al_2O_3$ layers even though only one of the oxide layer is in direct contact with asymmetric metals (see Fig. 5-3). The difference in the work function between the top-most electrode and the bottom-most electrode will be distributed across all the oxide layers in the stack and manifest as the built-in electric field.

The intervening WN metallic layers have no built-in electric field because sheet

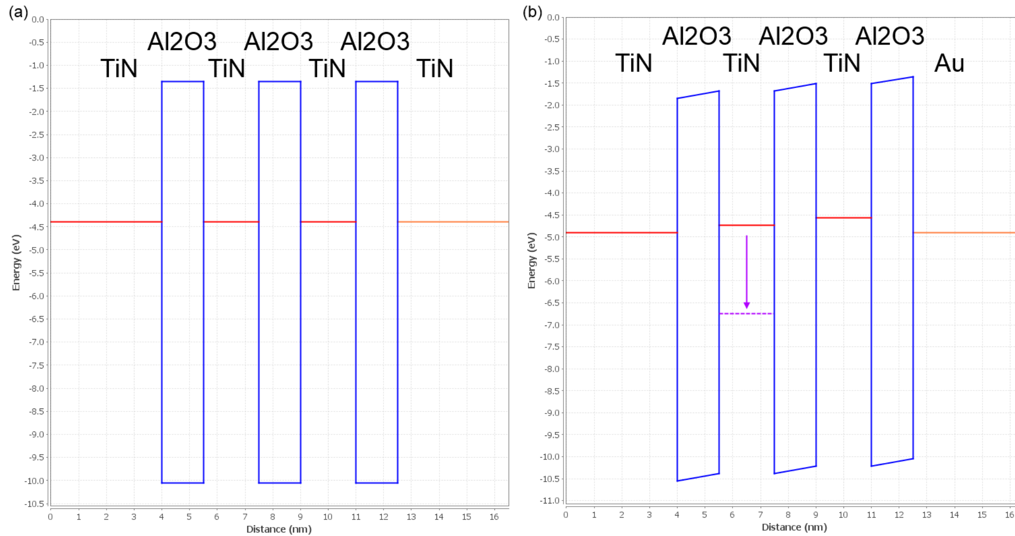



Figure 5-3: In (a), no built-in electric field is observed if the top and bottom electrode material is the same. In (b), a built-in electric field will exists across all the 3 oxide layers as long as the workfunction of the top and bottom electrode is different. If one of the intervening conductive layer was replaced by another metal with a workfunction of 2.0eV higher than than of TiN, the only effect to the entire band diagram is the shifting of the metal potential downwards by 2.0V (shown in purple dashed line) with other parts of the band profile left unchanged.

charges will appear at both its top and bottom faces to shield its center from having an electric field. This is expected behavior for a metal.

Thus, each one of the 3 oxide layers are expected to switch in a bipolar fashion as a result of the different top and bottom electrode, regardless of the materials of the intervening conductive spacers.

## 5.3 Transient as-fabrication features

Multilayer-oxide RS devices processed with FIB do not require electroforming at a large voltage. As supported by DFT simulations and experiments, the presence of noble metal atoms lowers the formation energy of oxygen vacancy (see. Table 4.3) and this is expected to lower the electroforming voltage to potentially within the normal switching voltage range. Furthermore, the 30 keV high energy ion implantation used could be seeding the oxide layer with oxygen vacancies from beam-induced damage,

so further production of more oxygen vacancies with electroforming is not necessary.

However, the lack of need of electroforming does not mean that an as-fabricated device exhibits the steady state highly-consistent switching straight away. The first cycle of our RS device shows a transient behavior before the steady state cycles can be obtained.

Newly fabricated devices are always first characterized on the electrical probe station with a basic set of tests (as listed in appendix C). The test set first applies a low range of voltage to the devices to extract the as-fabricated device resistance before exposing the device to two rounds of incrementally increasingly voltages. The first round of this test shows a transient behavior (see fig. 5-4a) before setting into the steady state behavior (see fig. 5-4b) that is displayed in the 2nd and also subsequent rounds of I-V sweeps.

FIB-processing removes electroforming that is otherwise required at large voltage of >8V and reveals a transient behavior which can offer more insight into the device operation. The transient I-V curves shows that the as-fabricated device stays in a higher resistance state up to a large voltage before gradually settling into the steady-state LRS limit as the maximum voltage applied to the device increases from 2.5V to 3.5V. It is only after this initial round of voltage application that the RESET behavior achieves its characteristic abrupt transition. This transient behaviors are consistently observed in all the devices characterized.

This repeatable transient behavior would be useful in future models for fitting the distribution of metal dopant atoms in as-fabricated FIB-processed device. This could provide insights into the implantation depth, concentration or potential clustering of implanted Au atoms. This is left as future work because coarse estimates for such a distribution will first need to be obtained through ion implantation simulations or through challenging measurements of the Au concentration such as by HRTEM or APT. Qualitatively, this transient behavior provide evidence that atoms are being implanted sideways into the device and that the initial I-V cycles serve to sweep stray materials, either excess Au dopants or defects created by implantations, away from the oxide bulk. This then allows remaining materials to take part in filamentary formation

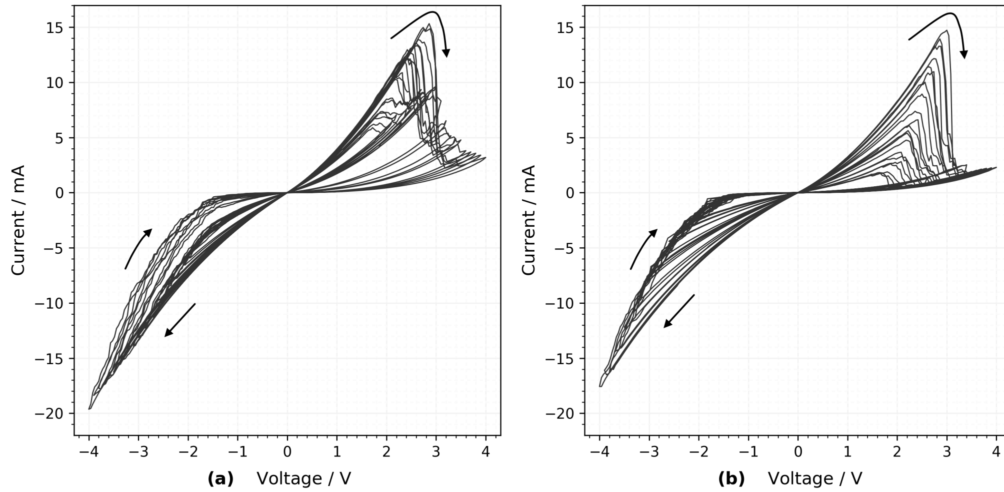



Figure 5-4: The first round of incremental voltage application on a freshly fabricated device (a) shows transient effects where the device current is higher than when during the second (b) and subsequent rounds of I-V sweeps. The applied voltage values in (a) and (b) corresponds to standard test 4 and 5 in Appendix C.

and dissipation with good characteristics without interference from unwanted stray materials.

## 5.4 SRIM ion-implantation simulation

A well-known simulation program, SRIM, was used to estimate the depth of implantation of various atoms into the device during the FIB milling step. SRIM stands for "Stopping and Range of Ions in Matter" and has been periodically updated since 1983, with the last update in 2013. This is a Monte Carlo simulation which allows the distribution of implantation atoms in a multilayer planar device to be calculated when given implantation atoms of a specific element and a specific energy.

SRIM simulations are limited to implantation on planar devices (see Fig. 5-5a), but we can use it to first determine the implantation depth of the key elements present (Ga and Au) into the different materials in our device (Au, Al2O3 and WN) as shown in Fig 5-5b. The result are shown for Ga and Au ion beams of 30keV since FIB milling on our devices was done using a Ga-ion source at a 30kV acceleration. These ions can penetrate between 5 to 40nm deep into each isolated material. By comparison, the

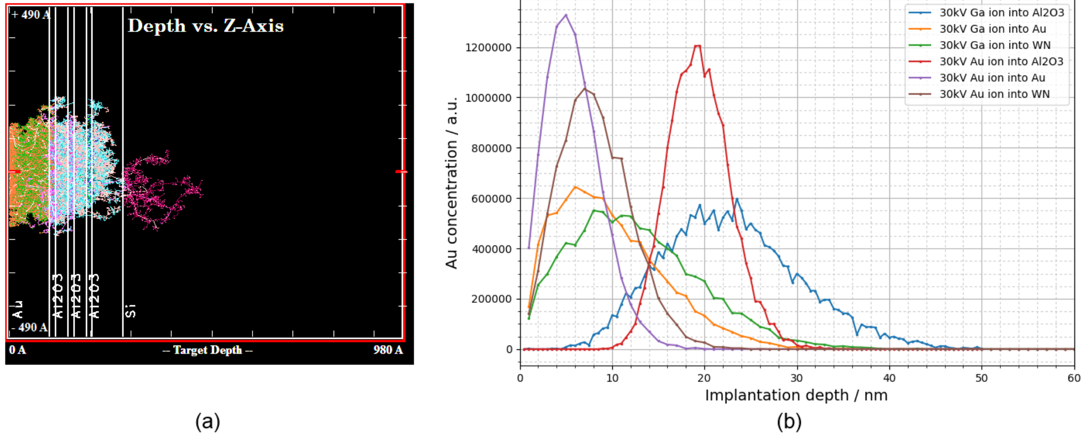



Figure 5-5: (a) Screenshot from a typical Monte-Carlo SRIM simulation showing implantation of 30kV Au dopants from the left edge into a multilayer device stack. The traces shows displacement of the original device atoms and also stopped Au ions. (b) Profile showing the implantation depths of various ions into a variety of bulk materials. In these simulations, only a single material was specified unlike the multilayer geometry shown in (a) so as to first understand penetration depths in the various materials present in our device.

combined triple-layer oxide (3x Al2O3 and 2xWN layers) stack only 11nm thick.

SRIM was also used to provide guidance on substitution ion sources. A 30kV Ga ion source was used in our devices to implant Au atoms from the surface of the device deeper into the oxide stack. Instead of Ga, other heavy mass ions such as Ar were also suitable for this implantation. On the other hand, He ion sources were confirmed not to be suitable for activating RS devices because the light mass He ions have low momentum and are not able to implant Au deeper into the device. A He ion beam was also experimentally tested at 15kV or 30kV and at dosages of 85pC up to 6.6$\mu$C for milling into device before later deposition of a Au layer. The He ion did not produce any noticeable change to the device, which is to say that these devices required electroforming and did not switch well after He ion implantation. This either implies that He ion are ineffective at ion implantation, or that the presence of a Au layer for the implantation is required.

The accuracy of the SRIM simulation can be checked by comparing the calculated dosage required to mill away the deposited layers to reach the Si substrate. Experimentally, the dosage used was determined by testing various dosages and using SEM

imaging to find a large enough dosage to expose the underlying Si substrate. The SRIM simulation can calculate the sputter yields but is not able to automatically update the device geometry to account for sputtering material losses after each step, so secondary calculations have to be made manually to derive this suitable dosage. The simulations yield a sputter dosage requirement of 0.529pC/$\mu$m$^2$ for each unit nm of Au, 2.07pC/$\mu$m$^2$ for Al2O3 and 1.02pC/$\mu$m$^2$ for WN. This sums to an simulated dosage of 50pC/$\mu$m$^2$ to mill away the deposited layers and is close to the dosage found experimentally of about 50 to 100 pC/$\mu$m$^2$.

However, the implantation distribution of Au cannot trivially be determined using SRIM because the geometry of the device is not updated as milling proceeds. The complication for our device is that Au atoms is not just sputtered downwards from the top electrode into the oxide layers, but can be sputtered out from the outer perimeter of the milled trench and implanted sidewalls into the inner perimeter along the actual device mesa. The sideways implantation can be experimentally verified by setting the ion beam to mill with a zig-zag fashion either outwards or inwards towards the device mesa. Changing the direction to mill outwards causes more materials to be ejected from the outer edge being milled and implanted into the inner stationary edge. Measurements of the as-fabricated resistance of this device shows that it has a higher initial conductivity as compared to a device milled with a zig-zag fashion going inwards (see Fig. 5-6). The higher conductivity is expected from the additional concentration of implanted atom under this milling direction.

Nonetheless, the post-milling distribution of Au can be crudely estimated by combining the sputter yield of Au from the 30kV Ga ion beam and the 2D distribution of implantation of Au atoms at different various energies and incident angles into single materials. The lateral FWHM of downwards implantations is about 1nm to 4nm from the edge of the device perimeter, with approximately 0.2 Au atoms to be found for each original atom in Al2O3. The lateral FWHM from sideways implantation will follow a distribution similar to Fig. 5-5b and will be implanted deeper of more than 1$\mu$m deep but with a lower implanted concentration. However, even if SRIM can modified to provide an accurate as-fabrication distribution of Au, this would still

| | Milling sequence | Current at 2V / mA |
|---|---|---|
| 1 | 30kV parallel, 1µm | 32, 32 |
| 2 | 30kV inwards, 1µm | 21 |
| 3 | 30kV outwards, 1µm | 53, 53 |
| 4 | 5kV outwards, 1µm | 3.2 |
| 5 | 5kV inwards, 1µm | 2.0 |
| 6 | 30kV outwards 4µm, then 5kV inwards 2µm, with 1µm overlap | 7.4 |
| 7 | 30kV inwards 4µm | 18 |

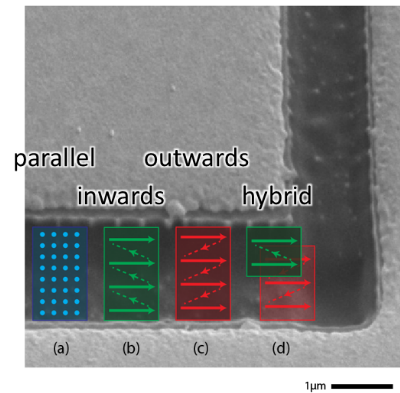



Figure 5-6: Effects from FIB milling direction on device properties. The SEM image shows a trench milled to create a mesa to define the actual device footprint. The top-left corner belongs to the device mesa. Comparing milling sequence 1,2,3, an outwards milling strategy creates a device with a lower resistance (higher current). This can be attributed to a larger amount of materials implanted sideways into the device sidewalls. Comparison of milling sequences (4,5) and sequences (1,2,3) shows that a higher FIB milling voltage gives strong filaments, implying that more materials are implanted successful at a higher voltage. Sequence 6 attempts to use a lower voltage milling to clean up the device sidewall with a gentle 5kV inwards milling after heavy implantation with a 30kV acceleration. Comparing sequence 6 and sequences (4,7) shows that filament are embedded more than 1$\mu$m deep.

be insufficient to relate models to experiments because an initial voltage sweep will remove an undetermined amount of implanted materials from the device (as in Fig 5-4). An accurate determination of the steady state distribution of Au dopants will require both an extension to the SRIM program and an a-priori model to handle electromigration. Unfortunately, it is still an ongoing effort in this community to develop a robust model for simulating electromigration effects which can hold across different material systems and devices.

## 5.5 Network-level simulation of filaments

Although many questions still remain in this field with regards to the mechanism of switching, we can treat the migration problem of each filament as a black box and see if we can derive higher-order models for explaining the benefits of multilayer oxides versus that of single layer oxides. In the case that we have a physical web of interconnected filaments in each oxide layer, we can still approximate the web as a layer of filament sites in parallel.

The hypothesis with multilayer oxides is that only one of the oxide layer will have filaments which switches while other oxide layers remain passive[63]. We find from simulations that this arrangement leads to an I-V behavior with a gradual SET and an abrupt RESET which is characteristic of our device.

The model simulates a two oxide layer device. One of the oxide layer is kept as a passive layer and defined with a number of filament sites, $N_p$, all fixed to be in a low resistance state (LRS). The other oxide layer is defined as the active layer and defined with another number of filament sites, $N_a$, which can either be in a low or high resistance state (HRS). We shall later verify the correctness of simplifying the model to keep all filaments sites in the passive layer to be fixed in the LRS.

In the model, each filament site hosts a potential filament which can either be in the LRS or HRS (see Fig. 5-7). The I-V responses of a single filament in the LRS and

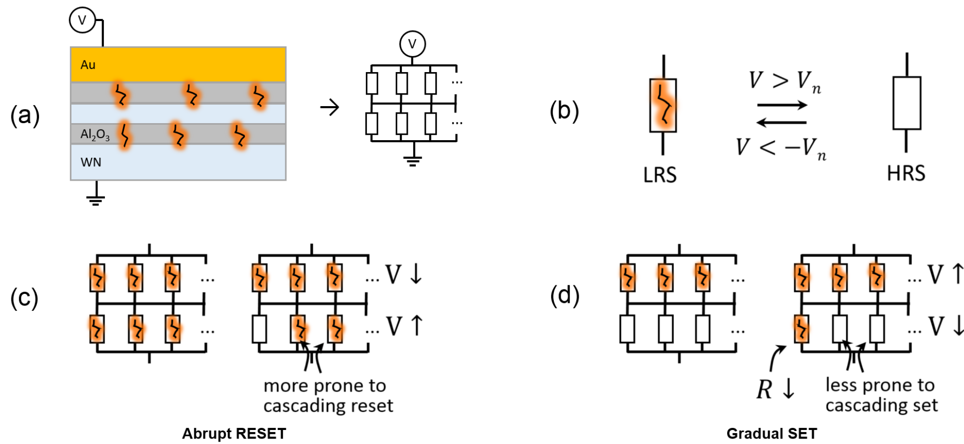



Figure 5-7: Schematic for network-level simulation of a multi-layer oxide structure.
(a) A 2-layer oxide device was simulated. Each oxide layer was simulated to host $N$ number of filament sites, where each filament site can either host 1 or 0 filament corresponding to either the LRS or HRS. This network of filaments and their respective resistances can be converted into a network of resistors.
(b) Each filament site can be switched between the LRS or HRS based on whether the voltage across that filament site is above or below some preset threshold. The simulation randomizes each filament site $n \in \{1, 2, 3, ..., N\}$ to have a fixed threshold $V_n \in \{V_1, V_2, V_3, ..., V_N\}$ drawn from a random distribution.
(c) The top layer is assumed to have all filament sites fixed in the LRS. When the bottom layer starts to RESET, the total resistance of the bottom layer increases. This causes the voltage across the bottom layer to increase and creates a cascade of RESET actions. This creates an abrupt RESET transition.
(d) When the bottom layer starts to SET, the combined resistance of the bottom layer decreases and the voltage across the bottom layer drops. The applied voltage must be strengthened to compensate for this voltage drop before more filament sites can be SET. This creates a gradual SET transition.

HRS were fitted empirically to experimental data to be of the form,

$$I_{\text{LRS}} = A_o V + A_h \ e^{3.8\sqrt{V}}$$

$$I_{\text{HRS}} = A_h \ e^{3.8\sqrt{V}}$$

, where the non-linear dependence on voltage, $e^{k\sqrt{V}}$ is taken from a Poole-Frenkel hopping conduction model (eq. 2.2) with $k = 3.8$ is obtained empirically for our devices (see Fig. 5-8a). The non-linear voltage term in $I_{\text{LRS}}$ was needed to fit the deviation away from a perfect ohmic behavior at higher voltages of >2V. On a side note, this suggests hopping conduction is present even when a device is in the LRS, potentially from stray filamentary materials or defect sites that exists in addition to the formed filament. A typical fit on I-V data yields $A_{\text{LRS},o} = 2.2c \times 10^{-3}$ and $A_{\text{LRS},h} = 2c \times 10^{-7}$, where $c$ is a scaling factor that is determined by the cross-sectional area of each filament and the total number of filament in the actual device. For simulation, $c$ was taken to be 0.005.

Each one of $N_a$ active filament site was randomly assigned a transition voltage $V_1, V_2, ..., V_{N_a}$ from a random normal distribution centered at $V = 3.5V$ and a standard deviation of $0.8V$. This transition voltage $V_n$ remains constant through the simulation and is used to determine whether the filament will SET or RESET. When the voltage across filament site $n$ exceed $V_n$, the filament will RESET. When the voltage across filament site $n$ is smaller than $-V_n$, the filament will SET. The simulation starts with all filament sites in the LRS and incrementally varies the applied voltage across the device from zero to a negative voltage, then to a positive voltage and back to zero. At any step of the simulation, the voltage across each filament site can be evaluated by treating all the filament sites as a network of resistors and checking whether the filament sites would SET or RESET.

This bare model is sufficient to show that a device with 2 or more oxide layer can have an abrupt RESET and a gradual SET (see Fig. 5-8b,c) despite each filament site having a threshold voltage with a large spread (recall that $V_n \sim \text{Gaussian}(\mu = 3.5V, \sigma = 0.8V)$). A closer analysis of the distribution of voltage across each filament

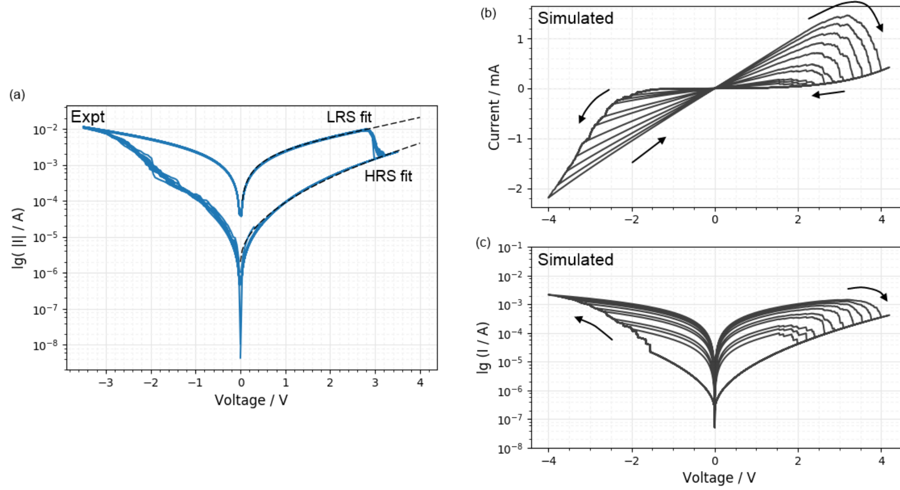



Figure 5-8: (a) Shows the form of the LRS and HRS (dashed lines) used in the simulation as fitted to experimental I-V plots (blue lines). The LRS fits well to an Ohmic behavior while the HRS fits well to a Poole-Frenkel hopping conduction mechanism. (b) and (c) shows the linear and log plots of the voltage sweeps from the simulation. The simulation shows a gradual SET and a relatively abrupt RESET similar to what is observed experimentally.

shows that when a single filament RESETs, the overall resistance in that active layer increases, causing the voltage experienced by other filaments in the same layer to increase by the voltage divider rule, thus triggering a sharp cascade of RESETs. On the other hand, SET is smeared out to take place over a wide voltage because a filament site which has SET would have a lower resistance. This causes the voltage across the active layer to be lowered, and subsequently the applied voltage have to be increased first to compensate for this drop, before increasing further to trip the next filament site with the next smallest transition voltage to SET.

We can also now check our initial assumption that the passive layer does not switch. Referring to Fig. 5-8c for RESET, we see that the already-formed filaments in the passive layer will see decreasing positive voltages when filaments in active layer breaks. This means that with each additional filament which is broken in the active layer, the filament in the passive layer is further protected from RESET, and thus the filaments in the passive layer stay unbroken. In the SET phase (in Fig. 5-8), there will be increasingly negative voltages across the passive layer with each filament in the active layer which forms, but there are no changes in the passive layer since all

the filament sites in that layer already house formed filaments. These checks validate our initial assumption that the passive layer can be assumed to be passive.

If desired, heating effects, separate transition voltages for SET or RESET, or other secondary effects can be introduced to better fit the results from this model to experimental data, but at the expense of losing prediction power from overfitting when additional parameters are available to be tweaked. Our model is already capable of fitting the general I-V behavior using few variables.

This chain effect at RESET and moderating effect at SET shows how a change from a single to multilayer device can reduce the dependence of the device's behavior on variable fabrication conditions and instead depend on the controllable device stack-up to give predictable resistive switching cycles.

# Chapter 6

# Pulsed measurements

We had shown novel strategies (chapter 4) in fabricating resistive switching (RS) chips to guarantee near-perfect device yields and consistent switching cycles that holds from cycle-to-cycle and device-to-device. In this chapter, we first show that FIB implantation has no detrimental effect to pulsed mode operation, and further show how pulsed mode measurements can be used to clearly identify RS devices with high operating consistencies. We also find that the endurance of a device under pulsed mode depends strongly on the pulsing parameters, implying that the future control circuitry for ReRAM arrays will have an important role both in optimizing the switching success and extending the life expectancy of individual ReRAM devices. In addition, we explain how new insights with regards to the mechanism of RS devices can be derived with closer analysis on pulsed mode behavior. The results showed here will help both in fundamental modeling of RS devices and in future optimal operation of ReRAM devices.

## 6.1 Introduction

The maximum operating frequency is an important specification for any digital integrated chip since it specifies the fastest speed the chip is rated to operate beyond which there is risk of malformed bits or data loss[64]. The operating speed is a merit which extends beyond digital ICs and into analog ICs such as op-amps[65], where

operating beyond some rated frequency will lead to signal distortion and less-than-unity gains. Resistive switching (RS) devices will thus have to be evaluated for its switching speed just like any other electronic chips destined for use in any applications. Specifically, we have to verify that the novel ion implantation techniques which we use for our devices are providing the high switching consistency and device yields not at the expense of switching speed. This will require a switching speed somewhere in the range of 10ns to 1$\mu$s as reported by other devices in literature[66, 67, 68, 69, 70, 71].

The spread of RS devices reported in literature varies significantly for each of its performance metric. For the switching consistency metric, as mentioned previously in chapter 2, there is no single adopted methodology to quantify the consistency of a RS device. The metric for resistive switching devices in the realm of pulsed mode operation is even much less properly defined because several experimental challenges has slowed down the development of a common experimental approach[72, 68]. The first experimental limitation comes from the devide side. Pulsed mode measurements are only worth performing on devices which already demonstrates a good amount of consistency in basic I-V measurements, or else introducing another experimental variable (the pulse properties) will just add more stochasticity to the data and hinder interpretation. Secondly, additional hardware is required for generating well-defined sub-1$\mu$s pulses, while special attention to impedance matching and high speed signal propagation would be required for even shorter sub-10ns pulse. Even when these two factors are present, there remains a lack of an agreed procedure on how to use pulsed mode measurements to evaluate the RS device to gain deep insights and for easy comparison to other reported devices. We will address this issue by proposing a method to extend existing pulsed measurement schemes and by showing how it can be used to derive more insights into the behavior of an RS device.

## 6.2 Experimental method

A Tektronix AFG3052C arbitrary function generator was used as the pulse generator. This function generator is limited by a rise and fall time of 7ns each and a shortest

pulse width of 12ns. A computer controller randomly chooses the voltage amplitude and pulse width for each pulse sent by the function generator to the RS device under test (see Fig. 6-1a). The magnitude of the voltage is chosen randomly from a uniform distribution whereas the pulse width is chosen from a uniform distribution on its logarithmic value. Both parameters are bounded to some predetermined range chosen to avoid unnecessarily large values that may shorten the life of the device. This scheme is similar to other set-ups used in literature[73, 66, 67, 74, 75] with a popular choice for the pulse generator being the Agilent 81110A with a faster 2ns transition time and a shortest pulse width of 3ns.

The function generator has a floating output with an output impedance of 50Ω. This impedance oughts to be matched in the transmission cable and the RS device to avoid signal reflection. Impedance-controlled coaxial cables were used for pulse transmission between all devices. Ideally, the source impedance of the pulse generator should match the device impedance as well, but this was too tedious to implement because the function generator has a fixed output impedance of 50Ω and the device impedance will vary as it switches between resistive states. This means that any impedance matching circuit created would have to be an dynamic one to match the 50Ω transmission line to the approximately 1kΩ to 50kΩ ever-changing device impedance if perfect impedance matching is desired. For simplicity, such a matching circuit was omitted and this was later identified not to be a limited factor for our particular case anyway. The shortest pulses of about 50ns which can be transmitted by this limited arrangement without significant signal distortions to the RS device is faster than the switching speed of our RS device and is thus not a limiting factor in characterizing the performance speed of our devices. The integrity of the pulse transmission was verified by connecting an oscilloscope (Agilent MSO-X 2024A, 200MHz, 2GSa/s) at a 1MΩ input impedance across the RS chip to view the waveform of the pulse as observed by the RS device.

A Keithley 2450 sourcemeter was used to determine the resistive state of the RS device after each pulse to check if the previous pulse was successful in either SET a RESET transition. If a SET transition has occurred, subsequent pulses will be of a

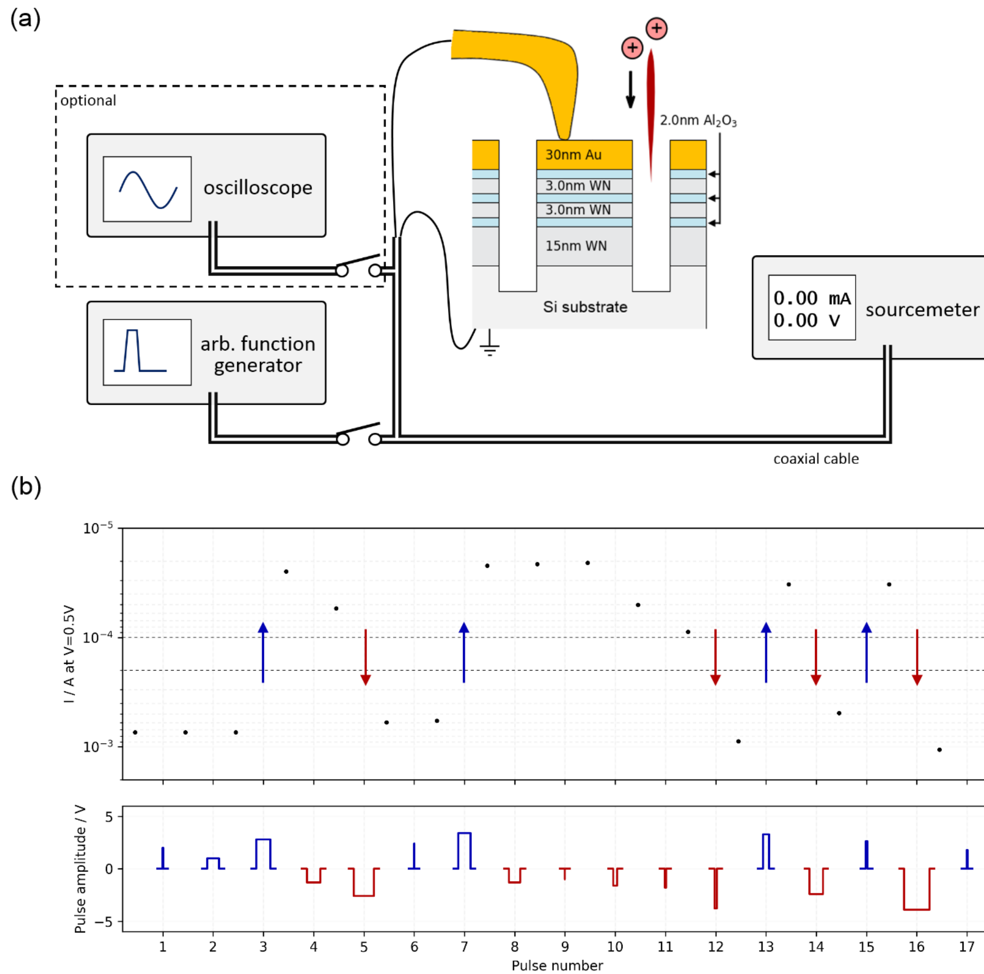



Figure 6-1: (a) Schematic for pulsed mode measurement on a resistive switching device. The illustrated device is a typical configuration for a highly consistent switching device fabricated with multilayer thin films and with focused ion beam (FIB) activation. An arbitrary function device generates >12ns pulses of either polarity and with variable amplitude. A sourcemeter determines the resistive state of the device between pulses and an oscilloscope is used to verify the integrity of the short pulse. (b) Pulses of random width and amplitude were sent to the device to attempt SET and RESET. A successful transition is verified with the sourcemeter to be when the measured electrical current at a test voltage of 0.5V crosses two predefined threshold current (shown by the two dotted lines). When the measured electrical current falls between the two threshold current, pulses of random polarity is sent to get the device out of this ambiguous resistance state.

positive polarity in order to attempt a RESET transition, and vice versa (see Fig. 6-1b). The resistive state was determined by measuring the device current at some small test voltage of 500mV and setting 2 thresholds for identifying the state to be the LRS or HRS. If the device current falls in between these 2 thresholds, the resistive state of the device is taken to be ambiguous and pulses of random polarity will thus be sent out to force the device into a more well-defined resistive state. These two thresholds vary with the stack configuration of the device-under-test, but a typical value pair would be 0.1mA and 0.2mA, good to tag almost all states as HRS or LRS given the large on-off ratio for our devices of between 10 to 100. When the sourcemeter is sampling the RS device, the function generator is electronically disconnected so that the measured current comes from the applied voltage across the RS device rather than from the 50Ω output impedance element of the function generator.

We tested two classes of devices under pulsed mode operation, the first being the typical FIB-processed triple oxide layer stack and the second where the Au atoms are co-sputtered during a reactive sputtering process.

## 6.3 Results

### 6.3.1 Pulsed operation behavior

The SET and RESET attempts were binned by pulse width and amplitude and the ratio of these attempts causing a success transition is plotted in Fig. 6-2 for each of two types of devices tested. The results immediately show that the switching speed is limited to 0.5$\mu$s, comparable to other RS devices reported in literature but limited in our case due to the larger device area not optimized for speed [69, 70, 71]. This is due to a large RC charging time constant corresponding to our 50$\mu$m × 50$\mu$m device, so a longer pulse is required for a switching. The need to reduce the device footprint in order to achieve faster switching has been described elsewhere[71].

Most pulsed mode measurements in literature have been performed by varying only either the pulse amplitude or the width separately. However, there are two

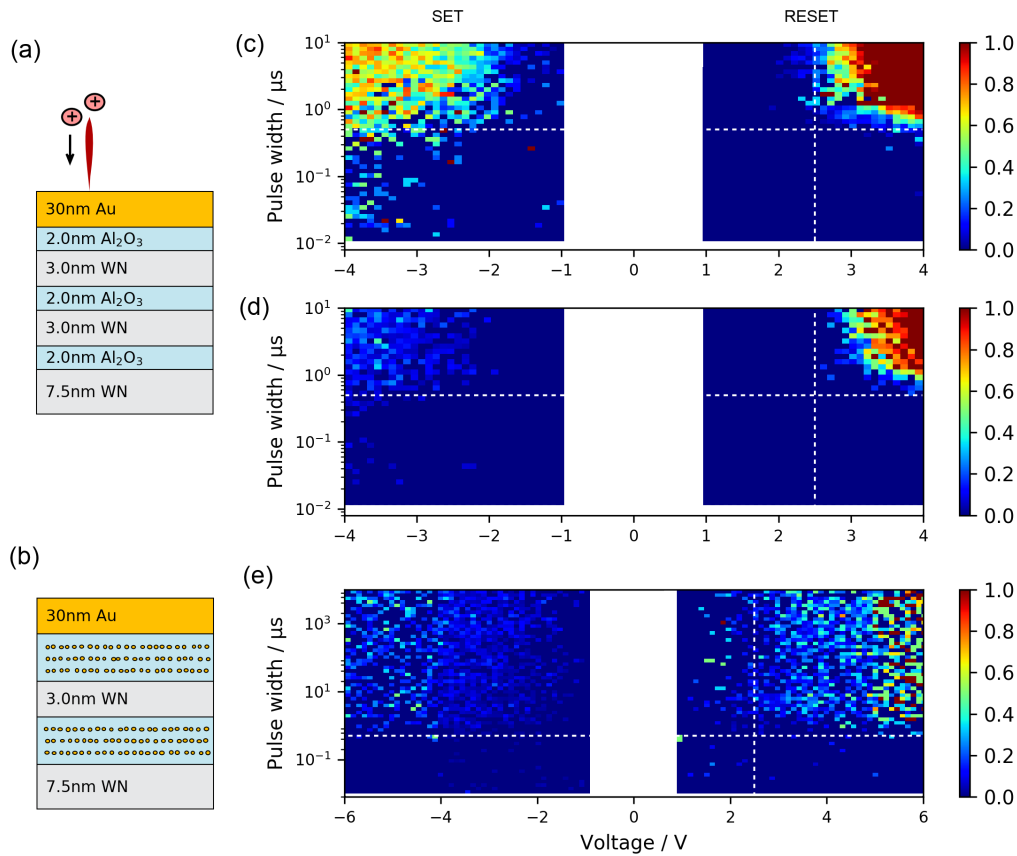



Figure 6-2: The pulsed mode behavior of two varieties of devices was measured. The first type of device is a FIB-activated RS device shown by (a), with pulsed characters shown in (c) and (d). This device was previously found to have highly consistent I-V switching cycles with an abrupt RESET for V>0. Plots (c),(d),(e) show the success in which a random pulse of the indicated voltage and pulse width caused a successful transition. The consistency of the device in (a) is evident in the V>0 region of (c) where there exists a time-voltage boundary beyond which random positive pulse will essentially guarantee a RESET. (d) tabulates switching attempts 1 to 25000 while (d) tabulates switching attempts 55001 to 80000. This shows that cycling a device causes the device to age from an increase in the voltage required for switching. Dotted lines are shown at +2.5V and 0.5$\mu$s as visual guides. The second type of device is where Au is co-sputtered along with reactive sputtered Al2O3 as shown in (b), with pulsed mode results shown in (e). The inferior switching consistency of this device shows up as an upper limit to the RESET success rate shown in the V>0 region.

benefits to presenting pulsed mode results as shown in Fig. 6-2. The first advantage is that key experimental procedure and range of experimental values (pulse width and amplitude) is immediately explained by the range of the plot while the plot values themselves demonstrate the characteristic of the device. This is very useful for the field of resistive switching devices where there is a wide variety of material systems and potentially different switching mechanisms, so distillation of results into condensed figures is essential for easy comparison to be made between different reported devices. Such a plot will help reduce the verbosity of experimental procedures in the future that accompanies pulsed mode measurement data, because the plot itself already implies that a range (rather than a single) of pulse width and amplitudes were tested and verification of a successful transition was performed after pulse. This will help standardise customized pulsed mode testing results currently undertaken by RS researchers.

The more significant advantage is that by tabulating pulse mode measurement attempts in Fig. 6-2, we can extract further insights or better quantify the switching character of a RS device which can help shed light in the future when deriving a full mechanistic understanding of the switching process. We will describe these in the next few paragaphs.

The FIB-activated RS device has earlier been shown to have a high consistently switching behavior, whereas the Au co-sputtered RS device is a promising device which helps to prove our mechanistic understanding of the former device but not yet optimized to have as good a switching consistency as the FIB-activated device. The $V > 0$ right side of Fig. 6-2 shows that the RESET attempts are divided distinctly into two different regions. RESET attempts of a sufficient large pulse width and pulse amplitude have an almost certain probability of causing a successful transition. This boundary is about 5$\mu$s at 3.2V, but is shorter at 1$\mu$s at a higher voltage of 4V. Such a relation has previously been identified although presented less visually[68, 72]. The trade-off relation between the pulse width and amplitude will guide future investigations into the switching mechanism. For example, not having a vertical boundary with a voltage threshold separating successful and unsuccessful

transitions provide support for time-dependent switching mechanisms such as heat induced filament breakage and undermines less time-dependent switching mechanisms such as voltage-only induced electromigration.

A comparison between the RESET probability distribution for the FIB-activated RS device and the Au co-sputtered device shows that differences in these two devices can be quantified by the character of their RESET success rates (compare Fig. 6-2a and 6-2c). For the Au co-sputtered device seen in Fig. 6-2b&e, the application of a large voltage of more than +5.5V and pulse width more than 100$\mu$s does not guarantee a successful RESET and has a stochastic behavior just like the SET behaviors in the V<0 regions of Fig. 6-2a, b, c. Stochastic behavior for RS devices are commonly presented in literature through cumulative distribution functions (CDFs) of the transition voltage or the resistance states. The switching success probability plots shown here go a step further in illustrating consistency by presenting experimental results across multiple dimensions, with differences easily visible for highly consistent switching as we have shown in Fig 6-2c vs 6-2e.

The FIB-processed RS device has an I-V curve (shown in previous chapters) which already hints at differences in SET and RESET where SET has a gentle transition and RESET has an abrupt transition. The pulsed success probability plot more clearly shows the differences between these two transitions. SET transitions are more stochastic, has a lower chance of success, and no sharp boundary beyond which a successful SET transition is guaranteed. This is evident in the V<0 region of Fig. 6-2 where each 2D bin on the plot has only a 30% chance of a successful transition for sufficiently negative voltages and sufficient long pulses, and that this probability does not approach unity even with more negative voltages or longer pulses. This is again an significant observation that will guide mechanistic studies. This distribution of probabilities suggest that the SET mechanism is governed not by long-term time dependence such as heating effects, because even an increase in the pulse width by 10x from 1$\mu$s to 10$\mu$s does not lead to a similar increase in the probability of switching success, but has this probability capped at about 30% per pulse. This implies that the SET mechanism might have a stronger dependence on on time derivative values

rather than on time directly. For example, this result would support theories where ions being accelerated by an electric field have a hopping success rate across the lattice which varies with the magnitude of acceleration. This can be verified in future experiments by tabulating the SET success probabilities with respect to the pulse slew rate (a.k.a transition time) rather than the pulse width and with an expectation that success rates will increased with faster slew rates. This might also be explained by migration mechanisms which depend on diffusion rather than drift for filament formation. In additional to fundamental understanding, this result will also guide the design for controller circuit of RRAM arrays. For example, if pulse width is indeed inconsequential for a SET transition, then the optimal strategy for SET will be to send a train of short but sharp pulses to a RRAM device to achieve the highest success rate of SET in the shortest possible time.

The granularity of information in the switching success probability plots also makes the plot useful for tracking degradation or aging effects in the RS device. Fig. 6-2a tabulates the switching success from the first 40000 pulsed switching attempts on a device while Fig. 6-2b shows the next 40000 pulsed attempts. The degradation of the device manifests as a uniform decrease in the SET success rates and a voltage up-shift in the threshold for successful RESET transitions. The decrease in the SET success rate imply that the ease of filament formation or vacancy migration in a lattice becomes more difficult with cumulative switches though the exact mechanism is still open for debate. The changes to the RESET characterization is more interesting for analysis. It is intuitive to expect that a heavily-cycled RS device undergoes some form of annealing which causes subsequent switching to become more difficult. However, the fact that this increase in difficulty shows up as an increase in the voltage requirements rather than in the pulse width requirements provide further hints at the changes due to heavy cycling. This implies that the RESET mechanism for our RS device has a fairly constant transient time of approximately $0.5\mu$s before reaching some steady state, such as the steady state temperature profile around a filament heated by Joule dissipation. This will be useful for empirically extracting time constants to be used in models for filament migration.

### 6.3.2 Device endurance

The pulsed mode switching experiments naturally leads to an evaluation about the endurance of our RS devices. The magnitude of both the SET and RESET pulse amplitude are nominally restricted to the range of 1 to 4V at the beginning. The cumulative switching success count and switching success rate were tracked and are plotted in Fig. 6-3. For a given specific range of pulse parameters, the switching success of the device decreases with cumulative switching attempts such that each device will have a limited endurance where the cumulative success count will plateau at some value. At first look, this is convenient because the endurance of a device can perhaps be quantified by this plateau value. However, we identify several shortcomings in the use of this metric. Firstly, the endurance of a device depends on the range of pulse amplitudes it has been subjected to. The use of higher voltage pulses has a higher tendency of decreasing the lifespan of a device and leads to a lower plateau height. This means that the plateau height will not be valid for comparing one device against another reported device because this height is strongly dependent on the experimental procedure. In other words, the endurance of a device can be inflated by repeated applications of the lowest voltage pulse able to cause switching. Such pulses will give a low switching success rate (as can be obtained from a plot like Fig. 6-2) but give a higher cumulative switching success count. The dependence of the device endurance on the switching parameters has already been reported previously under a more restricted set of experimental procedures.

Another issue limits the simple quantification of the endurance of a RS device. We find that there exists conditioning processes which can extend the life of a seemingly dead device. This is demonstrated by the several smaller increases beyond the initial plateau in switching success in Fig. 6-3. These additional rounds of switching came about after the random pulsed experiments was put on hold and a set of I-V curves were collected from the device under test. The I-V curve consisted of 60 rounds of I-V sweeps of increasing voltage ranges, but kept within to -4 to 4V, taking place over a 19 minute period. Since the voltage amplitudes seen in this I-V measurement set

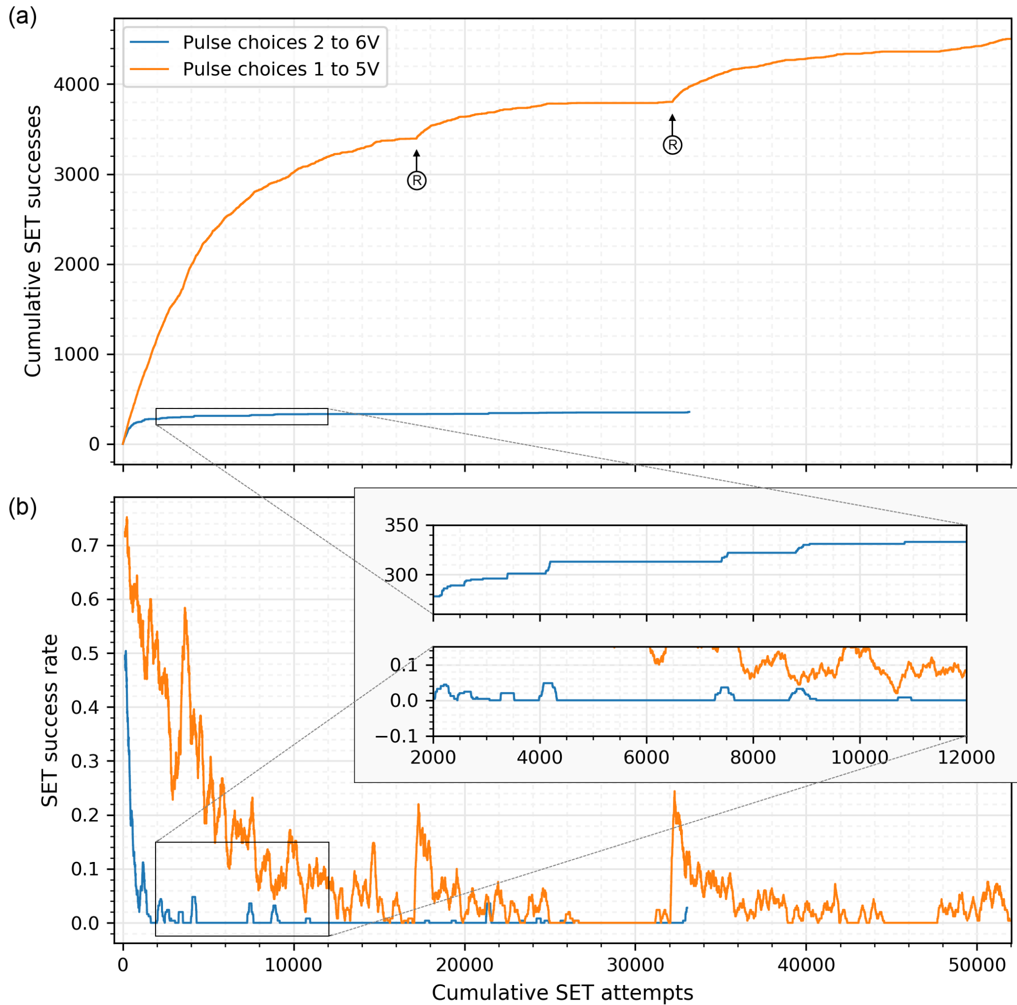



Figure 6-3: Subjecting RS devices to a continuous sequence of random pulses leads to a decline in switching rate over time and the number of successful transitions eventually reaches a plateau as shown in the top plot. The bottom plot shows the instantaneous switching success, averaged over a symmetric window of 250 switching attempts. The exposure of a device to larger voltage ranges (blue line) tends to give a lower device lifespan. Reconditioning cycles, indicated by the two circled R, can increase the lifespan of a seemingly dead device. Also, the switching success is not completely randomly but appears to be dependent on some hidden state of the RS device. This is shown in the inset as bouts of easy switching separated by periods where switching was practically not observable. RESET transitions follow a similar behaviour since each SET will be followed by a RESET.

is no higher than the pulse amplitude applied during the random pulsed test, it is likely that saturating the device over a prolonged period can help in reconditioning the device and in extracting more switching cycles before permanent deactivation of switching sets in. The existence of reconditioning protocols is good news for future use of RS devices in actual applications but with a downside of further complicating the standardization of procedures to evaluate the endurance of a device. Without an agreed set of procedures to evaluate device endurance, it will not be very meaningful to compare a quoted value from one device to another reported device since such a value can vary about an order of magnitude simply by small variations in experimental procedures.

A closer look at the successful cumulative switches yields another insight that will be useful for elucidating the switching mechanism. The inset of Fig. 6-3a shows the success of each switching is not independent but is rather a state variable which is correlated with recent history of successful switches. This is supported by the observation that increases in the cumulative switching success comes in bouts of intense activity rather than as a smoothly increasing function with time. This trend shows that switching success is not just dependent on the applied pulse voltage and width, but further as a state function of the device, likely to be a function of the physical lattice configuration of the switchable layer.

## 6.4 Further discussion

Beyond the identification of insights into RS behaviors and observations which will guide future attempts at a more complete mechanistic description of these processes, there are some further implications from our results. We have showed that pulsed mode measurement is great for a deeper characterization of RS devices because switching depends on pulse parameters and that this dependence is both a good indication for the switching performance of the devices and is rich with information useful for corroborating with proposed mechanisms or to be used in numerical models. However, there remains an issue in quantifying the endurance of a device because the limit of successful switching cycles of a device depends heavily on the experimental procedures

used. Since this limit will depend on whether re-conditioning procedures are used and on the intensity of the pulses used, there will need to be some agreed procedure to quantify this endurance. In the absence of a common accepted procedure, the next best option is for published reports to include details of the endurance test beyond the simple citation of a numerical figure. At the least, this will need to include the pulse width and amplitude used, and if re-conditioning of the device was done. On the other hand, a standardized procedure might for example be defined to use adaptive pulse applications determined to give some fixed rate of switching success over time. The applied pulse amplitude will then drift upwards as the RS device becomes worn, and the cumulative number of switching success when this pulse amplitude exceeds some large pre-defined voltage can then be taken as the endurance of the device.

This test protocol is flexible in accommodating the large variance in the transition voltages reported in RS devices since this procedure does not require a common pulse voltage to be used for comparisons, and that the adaptive pulsing scheme can be used to quantitatively equalize the varying efforts to which a RS device is tested before it is declared dead. At this preliminary stage, it is hard to tell whether it is more advantageous to operate RS devices in RRAM arrays to have a higher probability of success per switching attempt at the expense of shorter life using stronger pulse parameters, or vice versa to optimize for lifespan by using weaker pulse parameters. This choice is probably going to affected by the intended application of these RS devices of whether speed or lifespan is more desirable, but hopefully future studies will evaluate the relative advantages of operating RS devices with tuned pulses of for example, 1%, 50% or 99% switching success rates.

## 6.5 Conclusion

We show that pulsed mode measurements are valuable for adding another dimension by which RS devices can be characterized beyond that of basic I-V measurements. The variation of switching success probability across pulse width and amplitude is useful not just to further quantify the switching consistency, but can provide supporting evidence for further work when constructing a rigorous mechanistic model to describe

SET and RESET transitions.

We have also identified several dependence of pulse parameters and experimental procedures on the endurance of a RS device. This implies that future evaluation of endurance will have to be done under properly defined procedures so that reported numerical values between RS devices can be meaningfully compared to one another. Nonetheless, this is good news with regards to future applications of such devices because this implies that there will exist many operating strategies, for example by choice of gentler pulsing parameters or by using re-conditioning cycles, so that each RS device can reach the maximum operating lifespan possible.

# Chapter 7

# Future directions

We have demonstrated strategies to obtain near perfect device yields and highly consistency switching cycles that holds across cycles and across RS devices. However, plenty of work remains in this field and our work has only opened up more exciting possibilities for future exploration.

These new directions can broadly be classified into two classes. The first class entails a deeper look into the mechanism of switching now that it is possible to fabricate high-consistency devices where erratic switching behavior no longer hinders analysis. The second class entails future directions into how RS devices can be designed with the additional degrees of design freedom enabled by our work.

## 7.1 Analysis of second order effects in I-V curves

In chapter 4, we have only considered general trends and switching consistency in all the experimentally obtained I-V curves with a focus to support our hypothesis that implanted Au dopant atoms are responsible for highly consistent switching. In section 5.3, we briefly used the transient behavior in the first I-V sweep of an as-fabricated device to deduce that FIB-processing implants excess stray materials into the device stack which is removed in this first sweep. Similar analysis can be made on subtle but yet repeatable features of electrical measurement results to derive future insights into the switching mechanism.

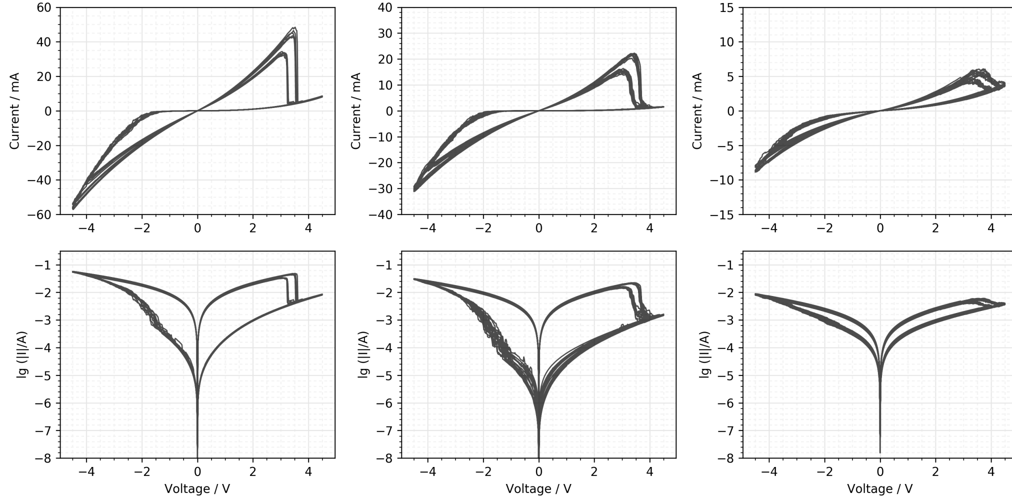



Figure 7-1: (a) shows the I-V curve from a fresh device cycled with the voltage waypoints of 0,-4.5,4.5,0,-3.5,3.5,0V. There are thus two visible RESET thresholds at 3.3V and 3.6V, but both transitions are abrupt. (b) shows that the resistance of both the LRS and HRS increases when the device is cycled beyond 200 times, but with the switching transition also becoming less abrupt. This presents one additional of clue for future model fitting. (c) With more cycling, the resistance of the device LRS further increases.

For example, fabricated RS devices would possess RESET transitions of variable abruptness in the different stages of their lifespan (see Fig. 7-1). This is useful for building more accurate RS models in the future because any worthy model oughts to also explain these experimental observations. The decrease in RESET abruptness can partially be explained by the network-level simulation model used in sect. 5.5 where continuously cycling could cause Au dopant atoms to migrate out of the bulk Al2O3 and can no longer help to form a filament. This RESET abruptness is also reduced when an excess ion milling dosage of 500pC/$\mu$m$^2$ is used instead of a more typical value of 50pC/$\mu$m$^2$ (see Fig. 7-2).

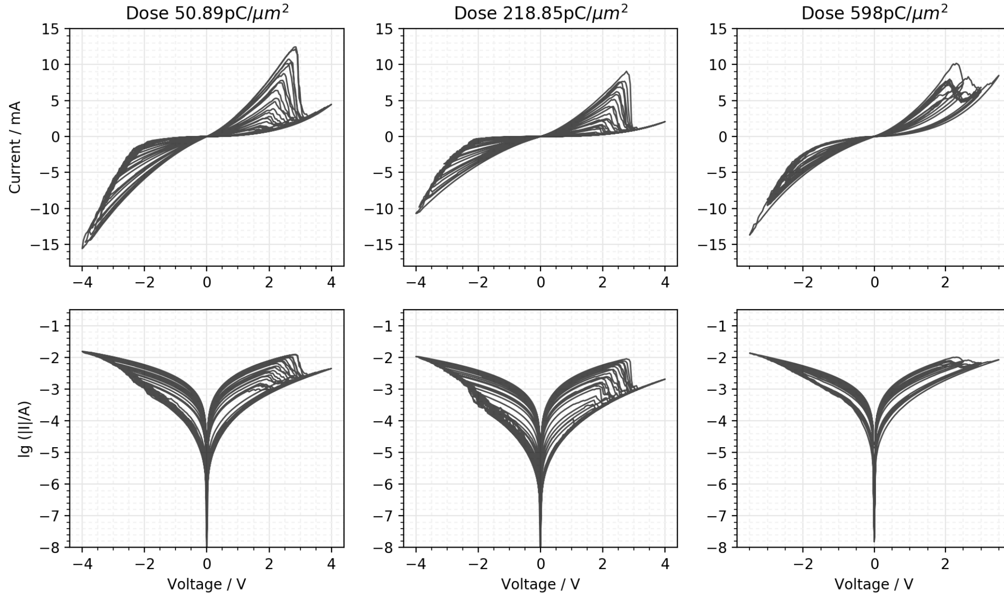



Figure 7-2: When the ion beam milling dosage increases greatly, the sweep cycles become distorted and the device cannot achieve a high resistance HRS. This is probably due to excess filamentary material embedded in the RS device which cannot be easily switched off. After 18 days, the device processed with the excessive dosage relaxes (not shown) and displays a I-V curve similar to devices freshly milled at lower dosages.

## 7.2 Atomic probe tomography

As mentioned in the introduction chapter, direct imaging of filaments, such as using HRTEM, is a challenging effort. Firstly, these filaments form at random locations within a device depending on local variations in the deposited films, so it is not certain where a lamella needs to be prepared from a device to capture a filament. Secondly, exposing a filament has unknown effects on the filament since these filaments are either mobile oxygen vacancies or conductive metal atoms. Oxygen vacancies could get annealed out of the film, or conductive metal atoms along a filament could cluster at the free surface of a lamella, making TEM observation on a lamella an inaccurate view of the actual filament formed within an RS device. This also causes in-situ RS TEM experiments, done on open slices, to be an inaccurate representation of RS in a typical device because the vast difference in diffusion coefficients in the bulk versus free surfaces are being ignored.

Direct imaging experiments, however hard both to execute and interpret, are

extremely valuable in further understanding of the RS mechanism as it provides physical dimensions useful for modelling. TEM imaging work on RS devices are often published in high impact journals and heavily cited as these experiments form a physical basis on which other models can be built. However, there now exists commercial equipment for atom probe tomography (APT) now in more than 100 locations around the world[76]. APT provides a richer 3D visualization of atomic identities within a sample by ionizing atoms from the surface and detecting their elemental identities using time of fight mass spectroscopy. The tool allows spatial localization of atoms down to 0.2nm. There have been some reported attempts of applying APT on RS stacks[77, 78]. However, such attempts are currently in the development stage since both fields have only gained prominence in the recent decades. Furthermore, the only published attempt [78] considered RS case mediated by oxygen vacancy filaments, but this is not a good use of APT because vacancies are not directly registered with APT.

Our RS device happens to be ideal for APT imaging and has the potential to provide a fundamental understanding into resistive switching. Metal dopant atoms have been found to be key in fabricating high yield and high consistently RS devices, but their distribution and size in our devices are unknown. The large mass difference between the doping Au versus the Al and O atoms in the matrix should simplify the requirements for these dopants to be imageable. Furthermore, our RS device can be fabricated with atomic layer deposition, so an RS stack can be directly deposited on an APT needle rather than having to do an additional FIB lift-out step.

An APT study would be useful to directly measure the distribution and identity of implanted atoms in the RS stack for subsequent use in computational models.

## 7.3 New degrees of freedom in stack design

We showed that changing a single layer RS device into a multilayer one and ion implantation of noble metal dopants can both improve the qualities of switching. However, we have barely scratched the surface of what is possible because we have

not yet optimized these two new sets of design parameters.

The ability to replace a single thick layer of oxide into multiple thinner ones to improve device switching consistency has great value. As mentioned briefly in chap. 4, this is useful because there is no need to search for a second oxide material to harness the high consistencies of bilayer oxide RS devices when there exist this simpler method of intercalating the single thick oxide with conductive spacers. Our method is also generalizable to beyond 2 and more layers. This might be useful in design restrictions where the HRS or LRS needs to be above some specified value, such as perhaps in meeting selector requirements in cross-bar arrays or to reduce power consumption during operation, where HRS or LRS can potentially be increased without causing switching stochasticity by simply increasing the combined oxide thickness and dividing this thicker oxide into a larger number of thinner oxides.

The use of metal dopants for improving device yield and increasing switching consistency is also rich for manipulation. The use of noble metal dopants, Au, Pd, Pt, have all been found (previously in Fig. ) to be favorable in improving switching consistency. However, effects of their concentration and implantation distribution have not been studied. There might be benefits to device longevity by having dopant ions at the center of an oxide layer, or by having dopant atoms aggregate to form a less mobile cluster.

We already see some promising signs that extended device lifespan is possible by replacing Au ion implantation by Au co-sputtered with Al2O3. Au was co-sputtered for bouts of 2 to 5s for an equivalent thickness of 1 to 3Å at periodic heights within a 2.0nm Al2O3 reactively sputtered film. The Au-doped Al2O3 film was repeated twice and separated by a WN spacer just like in our typical stack. This stack is not yet optimised, but already shows high consistencies in its LRS and HRS resistance and has remained switchable for a longer time of more than a month so far (compared to about two weeks for FIB-processed devices). This Au co-sputtered device is still lacking that it has erratic voltage thresholds for both the SET and RESET but future variations in process parameters can potentially solve this issue. This also shows that ion implantation can be replaced by a more accessible fabrication method, helpful in

enabling trials to be started by other research groups to replicate results as reported in this thesis.

## 7.4 Summary

In this work, we have shown strong experimental evidence that it is possible to fabricate high consistency RS devices using the strategies we propose, which is firstly, to use FIB-processing to implant noble metal Au dopants into oxides to serve as bridging atoms for oxygen vacancy filaments, and secondly, to divide single-layer oxides into multiple thinner oxide layers. These experimental evidence were supported by theoretical DFT calculations and network-level simulations to understand how Au or other noble metal atoms doped into multilayer oxides are able to aid in the RS phenomenon.

The greatest value in the devices fabricated in this work is not in the specific geometric configuration we have chose to pursue, but in the platform in which it has provided for accelerating future research into RS and resistive RAM (RRAM). This is because the deterministic manner in which we achieve near-perfect device yields and highly consistent RS behaviors can be transferable to other material or geometrical systems. We have demonstrated this flexibility in part by FIB-milling into e-beam deposited Au pads, by FIB-milling into other noble metal top contacts, and by replacing FIB-milling with Au co-sputtering, to show that our strategies still hold with these changes.

These highly consistent devices are not really intended to be used in consumer applications as-is, though it might suffice, but is intended for use by researchers as a reproducible base model on which to introduce changes and to study the effects of these changes. This will allow relations between causes and effects to be quickly established since the device-to-device and cycle-to-cycle variability no longer obscures the analysis of experimental data. The quantification method for switching consistency introduced in this work shall also come in handy under such an effort. As an example, our pulsed measurement results show the ease in which further mechanistic insights

can be obtained when a high consistency RS device is available for probing.

Looking ahead, this final chapter has described the spread of promising work which can be pursued next. These can either yield deeper insights into the RS mechanism, or can yield further meritorious resistive switching platforms which will not just have ideal RS properties but will have tunable behaviors so that the designed device can adapt to various applications as needed.

The work presented here will lower the barrier in finding the optimal device suitable for deployment in large arrays used for machine learning applications or hardware neuromorphic computing, such that isolated merits now seen in separate RS systems can potentially be brought together in a single ideal device to push the performance limits of resistive switching to unprecedented levels.

# Appendix A

# Microfabrication methods

## A.1 ALD deposition of WN/Al2O3 stack

N-type degenerate Si wafers were purchased from University Wafer. Wafers were dipped in 1:50 HF:H2O for 60s to remove native oxide and spin rinsed dried. Next, a wafer was loaded into a Oxford FlexAL ALD machine for plasma-enhanced deposition of alternate layers of WN and Al2O3 to give the stack Si/7.5nm WN/2.0nm Al2O3/3.0nm WN/2.0nm Al2O3/3.0nm WN/2.0nm Al2O3. The bottom electrode of the resistive switching device is the 7.5nm WN. The purpose of this WN is to build the stack starting from a well-defined layer to avoid wafer to wafer variations from an uncertified supply of Si wafers. WN is used instead of other commonly used metal due to CMOS requirements imposed on this particular ALD machine. For a 2-layer or 1-layer Al2O3 device, the thickness of each oxide layer will be increased to 3.0nm and 6.0nm respectively so that the combined oxide layer thickness remains constant. Deposition was done at 300°C. The deposition of WN is a N2/H2 plasma-enhanced reaction with bis(tert-butylimino)bis(dimethylamino)tungsten(VI) (BTBMW) precursor. The deposition of Al2O3 is a O2 plama-enhanced reaction with trimethylaluminum (TMA) precursor. Both recipes were supplied by the manufacturer. The thin film thickness were determined via X-ray reflectivity using a Rigaku SmartLab X-ray diffractometer, with both single films on wafers or composite films on wafers measured. The growth rate of WN and Al2O3 was deduced to be 0.5Å/cy and 1.0Å/cy respectively on the Oxford FlexAL ALD machine.

## A.2 Au deposition and FIB ion milling

Au deposition is typically performed on a Balzers tabletop sputterer at 130V and 40mA for 150s to give a film thickness of 30nm. There is no difference in device performance when Au is instead

being deposited with a AJA International sputterer or an e-beam deposition machine. No additional metal adhesion layer is used for Au deposition. A 30kV Ga ion beam on a FEI Helios NanoLab 600i DualBeam FIB/SEM was used to mill away material to produce a square mesa with a side wall length of 50$\mu$m. The SEM mode was used to image the chip to set-up for the FIB so there is no unintended FIB damage except as intended around the perimeter of the mesa. The milling was performed to a depth that exposes the Si substrate. This corresponds to an areal Ga ion dose of 80 to 120pC/$\mu$m$^2$. Resistive switching devices start in the LRS after ion beam milling.

## A.3 Reactively sputtered deposition

As an alternative to ALD deposition, the WN and Al2O3 can also be deposited by reactive sputtering. WN was reactively sputtered with a 29 sccm Ar and 6 sccm N2 mix at 3mT and DC 150W power on an AJA International Phase II J sputterer. Al2O3 was reactively sputtered with a 29 sccm Ar and 6 sccm O2 mix, with plasma struck at 30mT and deposition at 3mT. The deposition rate of Al2O3 was 0.11Å/s with a 300W RF power supply. If substituting FIB Au ion implantation, co-doping of Au can be done at this stage by flashing 2s to 5s of Au plasma at 20W DC for 3 to 7 times spaced evenly within each Al2O3 layer.

# Appendix B

# DFT method

The energetics of the $2 \times 2 \times 1$ perfect supercell (a= 9.62Å,c= 13.13Å) and all of the defects were calculated using density functional theory (DFT) using a plane-wave basis set, projector-augmented wave pseudopotentials and the Perdew-Burke-Ernzerhof (PBE) parameterization of the generalized gradient approximation (GGA) as the exchange-correlation functional, as implemented in the Vienna ab initio Simulation Package (VASP) v.5.4.1. A kinetic energy cutoff of 520 eV and a gamma-centered $2 \times 2 \times 2$ k-point mesh were used, resulting in a convergence accuracy of <1 meV/atom. All calculations were performed with a Gaussian smearing width of 0.05 eV and spin-polarized setting. Atomic positions were relaxed until the force on each atom was below 0.02 eV/Å.

The formation energy of $V_o^x$ in bulk Al2O3 was calculated as

$$E_f = E_{\mathrm{DFT}}^{V_o^x} - E_{\mathrm{DFT}}^{\mathrm{perf}} - \mu_o,$$

where $E_f$ is the formation energy of $V_o^x$ in bulk Al2O3, $E_{\mathrm{DFT}}^{V_o^x}$ is the DFT energy of the supercell with a $V_o^x$, $E_{\mathrm{DFT}}^{\mathrm{perf}}$ is the DFT energy of the perfect supercell with no defects, and $\mu_o$ is the chemical potential of oxygen in the system, calculated in the oxygen-rich limit, given below.

The formation energy of $V_o^x$ nearest neighbor (NN) to a dopant at the octahedral interstitial site of Al2O3 was calculated as

$$E_f = E_{\mathrm{DFT}}^{V_o^x - D_i^x} - E_{\mathrm{DFT}}^{D_i^x} - \mu_o,$$

where $E_f$ is the formation energy of $V_o^x$ at the NN site to the dopant; $E_{\mathrm{DFT}}^{V_O^x - D_i^x}$ is the DFT energy of the supercell with the dopant and $V_o^x$ at its NN site; $E_{\mathrm{DFT}}^{D_i^x}$ is the DFT energy of the supercell with only the interstitial dopant, and $\mu_o$ is the chemical potential of oxygen in the system, calculated

in the oxygen-rich limit, i.e.,

$$\mu_o\left(T, P_{O_2}\right) = \frac{1}{2}\left[E_{O_2}^{\text{DFT}} + E_{\text{over}} + \mu_{O_2}^{o}(T, P^o) + kT\ln\left(\frac{P_{O_2}}{P^o}\right)\right]$$

where $E_{O_2}^{\text{DFT}}$ is the DFT energy of the O2 molecule; $E_{\text{over}}$ is the correction for the O2 overbinding error caused by GGA, taken as 1.36 eV as identified by Wang et al.; $\mu_{O_2}^{o}\left(T, P^o\right)$ is the difference in chemical potential of O2 gas between T=0 K and the temperature of interest, at a reference pressure of $P^o = 1$ atm, as obtained from thermo-chemical tables; $P_{O_2}$ is the partial pressure of oxygen gas (1 atm in the O-rich limit). The binding energies of the complexes were calculated as

$$E_b(V_o^x - D_i^x) = E_f^{V_o^x} + E_f^{D_i^x} - E_f^{V_o^x - D_i^x}$$

Note that binding energy does not depend on the chemical potentials of the various elements involved. The chain formation energy (per $V_o^x$) was calculated as,

$$E_f = \left(E_{\text{DFT}}^{nV_o^x - D_i^x} - E_{\text{DFT}}^{D_i^x} - n\mu_o\right)/n,$$

where $E_f$ is the formation energy of $V_o^x$ chain with the dopant, $E_{\text{DFT}}^{nV_O^x - D_i^x}$ is the DFT energy of the supercell with the dopant and n $V_o^x$ (n=8) chain, $E_{\text{DFT}}^{D_i^x}$ is the DFT energy of the supercell with only the interstitial dopant, and $\mu_o$ is the chemical potential of oxygen in the system.

The chain binding energy (per $V_o^x$) was calculated as

$$E_b(nV_O^x - D_i^x) = \left(nE_f^{V_O^x} + E_f^{D_i^x} - E_f^{nV_O^x - D_i^x}\right)/n.$$

# Appendix C

# Standard electrical test set

The electrical characterization of a freshly milled RS device always first starts with the following set of I-V sweeps. This ensures that (1) as-fabricated resistance values can be recorded, and (2) comparisons can be made easily between devices fabricated across wafers or time by referring to their results from this test set. This ensures that subtle transient behaviors across different devices can be compared to one another.

| No. | Voltage waypoints / V | Purpose |
|---|---|---|
| 1 | (0,-0.5,0.5,0) | Extract as-fabricated resistance. |
| 2 | (0,-2,2,0) | Observe presence of stray as-fabricated materials. |
| 3 | (0,-3,3,0) | Observe presence of stray as-fabricated materials. |
| 4 | (0,-2.0,2.0,-2.1,2.1,-2.2,2.2,-2.3,2.3 ,-2.4,2.4,-2.5,2.5,-2.6,2.6,-2.7,2.7,-2.8 ,2.8,-2.9,2.9,-3.0,3.0,-3.1,3.1,-3.2,3.2 ,-3.3,3.3,-3.4,3.4,-3.5,3.5,-3.6,3.6,-3.7 ,3.7,-3.8,3.8,-3.9,3.9,-4.0,4.0,0) | Observe multi-bit switching character (with transient character). |
| 5 | (0,-2.0,2.0,-2.1,2.1,-2.2,2.2,-2.3,2.3,-2.4 ,2.4,-2.5,2.5,-2.6,2.6,-2.7,2.7,-2.8,2.8,-2.9 ,2.9,-3.0,3.0,-3.1,3.1,-3.2,3.2,-3.3,3.3,-3.4 ,3.4,-3.5,3.5,-3.6,3.6,-3.7,3.7,-3.8,3.8,-3.9 ,3.9,-4.0,4.0,0) | Observe multi-bit switching character. |
| 6 | (0,-3.5,3.5,-3.5,3.5,-3.5,3.5,-3.5, 3.5,-3.5,3.5,-3.5,3.5,-3.5,3.5,-3.5, 3.5,-3.5,3.5,-3.5,3.5,0) | Observe repeatability in switching. |
| 7 | (0,-3.5,3.5,-3.5,3.5,-3.5,3.5,-3.5, 3.5,-3.5,3.5,-3.5,3.5,-3.5,3.5,-3.5, 3.5,-3.5,3.5,-3.5,3.5,0) | Observe repeatability in switching. |

Table C.1: Basic test set used for electrical characterization. Voltage waypoints are shown in lieu of the actual voltage points used. A script converts these voltage waypoints into smoother voltage ramps with 20mV steps.

# Bibliography

[1] James M Tour and Tao He. Electronics: The fourth element. *Nature*, 453(7191):42, 2008.

[2] Dmitri B Strukov, Gregory S Snider, Duncan R Stewart, and R Stanley Williams. The missing memristor found. *nature*, 453(7191):80, 2008.

[3] Leon Chua. Memristor-the missing circuit element. *IEEE Transactions on circuit theory*, 18(5):507–519, 1971.

[4] Isaac Abraham. The case for rejecting the memristor as a fundamental circuit element. *Scientific reports*, 8(1):10972, 2018.

[5] Sascha Vongehr and Xiangkang Meng. The missing memristor has not been found. *Scientific reports*, 5:11657, 2015.

[6] Michael Chui and San Francisco. Artificial intelligence the next digital frontier? *McKinsey and Company Global Institute*, 47, 2017.

[7] Yu-Chi Chang and Yeong-Her Wang. Resistive switching behavior in gelatin thin films for nonvolatile memory application. *ACS applied materials & interfaces*, 6(8):5413–5421, 2014.

[8] Xingli He, Jian Zhang, Wenbo Wang, Weipeng Xuan, Xiaozhi Wang, Qilong Zhang, Charles G Smith, and Jikui Luo. Transient resistive switching devices made from egg albumen dielectrics and dissolvable electrodes. *ACS applied materials & interfaces*, 8(17):10954–10960, 2016.

[9] Eduardo Di Mauro, Olivier Carpentier, SI Yáñez Sánchez, N Ignoumba Ignoumba, Myriam Lalancette-Jean, Josianne Lefebvre, Shiming Zhang, Carlos FO Graeff, Fabio Cicoira, and Clara Santato. Resistive switching controlled by the hydration level in thin films of the biopigment eumelanin. *Journal of Materials Chemistry C*, 4(40):9544–9553, 2016.

[10] Julia Wunsche, Yingxin Deng, Prajwal Kumar, Eduardo Di Mauro, Erik Josberger, Jonathan Sayago, Alessandro Pezzella, Francesca Soavi, Fabio Cicoira, Marco Rolandi, et al. Protonic and electronic transport in hydrated thin films of the pigment eumelanin. *Chemistry of Materials*, 27(2):436–442, 2015.

[11] Julia Wünsche, Luis Cardenas, Federico Rosei, Fabio Cicoira, Reynald Gauvin, Carlos FO Graeff, Suzie Poulin, Alessandro Pezzella, and Clara Santato. In situ formation of dendrites in eumelanin thin films between gold electrodes. *Advanced Functional Materials*, 23(45):5591–5598, 2013.

[12] Niloufar Raeis Hosseini and Jang-Sik Lee. Biocompatible and flexible chitosan-based resistive switching memory with magnesium electrodes. *Advanced Functional Materials*, 25(35):5586–5592, 2015.

[13] Alexander Teverovsky. Reverse bias behavior of surface mount solid tantalum capacitors. 2002.

[14] Christian Terwiesch and Yi Xu. The copy-exactly ramp-up strategy: Trading-off learning with process change. *IEEE Transactions on Engineering Management*, 51(1):70–84, 2004.

[15] Wei Wu, Huaqiang Wu, Bin Gao, and He Qian. Si interface barrier modification on memristor for brain-inspired computing. In *Journal of Physics: Conference Series*, volume 864, page 012064. IOP Publishing, 2017.

[16] HY Peng, GP Li, JY Ye, ZP Wei, Z Zhang, DD Wang, GZ Xing, and T Wu. Electrode dependence of resistive switching in Mn-doped ZnO: Filamentary versus interfacial mechanisms. *Applied Physics Letters*, 96(19):192113, 2010.

[17] F Yang, M Wei, and H Deng. Bipolar resistive switching characteristics in cuo/zno bilayer structure. *Journal of Applied Physics*, 114(13):134502, 2013.

[18] David Robert Lamb. *Electrical conduction mechanisms in thin insulating films*. Halsted Press, 1967.

[19] Deok-Hwang Kwon, Kyung Min Kim, Jae Hyuck Jang, Jong Myeong Jeon, Min Hwan Lee, Gun Hwan Kim, Xiang-Shu Li, Gyeong-Su Park, Bora Lee, Seungwu Han, et al. Atomic structure of conducting nanofilaments in TiO2 resistive switching memory. *Nature Nanotechnology*, 5(2):148, 2010.

[20] Yasuo Takahashi and Masashi Arita. In-situ TEM observation of ReRAM switching. *2014 IEEE International Meeting for Future of Electron Devices, Kansai (IMFEDK)*, pages 1–2, Jun 2014.

[21] Yang Yang, Weiming Lü, Yuan Yao, Jirong Sun, Changzhi Gu, Lin Gu, Yanguo Wang, Xiaofeng Duan, and Richeng Yu. In situ TEM Observation of Resistance Switching in Titanate Based Device. *Sci. Rep.*, 4:3890, Jan 2014.

[22] Lide Yao, Sampo Inkinen, and Sebastiaan Van Dijken. Direct observation of oxygen vacancy-driven structural and resistive phase transitions in La2/3Sr1/3MnO3. *Nature communications*, 8:14544, 2017.

[23] Xiao-Jian Zhu, Jie Shang, Gang Liu, and Run-Wei Li. Ion transport-related resistive switching in film sandwich structures. *Chinese science bulletin*, 59(20):2363–2382, 2014.

[24] MNIST on Benchmarks.AI, May 2019. [Online; accessed 5. May 2019].

[25] Tayfun Gokmen and Yurii Vlasov. Acceleration of deep neural network training with resistive cross-point devices: Design considerations. *Frontiers in neuroscience*, 10:333, 2016.

[26] GW Burr, P Narayanan, RM Shelby, Severin Sidler, Irem Boybat, Carmelo di Nolfo, and Yusuf Leblebici. Large-scale neural networks implemented with non-volatile memory as the synaptic weight element: Comparative performance analysis (accuracy, speed, and power). In *2015 IEEE International Electron Devices Meeting (IEDM)*, pages 4–4. IEEE, 2015.

[27] Jae Sung Lee, Shinbuhm Lee, and Tae Won Noh. Resistive switching phenomena: A review of statistical physics approaches. *Appl. Phys. Rev.*, 2(3):031303, Sep 2015.

[28] Lei Li and Dianzhong Wen. Memory behavior of multi-bit resistive switching based on multiwalled carbon nanotubes. *Organic Electronics*, 34:12–17, 2016.

[29] Amit Prakash, Debanjan Jana, and Siddheswar Maikap. TaOx-based resistive switching memories: prospective and challenges. *Nanoscale research letters*, 8(1):418, 2013.

[30] Gang Niu, Hee-Dong Kim, Robin Roelofs, Eduardo Perez, Markus Andreas Schubert, Peter Zaumseil, Ioan Costina, and Christian Wenger. Material insights of HfO 2-based integrated 1-transistor-1-resistor resistive random access memory devices processed by batch atomic layer deposition. *Scientific reports*, 6:28155, 2016.

[31] Yanfei Qi, Ce Zhou Zhao, Chenguang Liu, Yuxiao Fang, Jiahuan He, Tian Luo, Li Yang, and Chun Zhao. Comparisons of switching characteristics between Ti/Al2O3/Pt and TiN/Al2O3/Pt RRAM devices with various compliance currents. *Semiconductor Science and Technology*, 33(4):045003, 2018.

[32] Wen-Yuan Chang, Hsin-Wei Huang, Wei-Ting Wang, Cheng-Hao Hou, Yu-Lun Chueh, and Jr-Hau He. High uniformity of resistive switching characteristics in a Cr/ZnO/Pt device. *Journal of the Electrochemical Society*, 159(3):G29–G32, 2012.

[33] Youngjun Park and Jang-Sik Lee. Flexible multistate data storage devices fabricated using natural lignin at room temperature. *ACS applied materials & interfaces*, 9(7):6207–6212, 2017.

[34] Yu-Chi Chang, Ke-Jing Lee, Cheng-Jung Lee, Li-Wen Wang, and Yeong-Her Wang. Bipolar resistive switching behavior in sol-gel MgTiNiOx memory device. *IEEE Journal of the Electron Devices Society*, 4(5):321–327, 2016.

[35] Zhongrui Wang, WG Zhu, AY Du, L Wu, Z Fang, Xuan Anh Tran, WJ Liu, KL Zhang, and H-Y Yu. Highly uniform, self-compliance, and forming-free ald $HfO_2$-based RRAM with Ge doping. *IEEE Transactions on electron devices*, 59(4):1203–1208, 2012.

[36] Dimin Niu, Qiaosha Zou, Cong Xu, and Yuan Xie. Low power multi-level-cell resistive memory design with incomplete data mapping. In *2013 IEEE 31st International Conference on Computer Design (ICCD)*, pages 131–137. IEEE, 2013.

[37] Haifeng Ling, Mingdong Yi, Masaru Nagai, Linghai Xie, Laiyuan Wang, Bo Hu, and Wei Huang. Controllable organic resistive switching achieved by one-step integration of cone-shaped contact. *Advanced Materials*, 29(35):1701333, 2017.

[38] Victor Yi-Qian Zhuo, Yu Jiang, Rong Zhao, Lu Ping Shi, Yi Yang, Tow Chong Chong, and John Robertson. Improved switching uniformity and low-voltage operation in $TaO_x$-based RRAM using Ge reactive layer. *IEEE Electron Device Letters*, 34(9):1130–1132, 2013.

[39] B Govoreanu, GS Kar, YY Chen, V Paraschiv, S Kubicek, A Fantini, IP Radu, L Goux, S Clima, R Degraeve, et al. 10× 10nm 2 hf/hfo x crossbar resistive ram with excellent performance, reliability and low-energy operation. In *2011 International Electron Devices Meeting*, pages 31–6. IEEE, 2011.

[40] Sandeep Munjal and Neeraj Khare. Valence Change Bipolar Resistive Switching Accompanied With Magnetization Switching in CoFe2O4 Thin Film. *Sci. Rep.*, 7(1):12427, Sep 2017.

[41] Zhongrui Wang, HongYu Yu, Xuan Anh Tran, Zheng Fang, Jinghao Wang, and Haibin Su. Transport properties of $HfO_{2-x}$ based resistive-switching memories. *Phys. Rev. B*, 85(19):195322, May 2012.

[42] Joonmyoung Lee, Hyejung Choi, Dong-jun Seong, Jaesik Yoon, Jubong Park, Seungjae Jung, Wootae Lee, Man Chang, Chunhum Cho, and Hyunsang Hwang. The impact of Al interfacial layer on resistive switching of La0.7Sr0.3MnO3 for reliable ReRAM applications. *Microelectron. Eng.*, 86(7):1933–1935, Jul 2009.

[43] Yi-Jen Huang, Tzu-Hsien Shen, Lan-Hsuan Lee, Cheng-Yen Wen, and Si-Chen Lee. Low-power resistive random access memory by confining the formation of conducting filaments. *AIP Adv.*, 6(6):065022, Jun 2016.

[44] Kyooho Jung, Y Kim, Hyunsik Im, Hyungsang Kim, and Baeho Park. Leakage transport in the high-resistance state of a resistive-switching NbOx thin film prepared by pulsed laser deposition. *J. Korean Phys. Soc*, 59:2778–2781, 2011.

[45] Chih-Yang Lin, Chen-Yu Wu, Chung-Yi Wu, Chenming Hu, and Tseung-Yuen Tseng. Bistable resistive switching in Al2O3 memory thin films. *Journal of The Electrochemical Society*, 154(9):G189–G192, 2007.

[46] Debanjan Jana, Subhranu Samanta, Siddheswar Maikap, and Hsin-Ming Cheng. Evolution of complementary resistive switching characteristics using IrOx/GdOx/Al2O3/TiN structure. *Applied Physics Letters*, 108(1):011605, 2016.

[47] Wanki Kim, Sung Il Park, Zhiping Zhang, and Simon Wong. Current conduction mechanism of nitrogen-doped $AlO_x$ RRAM. *IEEE Transactions on Electron Devices*, 61(6):2158–2163, 2014.

[48] Sir Nevill Francis Mott and Ronald Wilfrid Gurney. *Electronic processes in ionic crystals*. Clarendon Press, 1948.

[49] John G Simmons. Poole-frenkel effect and schottky effect in metal-insulator-metal systems. *Physical Review*, 155(3):657, 1967.

[50] H-S Philip Wong, Heng-Yuan Lee, Shimeng Yu, Yu-Sheng Chen, Yi Wu, Pang-Shiu Chen, Byoungil Lee, Frederick T Chen, and Ming-Jinn Tsai. Metal–oxide RRAM. *Proceedings of the IEEE*, 100(6):1951–1970, 2012.

[51] Qi Liu, Weihua Guan, Shibing Long, Ming Liu, Sen Zhang, Qin Wang, and Junning Chen. Resistance switching of Au-implanted-ZrO2 film for nonvolatile memory application. *Journal of Applied Physics*, 104(11):114514, 2008.

[52] B Gao, HW Zhang, Shimeng Yu, B Sun, LF Liu, XY Liu, Y Wang, RQ Han, JF Kang, B Yu, et al. Oxide-based RRAM: Uniformity improvement using a new material-oriented methodology. In *2009 Symposium on VLSI Technology*, pages 30–31. IEEE, 2009.

[53] Yan Wang, Qi Liu, HangBing Lü, ShiBing Long, Wei Wang, YingTao Li, Sen Zhang, WenTai Lian, JianHong Yang, and Ming Liu. Improving the electrical performance of resistive switching memory using doping technology. *Chinese science bulletin*, 57(11):1235–1240, 2012.

[54] Dongsoo Lee, Dong-jun Seong, Hye jung Choi, Inhwa Jo, R Dong, W Xiang, Seokjoon Oh, Myeongbum Pyun, Sun-ok Seo, Seongho Heo, et al. Excellent uniformity and reproducible resistance switching characteristics of doped binary metal oxides for non-volatile resistance memory applications. In *2006 International Electron Devices Meeting*, pages 1–4. IEEE, 2006.

[55] Shimeng Yu and Pai-Yu Chen. Emerging memory technologies: Recent trends and prospects. *IEEE Solid-State Circuits Magazine*, 8(2):43–56, 2016.

[56] Xiaolong Zhao, Sen Liu, Jiebin Niu, Lei Liao, Qi Liu, Xiangheng Xiao, Hangbing Lv, Shibing Long, Writam Banerjee, Wenqing Li, et al. Confining cation injection to enhance CBRAM performance by nanopore graphene layer. *Small*, 13(35):1603948, 2017.

[57] Cong Ye, Tengfei Deng, Junchi Zhang, Liangping Shen, Pin He, Wei Wei, and Hao Wang. Enhanced resistive switching performance for bilayer HfO2/TiO2 resistive random access memory. *Semiconductor Science and Technology*, 31(10):105005, 2016.

[58] Tingting Tan, Yihang Du, Ai Cao, Yaling Sun, Hua Zhang, and Gangqiang Zha. Resistive switching of the HfOx/HfO2 bilayer heterostructure and its transmission characteristics as a synapse. *RSC Advances*, 8(73):41884–41891, 2018.

[59] Masoud Akbari, Min-Kyu Kim, Dongshin Kim, and Jang-Sik Lee. Reproducible and reliable resistive switching behaviors of AlOx/HfOx bilayer structures with Al electrode by atomic layer deposition. *RSC Advances*, 7(27):16704–16708, 2017.

[60] Jill S Becker, Seigi Suh, Shenglong Wang, and Roy G Gordon. Highly conformal thin films of tungsten nitride prepared by atomic layer deposition from a novel precursor. *Chemistry of materials*, 15(15):2969–2976, 2003.

[61] Menelaos Tsigkourakos, Panagiotis Bousoulas, Vaggelis Aslanidis, Evangelos Skotadis, and Dimitris Tsoukalas. Ultra-low power multilevel switching with enhanced uniformity in forming free TiO2-x-based RRAM with embedded Pt nanocrystals. *Physica Status Solidi (a)*, 214(12):1700570, 2017.

[62] S Muraoka, K Osano, Y Kanzawa, S Mitani, S Fujii, K Katayama, Y Katoh, Z Wei, T Mikawa, K Arita, et al. Fast switching and long retention Fe-O ReRAM and its switching mechanism. In *2007 IEEE international electron devices meeting*, pages 779–782. IEEE, 2007.

[63] Jingon Jang, Han-Hyeong Choi, Sung Hoon Paik, Jai Kyeong Kim, Seungjun Chung, and Jong Hyuk Park. Highly improved switching properties in flexible aluminum oxide resistive memories based on a multilayer device structure. *Advanced Electronic Materials*, 4(12):1800355, 2018.

[64] Microchip Technologies. In *ATmega32U4 MCU datasheet*. 2015. Sect 29.

[65] Texas Instruments. In *THS3202 op-amp datasheet*. 2002.

[66] Chikako Yoshida, Kohji Tsunoda, Hideyuki Noshiro, and Yoshihiro Sugiyama. High speed resistive switching in Pt/TiO2/TiN film for nonvolatile memory application. *Applied Physics Letters*, 91(22):223510, 2007.

[67] N Xu, LF Liu, X Sun, C Chen, Y Wang, DD Han, XY Liu, RQ Han, JF Kang, and B Yu. Bipolar switching behavior in TiN/ZnO/Pt resistive nonvolatile memory with fast switching and long retention. *Semiconductor Science and Technology*, 23(7):075019, 2008.

[68] Gyoung-Ho Buh, Inrok Hwang, and Bae Ho Park. Time-dependent electroforming in NiO resistive switching devices. *Applied Physics Letters*, 95(14):142101, 2009.

[69] Andrew Lohn, Patrick R Mickel, and Matthew Marinella. Stages of switching in tantalum oxide memristor. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2013.

[70] N Panwar, A Khanna, P Kumbhare, I Chakraborty, and U Ganguly. Self-heating during submicrosecond current transients in Pr0.7Ca0.3MnO3-based RRAM. *IEEE Transactions on Electron Devices*, 64(1):137–144, 2017.

[71] Stephan Menzel, Moritz von Witzleben, Viktor Havel, and Ulrich Böttger. The ultimate switching speed limit of redox-based resistive switching devices. *Faraday Discussions*, 213:197–213, 2019.

[72] A Chen, S Haddad, YC Wu, TN Fang, S Kaza, and Z Lan. Erasing characteristics of Cu2O metal-insulator-metal resistive switching memory. *Applied Physics Letters*, 92(1):013503, 2008.

[73] Bin Yu, Xuhui Sun, Sanghyun Ju, David B Janes, and M Meyyappan. Chalcogenide-nanowire-based phase change memory. *IEEE Transactions on Nanotechnology*, 7(4):496–502, 2008.

[74] Zhongrui Wang, Saumil Joshi, Sergey Savel'ev, Wenhao Song, Rivu Midya, Yunning Li, Mingyi Rao, Peng Yan, Shiva Asapu, Ye Zhuo, et al. Fully memristive neural networks for pattern classification with unsupervised learning. *Nature Electronics*, 1(2):137, 2018.

[75] Selina La Barbera, Dominique Vuillaume, and Fabien Alibart. Filamentary switching: synaptic plasticity through device volatility. *ACS nano*, 9(1):941–949, 2015.

[76] Arun Devaraj, Daniel E Perea, Jia Liu, Lyle M Gordon, Ty J Prosa, Pritesh Parikh, David R Diercks, Subhashish Meher, R Prakash Kolli, Ying Shirley Meng, et al. Three-dimensional nanoscale characterisation of materials by atom probe tomography. *International Materials Reviews*, 63(2):68–101, 2018.

[77] Jochen M Schneider. Identification of atomic scale mechanisms active in resistance switching devices by atom probe tomography. Technical report, Rheinisch-Weefaelisch Technisch Hochschule Aachn Aachn Germany, 2018.

[78] JH Lee, EJ Cha, YT Kim, BK Chae, JJ Kim, SY Lee, HS Hwang, and CG Park. A study of threshold switching of nbo2 using atom probe tomography and transmission electron microscopy. *Micron*, 79:101–109, 2015.

[79] DR Lamb and PC Rundle. A non-filamentary switching action in thermally grown silicon dioxide films. *British Journal of Applied Physics*, 18(1):29, 1967.

[80] Myoung-Jae Lee, Chang Bum Lee, Dongsoo Lee, Seung Ryul Lee, Man Chang, Ji Hyun Hur, Young-Bae Kim, Chang-Jung Kim, David H Seo, Sunae Seo, et al. A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta2O5-x/TaO2-x bilayer structures. *Nature Materials*, 10(8):625, 2011.

[81] Rainer Waser and Masakazu Aono. Nanoionics-based resistive switching memories. In *Nanoscience And Technology: A Collection of Reviews from Nature Journals*, pages 158–165. World Scientific, 2010.

[82] Chiyui Ahn, Zizhen Jiang, Chi-Shuen Lee, Hong-Yu Chen, Jiale Liang, Luckshitha S Liyanage, and H-S Philip Wong. 1d selection device using carbon nanotube FETs for high-density cross-point memory arrays. *IEEE Transactions on Electron Devices*, 62(7):2197–2204, 2015.

[83] Zhongrui Wang, Saumil Joshi, Sergey E Savel'ev, Hao Jiang, Rivu Midya, Peng Lin, Miao Hu, Ning Ge, John Paul Strachan, Zhiyong Li, et al. Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nature Materials*, 16(1):101, 2017.

[84] Mirko Prezioso, Farnood Merrikh-Bayat, BD Hoskins, Gina C Adam, Konstantin K Likharev, and Dmitri B Strukov. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature*, 521(7550):61, 2015.

[85] Sung Hyun Jo, Ting Chang, Idongesit Ebong, Bhavitavya B Bhadviya, Pinaki Mazumder, and Wei Lu. Nanoscale memristor device as synapse in neuromorphic systems. *Nano Letters*, 10(4):1297–1301, 2010.

[86] Duygu Kuzum, Shimeng Yu, and HS Philip Wong. Synaptic electronics: materials, devices and applications. *Nanotechnology*, 24(38):382001, 2013.

[87] Xiaobing Yan, Jianhui Zhao, Sen Liu, Zhenyu Zhou, Qi Liu, Jingsheng Chen, and Xiang Yang Liu. Memristor with Ag-cluster-doped TiO2 films as artificial synapse for neuroinspired computing. *Advanced Functional Materials*, 28(1):1705320, 2018.

[88] Ming-Chi Wu, Yi-Wei Lin, Wen-Yueh Jang, Chen-Hsi Lin, and Tseung-Yuen Tseng. Low-power and highly reliable multilevel operation in $ZrO_2$ 1T1R RRAM. *IEEE Electron Device Letters*, 32(8):1026–1028, 2011.

[89] Jung Ho Yoon, Kyung Min Kim, Seul Ji Song, Jun Yeong Seok, Kyung Jean Yoon, Dae Eun Kwon, Tae Hyung Park, Young Jae Kwon, Xinglong Shao, and Cheol Seong Hwang. Pt/Ta2O5/HfO2-x/Ti resistive switching memory competing with multilevel NAND flash. *Advanced Materials*, 27(25):3811–3816, 2015.

[90] Geoffrey W Burr, B"ulent N Kurdi, J Campbell Scott, Chung Hon Lam, Kailash Gopalakrishnan, and Rohit S Shenoy. Overview of candidate device technologies for storage-class memory. *IBM Journal of Research and Development*, 52(4-5):449–464, 2008.

[91] An Chen. A review of emerging non-volatile memory (NVM) technologies and applications. *Solid-State Electronics*, 125:25–38, 2016.

[92] Shisheng Li, Shunfeng Wang, Dai-Ming Tang, Weijie Zhao, Huilong Xu, Leiqiang Chu, Yoshio Bando, Dmitri Golberg, and Goki Eda. Halide-assisted atmospheric pressure growth of large WSe2 and WS2 monolayer crystals. *Applied Materials Today*, 1(1):60–66, 2015.

[93] Vinod K Sangwan, Hong-Sub Lee, Hadallia Bergeron, Itamar Balla, Megan E Beck, Kan-Sheng Chen, and Mark C Hersam. Multi-terminal memtransistors from polycrystalline monolayer molybdenum disulfide. *Nature*, 554(7693):500, 2018.

[94] Yuchao Yang, Peng Gao, Siddharth Gaba, Ting Chang, Xiaoqing Pan, and Wei Lu. Observation of conducting filament growth in nanoscale resistive memories. *Nature Communications*, 3:732, 2012.

[95] Peifu Cheng, Kai Sun, and Yun Hang Hu. Memristive behavior and ideal memristor of 1T phase MoS2 nanosheets. *Nano Letters*, 16(1):572–576, 2015.

[96] Krzysztof Szot, Wolfgang Speier, Gustav Bihlmayer, and Rainer Waser. Switching the electrical resistance of individual dislocations in single-crystalline SrTiO3. *Nature Materials*, 5(4):312, 2006.

[97] Chikako Yoshida, Kentaro Kinoshita, Takahiro Yamasaki, and Yoshihiro Sugiyama. Direct observation of oxygen movement during resistance switching in NiO/Pt film. *Applied Physics Letters*, 93(4):042106, 2008.

[98] Chris J McDonald. The evolution of Intel's Copy EXACTLY! technology transfer method. *Intel Technology Journal*, 4(1):1–6, 1998.

[99] Takeshi Yanagida, Kazuki Nagashima, Keisuke Oka, Masaki Kanai, Annop Klamchuen, Bae Ho Park, and Tomoji Kawai. Scaling effect on unipolar and bipolar resistive switching of metal oxides. *Scientific reports*, 3:1657, 2013.

[100] Richard G Southwick and William B Knowlton. Stacked dual-oxide mos energy band diagram visual representation program (irw student paper). *IEEE Transactions on Device and Materials Reliability*, 6(2):136–145, 2006.

[101] Richard G Southwick, Aaron Sup, Amit Jain, and William B Knowlton. An interactive simulation tool for complex multilayer dielectric devices. *IEEE Transactions on Device and Materials Reliability*, 11(2):236–243, 2011.