Novel Electrical and Chemical Findings on SiO_x-based ReRAM Devices.

Luca Montesi

Supervisor: Prof. Anthony Kenyon

Thesis submitted for the degree of Doctor of Philosophy.

Department of Electrical and Electronics Engineering,
University College London.
London, UK

24th January, 2018

I, Luca Montesi confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

[Signature removed]

____

Luca Montesi 24/01/2018

ABSTRACT

Existing non-volatile flash memory technologies are characterised by slow access time, high power consumption and a quickly approaching scaling limit. Filamentary resistive RAM (ReRAM) is an emerging type of storage device that relies on the electrically driven change in resistance of a thin film sandwiched between two electrodes. The active region is often a binary oxide that develops a restorable conductive filament thanks to the electrically driven movement of oxygen. This technology offers potential sub-10 nm scalability, nanosecond programming with direct overwriting (unlike FLASH) and an appealing sub pJ/bit power consumption (compared to nJ/bit of FLASH).

In this thesis, metal-insulator-metal ReRAM devices with a TiN/SiOx/TiN structure are used. While other binary oxides have been used in the literature, SiO_x must be used in its amorphous form allowing for easier fabrication, and is an extremely well-studied material as its CMOS compatibility dates back 40 years.

Using the above devices, it was possible to observe data storage performance comparable to the one of other types of ReRAM. More interestingly, it was observed that the resistance states of this family of devices may be programmed using nanosecond pulses of identical magnitude, possibly leading to simple programming circuits. Consequently, it is shown that this programming method may also be used to gradually increase or decrease the device resistance state as well as have devices enter states that relax over time. These types of behaviour mean that SiO_x devices may be used in neuromorphic networks that require components whose behaviour resembles the one of the neuronal synapsis or the mammalian brain's forgetting process.

The literature reports on endurance-hindering electrode deformation phenomena during the operation of oxide-based ReRAM devices. A residual gas analyser (RGA) was used to detect that oxygen species are emitted during operation and therefore confirmed that such phenomena are caused by oxygen emission. Using SIMS (secondary ion mass spectroscopy) analysis on devices switched in atmospheres containing isotopically labelled oxygen, it was observed that, under deformed regions, it is possible to find incorporated atmospheric oxygen. Additionally, reducing atmospheric pressure had negative impact on device reliability.

SiO_x-based filamentary ReRAM is a strong candidate in the search for alternatives to flash memory. Moreover, these devices display behaviour that may be useful in applications trying to emulate the mammalian brain. Having observed device dependence on its atmosphere, endurance issues may now be addressed using electrodes capable of either adsorbing oxygen without bubbling or letting it go through without cracking.

To my father.

TABLE OF CONTENTS

Abstract ... 3

Table of Contents ... 6

List of Figures ... 10

List of Tables ... 13

Acronyms ... 14

1 Introduction ... 15

1.1 Research Roadmap ... 15

1.2 Rationale and Objectives of the Research ... 16

2 Background and Literature Review ... 24

2.1 Filamentary Resistive Switching ... 24

2.1.1 Extrinsic Mechanisms ... 27

2.1.2 Intrinsic Mechanisms ... 27

2.2 Electron Current Transport in Dielectrics ... 33

2.3 Why Silicon Oxide? ... 38

2.4 Relevant Literature ... 39

2.4.1 Regular Voltage Sweeps ... 39

2.4.2 Voltage Pulsing with Current Compliance ... 42

2.4.3 Fast Voltage Pulsing ... 44

2.4.4 Macroscopic Structural Changes in Oxide Following Electrical Stress . 46

2.4.5 Neuromorphic Behaviour ... 49

2.4.6 Labelled Oxygen Exchange 52
2.5 Conclusions 53
3 Experimental Method and Rationale 56
3.1 Sample Preparation 56
3.2 Experimental Tools 57
3.2.1 Hardware 57
3.2.2 Software 69
3.3 Experimental Techniques 72
3.3.1 Electrical Measurements in Air 72
3.3.2 In situ Switching 82
3.3.3 Surface Analysis and Tomography 89
4 Results from Electrical Measurements 92
4.1 Electrical Behaviour in Air 92
4.1.1 Current Compliance Voltage Sweeps in Air 92
4.1.2 Current Compliance Long Pulsing 105
4.1.3 Fast Nanosecond Pulsing 113
4.2 Electrical Behaviour in Modified Atmosphere 124
4.2.1 Electroforming 124
4.2.2 Resetting 129
4.2.3 Setting 133
4.2.4 Conclusions 137

5 Detection of Emitted Oxygen 139

5.1 Observations Attributable to Oxygen Movement 139

5.2 Oxygen Gas Detection 146

5.3 Ion Emission Detection 149

5.3.1 Channeltron Ion Detection without Mass Identification 149

5.3.2 SIMS Ion Detection with Mass Identification 152

5.4 Conclusions 154

6 $^{18}O_2$ Exchange 156

6.1 $^{18}O_2$-Exchange Results 156

6.1.1 Initial SIMS Studies 156

6.1.2 Further SIMS Studies 165

7 Summary, Conclusions and Future Work 169

A. Appendix 171

A.1 List of Publications Arising from This Work 171

A.1.1 Journal Publications 171

A.1.2 Conference Presentations Given 172

A.2 Description of Software Scripts and Drivers 173

A.2.1 Keithley 4200-SCS SMU Matlab Scripts 173

A.2.2 Keithley 4200-SCS PMU C Drivers 174

A.2.3 Keithley 4200-SCS PMU Matlab Scripts to Control C Drivers 175

A.2.4 Keithly 6430 Matlab Scripts 175

A.3 References 176

LIST OF FIGURES

Figure 1: Diagrams of a ReRAM 3D array. 21

Figure 2: Electrical switching in a ReRAM cell........ 25

Figure 3: Chemical differences are observed between states in intrinsic ReRAM cell 29

Figure 4: Schematic representation of fictional typical IV curves 41

Figure 5: IV curves of a hafinia-based device........ 42

Figure 6: Example of endurance and reliability testing 44

Figure 7: Optical microscope images of deformed regions 48

Figure 8: Schematic of a TiN/SiO_x/TiN wafer 57

Figure 9: Vacuum chamber 61

Figure 10 : Exchange rig used 62

Figure 11: Schematic of the steady gas flow probe chamber setup: 64

Figure 12: Linkam chamber photos 65

Figure 13: Schematic of the iONTOF TOF.SIMS[5] 68

Figure 14: Equivalent circuits showing how the oscilloscope was connected to the SMU output 74

Figure 15: V/T plots of output behaviour of a 4200-SCS SMU 77

Figure 16: V/T plots of output behaviour of the 6430 SMU 78

Figure 17: Oscilloscope plot of a sample PMU output from a short pulse-sequence. 79

Figure 18: Equivalent circuit used to keep track of device behaviour while using PMUs 81

Figure 19: Overnight RGA profile........ 84

Figure 20: RGA survey scan of the exchange rig high-vacuum chamber during $^{18}O_2$-atmosphere leak 87

Figure 21: Typical electroform, reset and set IV curves of a TiN/SiO_x/TiN wafer device......93
Figure 22: IV sweeps showing electroform, reset and set of a device. Both voltage and current were measured at each point......94
Figure 23: Collected in-air electroforming voltage data......95
Figure 24: Probability spread of reset and set voltages of TiN/SiO_x/TiN device. 96
Figure 25: Fowler-Nordheim plot......97
Figure 26: Poole-Frenkel plot......98
Figure 27: Log-Log plot of IV sweeps......99
Figure 28: TAT plot (zoom in)......101
Figure 29 :Thermionic emission plot......102
Figure 30: Plots used to detect the existence of quantum point contact conductance......104
Figure 31: Annotated version of IV curves......106
Figure 32: Cycling results from a TiN/SiO_x/TiN device......109
Figure 33: Reliability over time of a TiN/SiO_x/TiN- device tested at 85˚C......109
Figure 34: V/t plot of set......111
Figure 35: Voltage measured across a 90 Ω resistor in series with a DUT being reset by SMU......112
Figure 36: Switching between states using identical pulses......117
Figure 37: Device cycling using fast pulses without current compliance......118
Figure 38: Voltage measured across the DUT as a positive pulse sequence was applied. 120
Figure 39: Voltage across terminals of a device being electroformed using a fast pulse......121

Figure 40: Temporary device states 122

Figure 41: IV curves showing electroforming in different atmospheres 125

Figure 42: Column plot showing collected electroforming voltages........ 126

Figure 43: IV curves showing reset events in different atmospheres 130

Figure 44: Column scatter of reset voltages. 131

Figure 45: IV curves of set events that occurred in different atmospheres 134

Figure 46: Column scatter plot of set voltages 135

Figure 47: Unstable set in $^{18}O_2$ atmosphere 136

Figure 48: Surface deformations........ 141

Figure 49: a) AFM image of surface deformations 142

Figure 50: Surface deformations........ 143

Figure 51: Surface deformations........ 145

Figure 52: Detection of species between mass 24 and 34 over six sampling cycles 147

Figure 53: Oxygen gas detection over time........ 149

Figure 54: Streak seen on the focused ion beam microscope while forming or setting a device........ 150

Figure 55: Pixel intensity vs. time of a white streak........ 152

Figure 56: Detection of $^{16}O_2^-$ ions (superoxide) during device formation. 154

Figure 57: Oxygen depth profiles carried out by FIB-SIMS 159

Figure 58: Oxygen depth profiles carried by ToF-SIMS 161

Figure 59: Images of surface deformations 164

Figure 60: Oxygen profiles of top and bottom electrodes ramps 167

LIST OF TABLES

Table 1: Some dielectric conduction processes 34

Table 2: Summary of some literature that uses IV sweeps to characterise devices. .. 43

Table 3: Summary of some literature that uses pulses with compliance to characterise devices. 45

Table 4: Summary of some literature that uses fast pulses to characterise devices. . 47

ACRONYMS

RAM	Random Access Memory
ReRAM/RRAM	Resistive RAM
MIM	Metal Insulator Metal
CMOS	Complementary metal–oxide–semiconductor
flash	Not and Acronym. Refers to electronically programmable memory.
NVM	Non-Volatile Memory
CF	Conductive Filament
PCRAM	Phase Change RAM
FeRAM	Ferroelectric RAM
ECM	Electrochemical Metallization Memory
DUT	Device Under Test
LRS	Low Resistance State
HRS	High Resistance State
TMO	Transition Metal Oxide
ITO	Indium Tin Oxide
(C)AFM	(Conductive) Atomic Force Microscope/Microscopy
SIMS	Secondary Ion Mass Spectrometry
RGA	Residual Gas Analyser
SPICE	Simulation Program with Integrated Circuit Emphasis
FPGA	Field Programmable Gate Array
SMU	Source Measure Units
PMU	Pulse Measure Unit
ToF-SIMS	Time of Flight Secondary Ion Mass Spectroscopy
RGA	Residual Gas Analyser
FIB	Focused Ion Beam

1 INTRODUCTION

1.1 RESEARCH ROADMAP

Today's non-volatile solid-state memory is reliant on NAND flash. Unfortunately, this technology has been showing limits in terms of further scalability and lower power consumption. These shortcomings will only be rectified using new types of electronic components. New technologies reliant on devices whose resistance can be changed electrically have emerged and are being investigated as possible replacements to flash. One of these types of memories is known as filamentary oxide-based resistive RAM (ReRAM). This work is focused on addressing the main shortcoming of this novel silicon-based technology in its unipolar variant: its endurance. While this technology is not yet mature for commercialisation, this work aims at making it more appealing to industry and provides results that will help improve the endurance performance of future device iterations. While investigating reasons behind unsatisfactory endurance, device behaviour occurred which may be used to emulate biological synapses opening doors for further novel applications.

Chapter 1 explains the goal of the research that was undertaken. The chapter briefly introduces the current technological state of non-volatile solid-state storage, and outlines the problems and expectations associated with solid-state non-volatile memory (NVM) today. The chapter also addresses the inefficiency found in using today's binary central processing units (CPUs) running AI algorithms. The chapter ends by evaluating the advantages and disadvantages of resistive RAM (ReRAM) technology for binary storage and neuronal behaviour emulation.

Chapter 2 provides a look into the different kinds of electrochemical ReRAM available today, and opens by focussing on the physical and chemical aspects of the different technologies. The chapter also reviews devices found in the literature, and comparatively evaluates their electrical performance and limitations. Finally, this chapter looks at literature that helped compare the performance of devices in this work.

Chapter 3 details experiments on the current generation of unipolar silicon oxide based devices fabricated at University College London (UCL) and other facilities UCL collaborates with. The chapter discusses the experimental methods used, including the equipment and software. Chapters 4, 5 and 6 present results of experimental testing and compare these results with the literature introduced in chapter 2. Each chapter includes at least one novel contribution.

In conclusion, Chapter 7 presents accomplishments, limitations, and areas of improvement, and offers a path for continued improvement of the technology used.

1.2 Rationale and Objectives of the Research

Just like transistors, non-volatile memory (NVM) has never stopped evolving since the introduction of the digital computer. In the early days, two factors drove technological innovation of NVM: memory access time and the amount of data that can be stored in a given space (density). While memory access time and density remain relevant today, recent demand for portable and embedded devices requires a focus on reducing power consumption whilst increasing density. Additionally, the existing market demands memory with high 'endurance', which translates into longer device life. Accordingly, the factors that drive innovation of NVM today are

more complex and include memory access times, density or size, power consumption, and endurance. Data retention time is also an important measure of quality and 10 years are expected in [1-3].

Currently, the most popular solid-state storage element is flash memory [4]. However, this is characterised as having relatively low endurance (10^3 – 10^5 rewrites [1]), low erase speed (in the milliseconds) with no fast data overwrite possible, high write voltages (~15 V) [5, 6], and high power consumption (nJ/bit instead of pJ/bit found in other technologies) [7]. Furthermore, additional scaling-down of flash memory is difficult [5] due to device physics with only complex 3D stacking looking promising [8]. Accordingly, this is a good point in time to switch to a technology superior to flash memory [3, 9].

The above limitations of non-volatile memory have also been appearing in other silicon based devices as scaling approaches its limit due to fundamental physics of silicon. These concerns are voiced in recent updates of the International Technology Roadmap for Semiconductors (ITRS) [6]. The introduction and development of non-CMOS based technologies are described as an appealing future path [6, 10-13]. This envisages the possibility of having mixed CMOS and non-CMOS chips. When referring to storage devices, it appears from ITRS that [14] chemical reduction/oxidation or redox (as well as spin transfer torque (SST) RAM, more on this later) based ReRAM is to be pursued for commercialisation.

Many NVM technologies have emerged but have displayed downfalls. Ferroelectric (Fe) RAM [6, 15, 16] is an option based on DRAM design. It relies on a high dielectric ferromagnetic material instead of an insulator and shows short latency (< 300 ns), low programming voltages, power consumption and variability,

and high endurance (>10^{10} cycles). Unfortunately, this technology presents its own set of problems. While FeRAM has been used commercially, it shows low scalability, multi layer compatibility or 3D integration. It also has modest retention (>1 year) and a medium fabrication cost [6]. Others that its fabrication is not easy due to the ferroelectric layer being highly susceptible to oxidation [9].

Another alternative is magnetic resistive (M)RAM reliant on magnetic tunnel junctions (MTJs) [17, 18]. This technology has no known endurance hindering process, is easy to fabricate, and may be integrated in current CMOS technology [19]. Here one has a thin insulating magnet whose spin is electrically programmed and whose tunnelling resistance depends on the orientation of spin. Unfortunately, MRAM requires very high writing currents to be programmed quickly (> 1 mA) and resistance change between states is not very large [20].

One more technology, which also requires a high programming current, is phase change RAM (PCRAM) [9, 21-23]. This technology makes use of heat to change the phase of a material (a chalcogenide) between crystalline and amorphous and consequently its resistance. Despite its high programming current and power consumption, PCRAM has great scalability (< 10 nm) and resistance separation between states. It may be used in 3D arrays. Its fabrication cost, retention (> 1 year), latency (0.3 – 10 μs), endurance (< 10^{10}) and variability remain average [6, 23].

Programming current is lowered in spin-transfer torque (SST)-RAM, which is similar in approach to MRAM but uses a conductive layer instead of a dielectric. It appears appealing, though according to [14] its scalability is doubtful, it still requires relatively high programming currents/energy (~20pJ/bit [24]) and, just like all other mentioned technologies, it requires layers of materials not commonly found in

CMOS technology. These technologies are therefore not ideal candidates to replace flash.

Given the above, focus was given on a new kind of electrochemical memory not found above and known as 'filamentary Resistive RAM' or more commonly as ReRAM. This term generally only refers to this type of memory even though also all of the above NVM memory types (FeRAM, MRAM, PCRAM, and SST-RAM) rely on resistance change. Most of the theory behind ReRAM originated around the 1960s [25, 26] and 1970s [27]. This older research shows that thin dielectrics can conduct electricity after an initial electroforming step, which employs a high biasing voltage [28]. Following this step, which increased the conductivity of the pristine material, applying lower voltages allows a change in the amount of conducted current, leading to different resistance states. The process leads to two terminal devices with programmable resistance [29] by relying on a phenomenon known as resistance switching. In the late 90s, researchers discovered a new use for resistance switching: storing binary data. The discovery led to a vast interest from the research community ever since [5].

To better explain the behaviour of these devices, it is important to start by looking at their dielectrics, which are typically made out of metal oxides such as hafnium oxide, titanium oxide, tantalum oxide, or, more rarely, silicon oxide. The composition of the oxide will lead to different families of devices that are commonly placed under the ReRAM umbrella. All these devices have a dielectric usually with a thickness in the tens of nanometres [1] or less in certain cases [30, 31]. This insulator is sandwiched between two electrodes [32]. This film typically develops a conductive filament when treated with an electrical stimulus. This filament can then

be electrically broken and restored many times. These electric signals are generally not greater than a few volts, with programming time shown to be as fast as 100 picoseconds [15] leading to low power consumption. Breaking and rebuilding the filament results in an electrically programmable memory device. The state of a filamentary ReRAM device can be read using a non-programming low voltage sampling pulse. It must be noted, however, that some oxides used in ReRAM, such as perovskites, make use of conductive regions [33-35] rather than filaments. These regions are areas often as wide as the top electrode.

Because more than two resistance states are possible, multi-bit memory elements may be created. For instance, four possible resistance states would correspond to a two-bit memory cell [36-38]. Similarly, it is possible to create temporary states, which relax back to the initial state [39]. Accordingly, the filament that electroforms inside the dielectric must be controllable to achieve satisfactory device operation.

Filamentary ReRAM devices typically rely on a single conductive filament, whose size is in the range of nanometres, and electrodes that can be reduced to very small sizes. Theoretically, these contact points must only be as large as the filament. This leads to simple two-terminal devices, which, thanks to their size, may be placed in 2D arrays. These arrays have a high density but are problematic due to current routing. In these configurations, 'cross talk' occurs when adjacent cells are in an ON or low-resistance state. The use of diodes or resistors mitigates such issues [4, 5, 40]. Figure 1 shows a possible 2D array configuration as well as how diodes may be used to avoid crosstalk [5, 40]. 3D arrays are also possible [1].

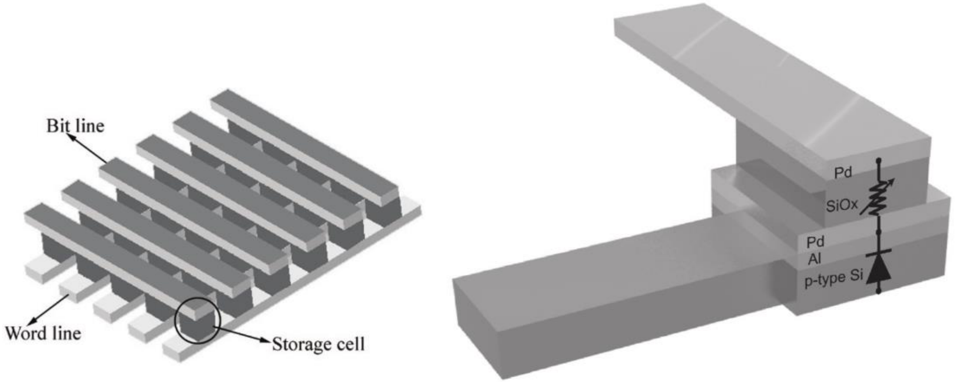


Figure 1: Diagrams of a ReRAM 3D array. On the left is an overview of how cells may be interconnected and addressed. On the right, it is shown how diodes may be used to avoid cross talk. Images taken from Ref. [5, 40] with permission obtained from publishers.

Most of the characteristics discussed make filamentary ReRAM devices very appealing compared to flash memory [9]. However, two important requirements to compete with flash memory still have not been discussed: cycling endurance and reliability. According to Ref. [21, 41], in order to be competitive with flash memory there is a need for an endurance of 10^3 to 10^7 programming cycles. Unfortunately, this requirement is still not easily met in unipolar devices. As discussed later, reports claiming high endurance rates seem to discard large amounts of data to portray their devices as appealing [9, 42-44]. Furthermore, today's expectations for reliability require that data be retained for at least 10 years under normal conditions. Fortunately, this requirement is exceeded by current commercial ReRAM devices [1, 45]. Therefore, the main area to improve is the cycling endurance of unipolar filamentary ReRAM [1].

ReRAM, including its filamentary variety, not only sees a use in ordinary data storage, which is the way flash is commonly used. It can also be used directly in current mode logic circuits to turn volatile devices such as look-up tables (LUTs), multiplexers and flipflops into non-volatile ones. The basis of current mode circuits is found in simple sensing amplifiers. ReRAM could enable field programmable gate

arrays (FPGAs) with instantaneous power on. Today, these programmable circuits require lengthy programming at start-up, generally from non-volatile flash memories [22, 46, 47]. In situ programming of ReRAM elements in current mode circuits remains however problematic due to programmer loading on the differential amplifiers. While ReRAM reliability over time remains a concern, programming power consumption, programming speed and endurance are not main concerns as FPGAs are not reprogrammed frequently. FPGAs are found in all digital applications where the design and fabrication of an application specific integrated circuit (ASIC) is not justified.

More recently, some filamentary ReRAM devices have also shown some similarities with neuronal synapses under certain kinds of electrical stimulation. Lots of research goes into the emulation of biological neuronal networks to implement artificial intelligence. These implementations were traditionally realised using highly inefficient approaches based on sequential super-computers. The use of filamentary ReRAM devices in machines whose objective is to implement artificial intelligence is appealing. Highly parallel, post-binary and asynchronous implementation would help achieve high overall efficiency. Filamentary ReRAM devices have a small device footprint that allows for high density and therefore parallelism. Different states allow for semi-analogue behaviour, and low programming power leads to low power dissipation [48-52].

This work's objective is further analysis of the current generation of unipolar silicon oxide devices, as questions are still present as to how these devices function. The devices in question rely on intrinsic switching in silicon oxide, as no dopants are needed in the switching layer and no ionic migration occurs to and from the

electrodes. Therefore, intrinsic devices are CMOS compatible and fabrication is cost effective. Electrical characterization provides insight into observable performance (switching time, endurance, retention, etc.). Knowing these dynamics allows for the simplification of programming circuits and improves the understanding of the electrochemistry that takes place. Visually observable changes in devices and the detection of oxygen absorption and emission help strengthen conclusions. For instance, it is reaffirmed that intrinsic devices rely on mobile oxygen ions and concluded that their escape is detrimental to device endurance.

2 BACKGROUND AND LITERATURE REVIEW

2.1 FILAMENTARY RESISTIVE SWITCHING

Focus will now be given to filamentary ReRAM devices by providing insight on how they function and can be programmed. As mentioned in the previous chapter, this type of memory encompasses devices that can be electrically stimulated to conduct electricity thanks to the electroforming (or forming) of a single nanoscale conductive filament [21, 53].

These devices are generally Metal-Insulator-Metal structures (MIM). In general, a pristine (never been biased) two terminal MIM device will act as an electrically open circuit (>~$10^{11}\Omega$, assuming an insulator thickness in the 10s of nanometres). Following electroforming (also known as a 'electroform'), it will act as a low resistance device (~600Ω). A return to a high resistance state (this is known as a 'reset') will increase resistance (~100kΩ), but not to a pristine state value. Bringing the device back to a low resistance state could follow this step and is known as 'setting' the device. Overall, the two states are referred to as high resistance state (HRS) and low resistance state (LRS), respectively, and involve the electrically driven rupture and rebuilding of the initially electroformed filament. These reactions are extremely non-linear, as a small change in voltage results in a large change in resistance between the two electrodes. Furthermore, control over the strength and consequent resistance of the filament is possible; therefore, different and intermediate resistance states can be created.

Electrical conditions to change a device state are very critical to ensure proper functionality. Electroforming is generally accomplished using a high voltage with a low current compliance (or CC, discussed later). 'Bipolar' devices are

therefore ‘reset’ using a smaller voltage of opposite polarity without current compliance. ‘Unipolar’ devices, instead, are ‘reset’ using a smaller voltage with the same polarity as the electroforming one but without current compliance. ‘Setting’ will generally be carried out with current compliance using a voltage of the same polarity (but usually smaller in magnitude) to the one used to initially electroform [21]. Nonetheless, certain devices do not require a specific ‘electroforming’ voltage as a ‘setting’ voltage is sufficient to electroform [54]. Devices that rely on opposite polarity resetting are generally asymmetric and commonly referred to as MIM’. Here the ‘prime’ symbol indicates that one of the two electrodes is made out of a different conductive material. Multiple metal layer electrodes are sometimes used to exploit the physical properties of one metal (outer metal) and the electrical ones of another (inner metal). The physical structure of a ReRAM element and its general electrical behaviour are shown in Figure 2 [4]. Cycling through the described electrical stimuli should lead to many successful switching events.

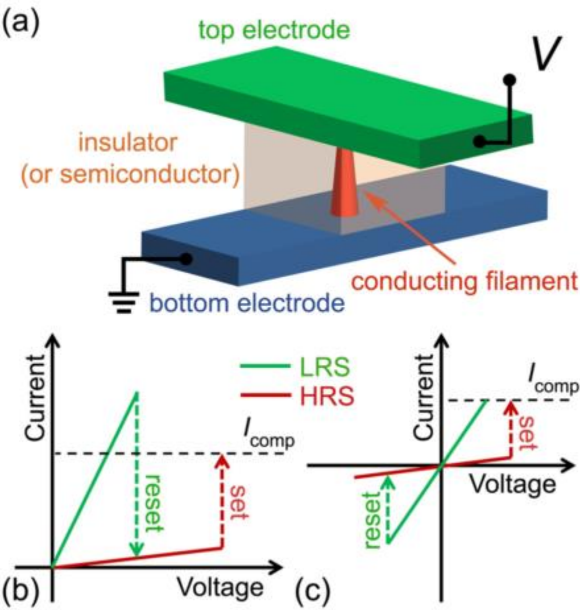


Figure 2: Electrical switching in a ReRAM cell. (a) Schematic representation of a ReRAM cell. (b) Typical switching characteristics of a unipolar device. (c) Typical switching characteristics of a bipolar device. I_{COMP} shows the point where current compliance kicked in while ‘setting’. Taken from [4]. Permission obtained from publisher.

The resistance state a of a device can be sampled using a low voltage. This voltage must be low to ensure that the device filament is not altered but, at the same time, the voltage must be high enough to induce a sufficiently high current to trigger a current-to-voltage amplifier which produces a binary reading [20, 21].

Current compliance is a biasing technique by which the maximum amount of current injected through a device can be limited. Once a specified current is flowing through a device, the voltage applied will be reduced by the biasing source in order not to exceed the specified current [21]. When it comes to filamentary ReRAM, this allows a filament to stop electroforming once it has become conductive. This limit in injected energy allows for the avoidance of disruptive thermal meltdowns. Therefore, it takes a certain time for the filament to electroform and thanks to current compliance, the necessary voltage is applied for this amount of time and not for longer. Voltage source current compliance insertion delay is therefore very important; a device may melt down in case of delay.

With previous knowledge of a device filament electroforming time, it is possible to avoid the use of current compliance. The use of a high-speed pulsing source will be, in this case, equivalent to the use of a standard (or slow) one with compliance if voltage is cut off as soon as the filament is electroformed. This, however, is not easy to accomplish because it is hard to predict programming time differences between devices. Filaments electroformed this way will generally be weaker than those electroformed with current compliance. While current compliance gradually adjusts the voltage as needed this other approach does not.

Filamentary ReRAM devices can be divided into two sub-families: extrinsic and intrinsic. Focus will remain on the latter, but it is important to mention the

former to appreciate their differences and similarities. A brief look will first be given at non-CMOS compatible extrinsic memories because of the relative simplicity of their mechanism of operation. A more in-depth review of intrinsic memory devices will follow.

2.1.1 Extrinsic Mechanisms

In extrinsic MIM' (asymmetric) devices, the insulating (and active layer) becomes conductive due to the movement of metal ions. One scenario sees the conductive filament being made up of indiffused atoms from one of the electrodes (generally, the anode as positive metal ions will be attracted to the cathode). This process is referred to as Electrochemical Metallization (ECM) leading to Conductive Bridge RAM (CBRAM) [21]. Alternatively, these filament-creating ions are already within the insulating layer. In this case, they migrate and rearrange into a conductive filament when an electric field crosses them. A combination of both phenomena is not to be excluded. In both cases, one can talk about extrinsic memory as the filament is created through atoms not normally found within the insulating material.

2.1.2 Intrinsic Mechanisms

Intrinsic filamentary ReRAM relies, like extrinsic memory, on the electroformation and manipulation of a single filament within an MIM structure. Switching occurs because of the intrinsic nature of the insulator. This type of ReRAM has been the focus of UCL research [2, 20, 22, 36, 37, 39, 53, 55-59]. Unlike extrinsic memory, intrinsic devices can be unipolar or bipolar. The former tend to have an MIM (symmetric) configuration, whereas the latter appear to be MIM' (asymmetric) devices. The choice of electrodes has shown to be important: their physical properties (strength, likelihood to oxidise, etc.) will affect device endurance whereas

their work function will affect their electrical behaviour [60]. As mentioned electrodes may be made of different metal layers to benefit from physical and electrical properties of different metals.

Intrinsic ReRAM relies on two different types of phenomena to create and destroy a conductive filament [53, 61-63]. The first phenomenon depends on ion and vacancy drift due to a stimulating electric field [37, 55, 60, 64]. The second phenomenon, in unipolar devices, makes use of the Joule or thermal effect that a current flowing through a material can induce [65]. These processes often occur concurrently and aid each other during operation [64]. In bipolar cells, an inverted electric field drives the second phenomenon.

UCL research group's main product is a TiN/SiO_x/TiN device which is initially a pristine unipolar MIM (symmetric) device [37]. Here, TiN is a conductive ceramic which, unlike a metal, should not easily ionise. SiO_x is silicon rich silica with $x<2$. In this device, the SiO_x film was grown by co-sputtering Si and SiO_2 using an Ar beam. The resulting deposited layer is non-stoichiometric, amorphous and exhibits columnar growth. This film has defects, and some silicon dangling bonds are bound to be present. These can be looked at as oxygen vacancies as they are ready to react with oxygen ions [37] (Figure 3 left). The amorphous nature of the material allows it to have random defects, which make it a useful material for ReRAM purposes. However, it is this stochastic nature that makes it a hard material to study and understand.

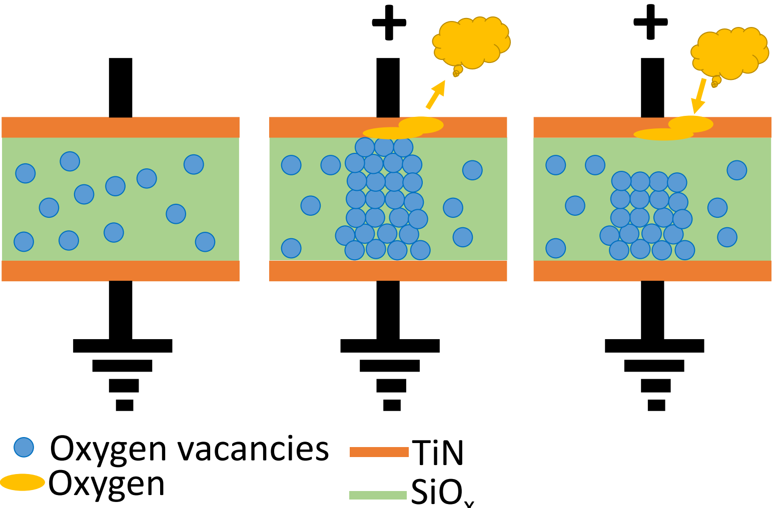


Figure 3: Chemical differences are observed between states in intrinsic ReRAM cell: left: a pristine device, centre: an electroformed device, right: a reset device. Circles represent oxygen vacancies whereas yellow regions represent oxygen.

The above situation changes upon applying an electroforming voltage to the TiN electrodes as this induces an electric field across the SiO_x layer. This bias has to be limited in time to avoid break down (current compliance or pulse shaping may be used [39]) and creates oxygen ions by breaking silicon-oxygen metastable bonds found throughout the SiO_x [60], adding to the number of oxygen vacancies already present. As per UCL's recent unpublished findings, this process appears to be aided by electron injection meaning that a field alone is not sufficient. This stems from a minimal current being needed to electroform. Oxygen ions, which are negatively charged, drift towards the anode, leaving a path of oxygen vacancies along the columnar structure of the material [37]. Since these vacancies allow for trap assisted tunnelling (TAT) [55] (more on this in section 2.2), a conductive filament with a low electrical resistance has effectively been created (Figure 3 centre). The device has therefore entered a low resistance state (LRS). Per simulations carried out at UCL, electroforming does not appear to be aided by Joule heating, due to a low maximum current. This overall process is very important for subsequent switching as the bulk of the filament will stay in place from now on. The process stops as soon as a

conductive filament has electroformed as current compliance (or pulse end) will kick in and reduce the voltage applied to the device terminals. The oxygen ions that have drifted towards the anode will accumulate near the interface, in the anode, and some may escape. Such accumulation of oxygen in the anode may lead to physical damage [60] or to oxygen emission as these ions will try to escape through the positive contact. The device is now conductive and in a low resistance state.

During the unipolar reset process, a bias lower than that used for electroforming is applied across the device electrodes. This bias will not be limited in time or current as it has to break the conductive filament through Joule heating. Current will now exceed the one used during electroforming and required time will be longer than to electroform. While looking at unipolar devices, some unwanted electric field will be present but it will not be strong enough (thanks to the now lower voltage) to re-form the filament to be broken (or reset). The introduced heat will increase the mobility of oxygen ions, which will attempt to find a new state of equilibrium by re-bonding with the silicon dangling bonds (i.e. the conductive oxygen vacancies previously generated). It appears that atmospheric oxygen ions are also used in this process. This leads to a new state of the filament. The bulk of the filament structure, which is made by conductive oxygen-vacancies is still present. However, a small part of it has now been destroyed due to oxygen recombination (Figure 3 right). It appears that some of the oxygen ions involved in re-oxidising the filament are acquired from atmospheric oxygen surrounding the sample. Device resistance will now be between that of a pristine device (where no filament was present) and that of a device in the LRS in which a full filament is present. This device is labelled as being in a reset or high resistance state (HRS).

Resetting bipolar devices bipolarly follows a similar process to the one above. In this case, resetting a device is carried out by applying a smaller and inverted voltage compared to the one used to set. When breaking a filament, setting an upper boundary for pulse time or current compliance is not critical. As in the unipolar case, the Joule effect may help to break the filament. However, an opposite electric field will now play a major role in assisting the oxygen recombination process. The filament is therefore reset or broken [59] and the device ends up in an HRS. Work presented in this thesis did not involve the use of bipolar devices.

With both unipolar and bipolar devices, setting is done in a way similar to electroforming: by applying an electric field for a short time or using current compliance. However, the necessary voltage is now lower: a smaller gap needs to be ionised (and therefore turned into oxygen vacancies) than when one had a pristine insulating layer. This is the gap that is created by the reset processes just described. This process, which can be repeated, appears to sometimes increase top electrode damage and release oxygen species into the atmosphere. The device can now cycle between set and reset states. Eelctroforming is no longer necessary.

Unipolar and bipolar devices provide the possibility of sampling their resistance state by applying a low voltage to the device terminals. This can be a positive or negative bias. The low voltage and resulting current will not affect the resistance state of the device. The magnitude of the flowing current is used to determine the resistance state the device is in.

This family of filamentary devices can be observed to only use one single filament, as electrode size does not seem to affect performance. If this were not the case, a larger electrode area would allow for more filaments to create and affect

behaviour. Heating up these devices decreases their resistivity. This shows that the filament is of a semi-conductive nature as metals generally increase in resistivity when heated up [55]. This supports the idea that with intrinsic switching no metallization is involved. It is also interesting to see that as a filament is gradually set using increasing current compliance, device resistance resistivity drops in quantized steps. Here each step represents a $G_0/2$ or half a quantum of conductance (where $G_0=2e^2/h$). This may be explained by looking at the filament as many atomic-sized constrictions assembled together [36] (more on this in 2.2). Similar results are found in hafnia-based devices as shown in Ref. [66]. In extrinsic devices, this phenomenon seems to occur by the full G_o [67, 68].

This type of device may also be switched using identical fast pulses (without current compliance) leading to novel applications [53]. For instance, filament strength [69] (or device resistance) can be changed in small steps [70] by applying trains of nanosecond-domain pulses. When pulses are too short, transient state change takes place. These phenomena make devices tempting for use in modern highly parallel hardware neural networks.

Nevertheless, the switching mechanism behind intrinsic devices remains quite controversial [71] and several different theories and categorisation methods exist [4, 21, 72, 73]. The one described above is the one that most closely appears to model the devices here characterised. Similarly, many different devices exist; some are fully intrinsic and some are not. Section 2.4 provides an overview of ReRAM devices found in the current literature. Although, most of the literature discusses (without positive confirmation) oxygen as the principal actor behind device

switching in devices involving oxides [5, 21, 25, 28, 37, 38, 44, 51, 55, 59, 60, 64, 74-84], not all theories explain functionality in an identical manner.

2.2 Electron Current Transport in Dielectrics

Section 2.1.2 provided an explanation on how switching occurs in UCL's ReRAM devices. It emerged that different types of electron conduction methods must occur within the dielectric active layer for switching to occur as required. This section is used to provide further insight into these different dielectric conduction processes [2].

A dielectric should ideally be insulating. However, this is not always the case especially when it comes to thin dielectric layers. Electrons can move through dielectrics by many different processes [85] as shown in Table 1.

Filamentary oxide-based ReRAM generally relies on a few dielectric electron transport methods. Some methods depend mainly on the electrode-dielectric interface barrier height: direct tunnelling, Fowler-Nordheim tunnelling, and thermionic emission, which is also known as Schottky emission. Some other methods depend mainly on the dielectric trap energy level: Poole-Frenkele emission, TAT (trap-assisted-tunnelling), DC Ionic conduction, and Ohmic conduction [86]. Brief explanations of each one of these processes of interest follows.

Process Name	**Band Diagram Process Description**	**Simplified Band Diagram**	**Type of Mathematical Model for Conductance: *J*=**	**Linearized Plot Axes**
Tunnelling	Electron are waves which, according to quantum mechanics, have a non-zero probability of transmitting through a barrier.	e^-	$k_{tun}E_i^2 e^{\left(-\frac{4\sqrt{2m^*}(q\phi_B)^{3/2}}{3q\hbar E_i}\right)}$	
Fowler-Nordheim Tunnelling	Occurs when electrons tunnel through a triangular barrier and continue in the oxide conduction band. The barrier thickness decreases with applied field.	e^-	$C_{FN}E^2 e^{\left(-\frac{4\sqrt{2m^*}(q\phi_B)^{3/2}}{3q\hbar E_i}\right)}$	$x = \frac{1}{V}$ $y = log(I/V^2)$
Thermionic Emission	Thermal energy provides electrons with energy to hop over a triangular barrier and enter a conduction band.	e^-	$A^{**}T^2 e^{\left[-\frac{q\left(\phi_B - \sqrt{qE_{i/4\pi\epsilon_i}}\right)}{kT}\right]}$	$x = V^{\frac{1}{2}}$ $y = \log(I)$
Poole-Frenkel Emission	Thermal energy provides energy to trapped electrons to hop over a triangular barrier.	e^-	$qn_0\mu E_i e^{\left[-\frac{q\left(\phi_B - \sqrt{qE_{i/\pi\epsilon_i}}\right)}{kT}\right]}$	$x = V^{\frac{1}{2}}$ $y =$ Log(I/V)
Trap Assisted Tunneling	Electrons tunnel through a barrier to a well leading to an equal or lower energy state. The process repeats.	e^-	$k_{TAT}e^{\left(-\frac{8\pi\sqrt{2qm^*}}{3hE_i}\Phi_t^{\frac{3}{2}}\right)}$	$x = \frac{1}{V}$ $y = log(I)$
Ohmic Conduction	Electrons stay in the conduction band		$k_{OC}E_i e^{-\frac{\triangle E_{ac}}{kT}}$	$x = \log(V)$ or V $y = \log(I)$ or I
DC Ionic Conduction	Electrons are injected as negative ions migrate towards the positive interface. Current plateaus once TAT flow is established.		$k_{DCIC}\frac{E_i}{T}e^{-\frac{\triangle E_{ac}}{T}k}$	
Space-charge-limited current	A positively charged sample will acquire electrons to reach a non-charged state		$\frac{9\epsilon_i\,\mu V^2}{8d^2}$	$x = V^2$ $y = I$

Table 1: Some dielectric conduction processes [2, 85, 87]. J=current density, ϕ_B=barrier height, E_i=electric field in insulator, ϵ_i=insulator permittivity, m^*=effective mass, d=insulator thickness, $\triangle E_{ac}$=activation energy of electrons, q=electron charge, h=plank's constant $\hbar$=reduced plank's constant, k=Boltzmann constant, A**=effective Richardson constant.

With traditional tunnelling, an electron crosses a thin dielectric or an entire rectangular potential barrier whose energy is greater than the one of the electron. This is explained in quantum mechanics by describing the electron as a wave function. Solving this function in the presence of an energy barrier will lead to solutions where the wave is found past the barrier. This leads to a non-zero probability that an electron can travel through such barrier. The thicker the dielectric, the fewer electrons will manage to 'tunnel' through it.

Fowler-Nordheim tunnelling sees a trapezoidal barrier. Here the electron only has to tunnel through the thin triangle before it will be able to continue in the conduction band of the oxide. The barrier is triangular as it becomes thinner with increasing applied field. In other words, the electron will be able to continue through the dielectric once it has entered its first layer. With TAT, electrons use traditional tunnelling to hop between low energy wells. In other words, electrons tunnel short distances between sites that are ready to temporarily accept them (dangling bonds, etc). This process is referred to as 'elastic' when the electron does not lose energy between wells and as 'inelastic' when it does. All processes involving tunnelling see a current that is related to voltage by relationships with no temperature dependency.

Tunnelling is not always necessary for electrons to move through a dielectric. Thermal energy can excite electrons sufficiently to hop over an energy barrier and see them enter a conduction band. In other words, thermally excited electrons get across the dielectric using its conduction bands following thermal energy excitement. Electrons sometimes get trapped again into a new energy well. This is now Poole-Frenkel emission, which describes thermally excited electrons moving

across a dielectric by hopping from one energy well to the next by only temporarily entering the dielectric conduction bands. These two conduction processes have similar governing equations but thermionic emission has a higher dependency on temperature.

With Ohmic conduction, which generally occurs at high temperature and low field, one has a behaviour that is linear with respect to voltage, exponential with respect to temperature and occurs due to thermally excited electrons hopping from one isolated state to the next. DC ionic conduction describes the electron flow into the dielectric due to electron injection. The influx of electrons displaces negatively charged ions towards the anode leading to a TAT-driven conductive path. This is a process that sees current plateau over time and has a linear dependency with respect to voltage but an exponential one with respect to temperature. Space-charge-limited conduction leads to an electron flow into the dielectric because the dielectric is positively charged and electrons neutralise these charges. This time-limited process has a square dependency with respect to voltage.

Experimentally obtained current versus voltage curves of dielectrics can be used to understand what conduction process (or processes) occur in a device under test (DUT). These are generally plotted on special x-y plots with special axes. The choice of axes depends on the expected conduction process and will lead to a straight line if the expected process is the one occurring in the dielectric. From Table 1, the following can be extracted: conductions processes relying on Fowler-Nordheim tunnelling will lead to a straight line when the x-axis represents $\frac{1}{V}$ and the y-axis $log(I/V^2)$. Thermionic emission will lead to a straight line when the x-axis represents $\sqrt{V}$ and the y-axis $log(I)$. Poole-Frenkel Emission will lead to a straight

line when the x-axis represents $\sqrt{V}$ and the y-axis $ln\left(\frac{I}{V}\right)$. Ohmic and DC ionic conduction will lead to a straight line when the x-axis represents V and the y-axis I. Alternatively, linearity may be enhanced by using a log-log plot. Space-charge-limited conduction will lead to a straight line on a plot where the x-axis represents V^2 and the y-axis I. When multiple conduction methods occur at once, linear regions are expected to be found for the same region of the original IV plot on different modified axes plots. Extracting physical constants when multiple phenomena occur at once may lead to inaccurate results.

Looking at the literature, it is possible to find permittivity and barrier height data for co-sputtered SiO_x thin film in Ref. [2] extracted through IV curve fittings. Poole-Frankel hopping leads to a relative permittivity between 6.3 and 42, thermionic conductance leads to similar relative permittivities between 0.46 and 53.3. Trap assisted tunnelling leads to a barrier height between 0.22 eV and 0.85 eV. The TAT barrier extracted in Ref. [55] is however different and of 0.086 eV and 0.52 eV for LRS and HRS states respectively. The TAT barrier height extracted in Ref. [88] is 0.879 eV. Extracting barrier height through Fowler-Nordheim tunnelling in thin film silica leads to around 3.15eV [89]. However, this value is only about 0.35 eV or 0.66 eV in silicon rich silica in Ref. [90] and Ref. [55], respectively. It must be noted that some conduction methods, such as Fowler-Nordheim tunnelling, require the use of correction factors during extraction. These adjustments help alleviate error introduced by electrode shape [91] and material [89].

As discussed in the previous section (2.1.2), some show evidence of quantum point of contact conductance by showing changes in current by the half or full G_0. This is due to electron conductance occurring in narrow nanometre-scale filaments.

Here the conductive constriction is in the order of an electron fermi wavelength and therefore the amount of electrons going through per unit of time is quantised in multiples of G_0 ($= 2e^2/h$), which is the fundamental quantum of conductance. Electrons move through the constriction ballistically and without scattering. Evidence of current or resistance changes in multiples of G_0 or $G_0/2$ therefore implies the presence of a quantum point of conduction mechanism [36]. Changes in multiples of G_0 imply conduction on metallic (extrinsic) filaments. Conduction in multiples of $G_0/2$ imply conduction on semi conductive filaments. In Ref. [37] one can see steps of $0.5G_0$, G_0, $1.5G_0$, 2G0, $2.5G_0$, $3G_0$, $3.5G_0$, etc on poly-Si/SiOx/p-type-Si/Cr-Au devices. Here conductance appears to range between 0 and $8G_0$. Other works analysing quantum conductance in hafinia-based devices are found in Ref. [66, 92].

2.3 WHY SILICON OXIDE?

Filamentary intrinsic ReRAM exists based on several different oxides. As it will be show in section 2.4, these devices perform similarly. However, not all are equally compatible with the standard CMOS process found at most IC foundries. CMOS compatibility translates into being able to cheaply integrate ReRAM storage on pre-existing silicon chip designs. This is what will certainly speed up the adoption of this new technology. SiO_x is a material that is commonly used in many CMOS devices currently fabricated, as well as being one of the most commonly available materials on the planet [55]. Finally, SiO_x may be used in its amorphous form leading to high device yields.

Other research groups [43, 93-95] have previously talked about SiO_x-based devices, which require a vacuum to operate and do not rely on filament switching.

Similarly, there are also extrinsic SiO_x-based devices [96], which rely on metallization. UCL's devices [2] are the first to be intrinsic SiO_x-based and are appealing thanks to in-air operation and high scalability due to their filamentary nature.

2.4 RELEVANT LITERATURE

Characterisation of manufactured devices is mostly done electrically. This means that experimental devices are generally studied through regular voltage sweeps, long pulses with current compliance, or fast pulses (which may not have current compliance). This last method also appears in the literature to use these devices as emulators of certain neuron functions. However, characterisation is not always purely electrical; it also involves the observation of physical changes. A section on literature on $^{18}O_2$ exchange in oxides has been included as this will provide a background for the interpretation of some results.

2.4.1 Regular Voltage Sweeps

As a first step, most ReRAM devices (intrinsic, extrinsic, non-filamentary, filamentary, etc.) are characterised using a voltage sweep. A schematic representation of fictional typical sweeps is show in Figure 4a for unipolar devices and in Figure 4b for bipolar ones. Here a voltage is swept from 0 V to a certain positive or negative value and then, typically, back to 0 V. When electroforming, or setting, a positive jump of at least an order of magnitude in current represents an abrupt drop in resistance showing that a conductive filament has electroformed or been restored at a specific voltage. Sometimes, more than one jump may occur in place of just a large one, showing potentially useful intermediate resistance states (multi-state memory, neuromorphic applications, etc.). These jumps show that this

family of devices is highly non-linear. If current compliance is used during a sweep, a flat line will generally follow the last one of these jumps. This means that the voltage source has dropped its voltage in order not to exceed a certain current. Sweeping backwards towards a voltage of 0 V shows that the device resistance has successfully changed. This second curve appears different from the first and it allows for direct observation of the change in state. What happens during a reset sweep is similar, except that now the expectation is a jump (or many, again, useful gradual jumps) that shows a decrease in current. This corresponds to the filament breaking and therefore a reduction in the current flow. During this phase, current compliance is never used as there is no danger; resetting the filament automatically limits current due to the expected increase in resistance. Reset current generally exceeds electroforming current in unipolar devices; the same is not always true for bipolar ones. Also, during reset devices show highly non-linear behaviour. During some sweeps, it is possible to observe the stochastic nature of these devices: resistance does not always increase or decrease steadily before switching. Often, resistance increases and decreases seemingly randomly. This is due to competing setting/electroforming and resetting processes occurring within the device and opens doors to further possible neuromorphic applications as discussed in 2.4.5. The use of sweeps is therefore very common to find out at which voltages a new device will switch and observe the stability of its highly non-linear behaviour. Switching voltages determined using this method may then be used to program devices using short pulses rather than long sweeps. Figure 5 [83] shows the discussed aspects of commonly used IV curves.

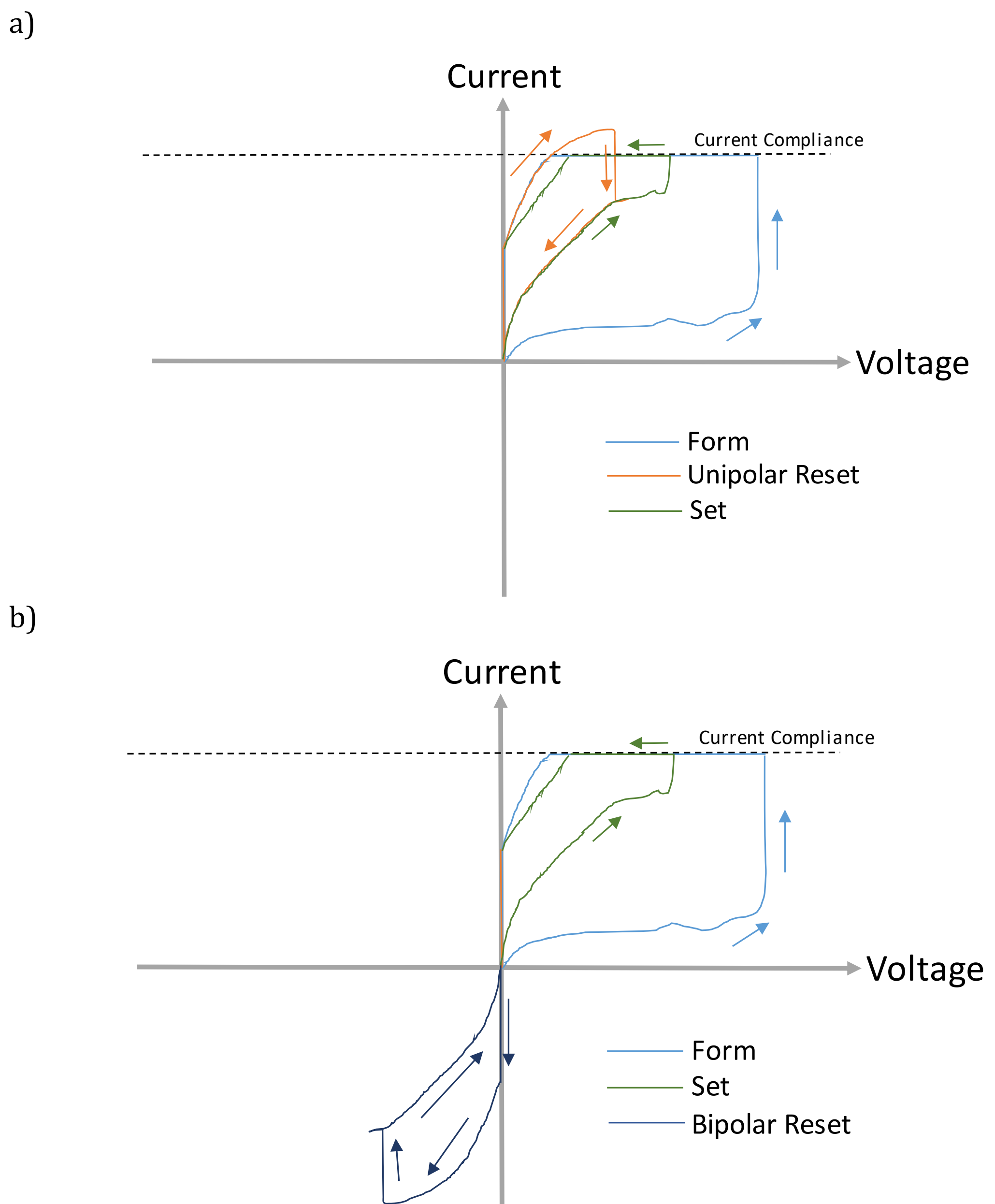


Figure 4: Schematic representation of fictional typical IV curves for unipolar (a) and bipolar (b) devices. Electroforming and setting are limited in current through current compliance. Unipolar resetting exceeds this current compliance. Bipolar resetting may or may not exceed the previously used current compliance. Any of these curves may also occur with inverted polarity.

Table 2 shows a summary of some experimental (and therefore non-commercial) intrinsic and extrinsic ReRAM devices found in the literature that were characterized using IV sweeps. It shows that this type of characterisation is broadly used, gives the idea that many different technologies exist, that low setting and resetting voltages are common as they ease integration into current CMOS

technology, and that endurance in the thousands using IV sweeps is possible. High electroforming voltages are not a concern as this step may be carried out at the foundry using off-chip programmers. High non-linearity is present in all devices covered.

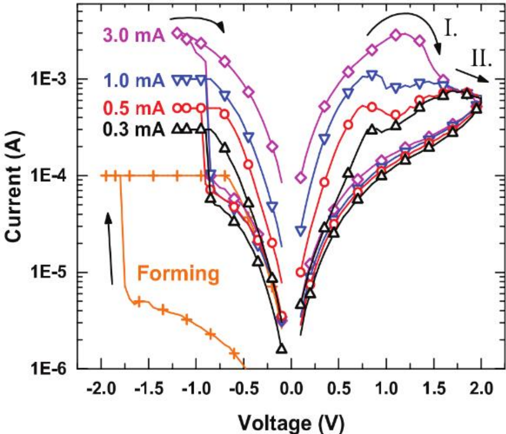


Figure 5: IV curves of a hafinia-based device found in [83]. With this bipolar device, current compliance is used during electroforming and setting (marked as 3.0 mA, 1.0 mA, 0.5 mA, and 0.3 mA). Different levels of current compliance lead to different resistance states. While resetting as well as sweeping the LRS some competing processes may be observed. Permission obtained from publisher.

2.4.2 Voltage Pulsing with Current Compliance

Through repeated voltage sweeps, it is possible to determine the switching voltages of a device. Therefore, to achieve faster switching, it is possible to directly apply these determined voltages, along with the necessary current compliance. As already mentioned, following a 'programming pulse', a smaller 'reading pulse' will be used. This second pulse will be at a voltage low enough not to change the state of the device and high enough to allow for the reliable sampling of current. This leads to a method to quickly cycle devices through states and test their cycling endurance. Sampling the resistance state at certain time intervals is, instead, used to learn about device reliability over time. This shows how stable a programmed device remains over time. Pulsed switching is therefore possible following initial characterisation using sweeps.

Layers	Technology	Switching Voltages	Notes
ITO/SiO_x/p-type Si devices	Intrinsic	Setting and resetting at ±5 V	UCL's previous work [55]. Similar curves found in [59] and [36] also to show quantized conductance.
Ti/TiO_x/Au	Probably intrinsic	Set and rest occurring within ±4 V sweeps. 10 mA CC used for safety.	Bipolar. No initial electroforming. Sweeps used to test endurance but not fully reported. [81]
TaO_x	Probably intrinsic	Set and reset at ~±0.7 V. CC tested between 0.2 mA and 1 mA affects these voltages.	~3150 cycles reported but not shown. [97]
Pt/$SiO_{0.73}$/Pt	Intrinsic	Set: 1.7 V Reset: 0.7 V	Sweeps shown over 30 cycles. [60]
Pt/HfO_2/TiN	Intrinsic	Eelctroform: ~-1.75 V Reset: ~-0.75 V Set: 1.75 V	Active layer is 5 nm, in a vertical stack of two devices. Current compliance used from 0.3 mA to 3 mA [83].

Table 2: Summary of some literature that uses IV sweeps to characterise devices.

Given today's characterisation systems [98], these pulses are usually in the millisecond-domain and current compliance circuits kick in within nanoseconds (more in section 2.4.3), preventing thermally driven phenomena from burning devices to an unrecoverable state. This also means that these slow pulses with current compliance may never reach the specified voltage and be undefined in magnitude. These pulsed results will provide complementary information to what was learned through sweeps.

The literature shows many devices that have been characterised using pulses with current compliance. Some of these devices are summarised in Table 3. Out of this characteristic pool of devices, endurance is shown using cycles but resistance is often sampled only after a certain number of switching cycles (Figure 6) [44]. All

devices are oxide-based and intrinsic except one, which relies on metal dopants in the active layer. When long (microseconds or longer) pulses of a few volts are used on devices with a low resistance LRS (<1 kΩ), it can be assumed that devices are programmed using current compliance; otherwise, the thin active layers would most probably melt down. It is common for published papers not to report whether current compliance is used while setting or electroforming.

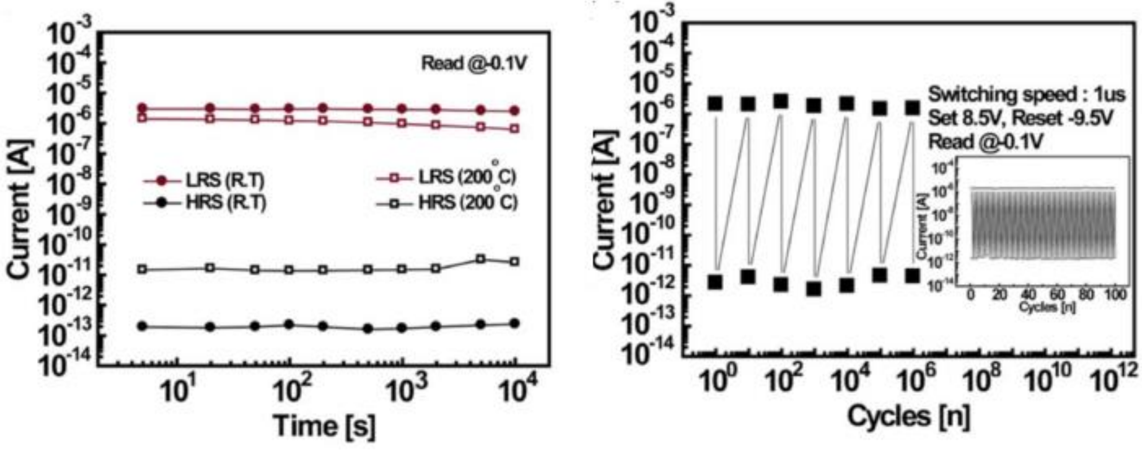


Figure 6: Example of endurance and reliability testing (left and right, respectively). The characterisation method of pulsing with current compliance was used in these examples. Taken from [44]. Publisher permits reproduction: © IEEE 2011.

2.4.3 Fast Voltage Pulsing

As stated above, devices can also be switched with short pulses generated by a programming device that does not make use of current compliance. This is possible because the pulse is short enough to avoid disruptive thermal behaviour. This allows for a smaller programmer as fewer transistors are needed. The drawback of this technique is a need for prior knowledge of electroforming or setting switching time. However, some devices that do not require current compliance at all are also found in the literature [64]. These seem to have sufficiently high resistance, which leads to low and safe currents.

Layers	Technology	Switching Pulses	Notes
Pt/TiO_x/Pt	Intrinsic	900 ms ±6 V.	Reading using +1 V pulses. -1 V pulses greater variation between states. Hence, diode-like rectification. Five cycles shown. [78]
Mo/SiO_x/Pt	Intrinsic	±10 V; 1 ms and 1 s also used but harder to reset.	Current compliance only used on set. 100 cycles all shown. Cycling up to 10^6 cycles is shown with resistance only sampled 14 times. Figure 6. [44]
Ir/TiO_x/TiN	Intrinsic	1 µs, voltage depends on previous state.	Complex negative feedback programmer. 0.1 V used to read. Current compliance probably used given length of pulses. 500 cycles shown with four different states. [38]
Pt/SiO_2:Pt/Ta	Hybrid Extrinsic/Intrinsic	10 µs pulses	More than 10^7 cycles but not all shown. Current Compliance probably used due to pulse length. [45]
Pt/$SiO_{0.73}$/Pt	Intrinsic	Set: 1.7 V, Reset: 0.7 V with 1 or 10 mA current compliance. Pulse length is not specified.	Endurance > 100 cycles. Read at 150 mV. Unusual: has inverted voltages and x<1. [60]

Table 3: Summary of some literature that uses pulses with compliance to characterise devices.

Fast pulsing is commonly used to extract device performance metrics such as endurance (number of cycles), data retention and the type of programming pulses required. Looking at Table 4, it appears that when endurance is reported as greater

than 1000 cycles [83], device state is not sampled after every switching event [42, 43, 83, 99]. This leads to have doubts as to whether the device switched following every programming pulse. Therefore, claims as high as 10^8 cycles [99] must be looked at carefully. Programming pulses are generally 100 nanoseconds or less [42, 55, 64, 83, 99] with one device using 100 picosecond pulses [45]. One device uses 100 nanosecond set pulses but 1 microsecond pulses to reset. This is not surprising as resetting is a self-limiting process with no risk of device meltdown. Programming voltages are spread over a wide range: from 1.5 V to 20 V. Some devices do not require an electroforming step [81, 83]. Device reliability is generally measured at above room temperature and for at least 10^4 seconds [42, 64, 83, 99]. It must be noted that devices based on SiO_x are the most similar to those developed at UCL.

2.4.4 Macroscopic Structural Changes in Oxide Following Electrical Stress

Ever since the 1930s [28, 100], reports have emerged hypothesising that oxygen species must be emitted from devices during their operation. This is because the positive electrode (or anode) sometimes changes form from smooth to deformed. This is presumed to reduce device endurance.

Changes in top electrode surface appear on devices during operation in [25, 60, 76, 82]. These are devices based on binary-oxide active layers (such as SiO_x, SiO/V_2O_5, and TiO_2) and different metal top electrodes (such as Cu, Ti or even inert Pt). In all cases these deformations (Figure 7 [60]) are empirically attributed to oxygen emission. A discussion is found in [5] stating that when the top electrode cannot contain oxygen, this element creates deformed regions to escape. It is concluded that it is important to prevent oxygen from escaping.

Layers	Tech-nology	Switching Pulses	Notes
Double HfO_x layer	Intrinsic	±3.5V, 100ns	Endurance: 10^8 switching cycles but only 4 samples per decade, Retention: 10^5 s at 125°C [99].
Poly-Si/SiO_x-/Poly-Si	Intrinsic	Set: 13V, 50ns Reset: 6V, 100 ns	Endurance: 10^5, only 300 cycles sampled Retention: 10^5 seconds. Repeated read pulses to ensure reliable reading (?) [64]
W/SiO_x/N:Si	Intrinsic	Electroform: 11V (with CC) Set: 4.5V, 100 ns Reset: 12V, 1 µs	Vacuum operation only, Endurance: first 100 cycles shown, 10^4 cycles only sampled at decades. [43]
ITO/SiO_x/p-type silicon	Intrinsic	Set: 8V, 90ns Reset: 12V, 90ns	Reading: 1V repeated pulses Endurance: 5 cycles [55]
N-type:Si/SiO_x/p-type:Si	Intrinsic	Programming : ±20 V	Read: 2V [55]
TaO_x-based	Intrinsic		Fast pulsing used to enhance device endurance by reducing filament temperature and size [80]
Pt/HfO_2/TiN	Intrinsic	set: −2.5 V, 30ns reset: 3.8 V, 30ns or reset 4.8V, 10ns	Read: 0.5V or 0.3V Endurance: 1000 cycles shown, sampled every time; 10^7 sampled every 1000. Retention: ~7 hours at 400K. No need to electroform [83]
Ti/TiO_x/Au	Intrinsic	Set: 4V Reset -5V Pulse length between 10 and 1000 ns change log(ΔR)	Read: 1V. Endurance: 100 cycles sampled a few times. Top electrode as oxygen reservoir. No electroforming. [81]

Table 4: Summary of some literature that uses fast pulses to characterise devices. Devices based on SiO_x are the ones most similar to UCL's devices.

Reports of oxygen movement are not limited to the observation of surface deformation. [75] discusses experiments with devices based on NiO thin films in a Conductive Atomic Force Microscope (CAFM [101]). It is observed that switching is hindered in vacuum showing a need for environmental oxygen. Interesting behaviour is shown when the CAFM chamber is filled with $^{18}O_2$ gas. This new

pressurized environment allows for switching to occur again. SIMS analysis of the active layer, however, shows a decrease in ^{16}O and an increase in ^{18}O when a positive bias was applied. ^{18}O only appears to penetrate 2 nanometres into those regions that were positively biased. No change was evident when a negative bias was applied. This implies oxygen dependency in filament formation and disruption.

Ref. [102] discusses how O^{2-} oxygen species might be created during electroforming and setting because of electron injection. This implies that electric field without current flow may not be sufficient to electroform or set a device. The study is purely theoretical. From a more experimental point of view, some studies have been carried out on oxides and other materials showing that superoxide (O_2^-) may be produced by oxides under particular chemical or thermal stimulation [103-105].

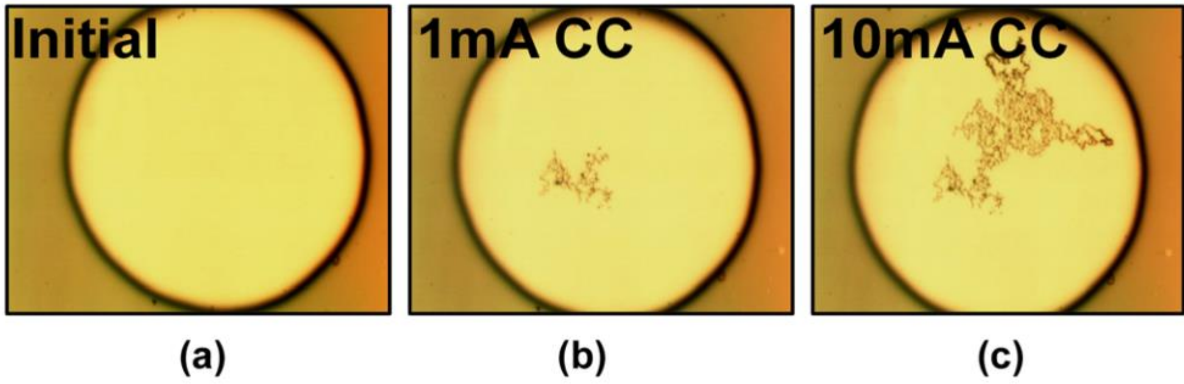


Figure 7: Optical microscope images of deformed regions on a (a) prisitne Pt/SiO$_x$/Pt device, on a (b) device electroformed with 1 mA current compliance and on a (c) device electroformed with 10 mA current compliance. These show that a gas (probably oxygen) is emitted during operation of the memory cell. Taken from Ref. [60] with permision obtained from publisher. Oxygen therefore appears to be key to the functioning of oxide-based ReRAM devices. Structural changes occur due to its supposed release, performance decreases in vacuum and oxygen appears to be absorbed while cycling. These oxygen emission hypotheses have been eventually confirmed by this work, which is included later on in this thesis [58].

2.4.5 Neuromorphic Behaviour

ReRAM devices have applications outside of computer memories. Given that these devices have been shown to be able to enter multiple intermediate states, post-binary or semi-analogue applications seem only natural.

One of the above applications is the direct emulation of some of the functions of a biological neuron synapse. It appears possible to emulate potentiation and depression. These correspond to a gradual increase and decrease of a device resistance state. In addition, the forgetting process may be emulated along with plasticity or self-reconfigurability. The latter allows for defective ReRAM devices to be excluded from a neuronal network. This leads to the lack of a need for perfect and deterministic devices in the design of useful machines. It is the stochastic noise present in devices that keeps systems going by preventing them from getting stuck in minimum or maximum states. It is the above device behaviours that are referred to as neuromorphic.

The integration of these devices into networks can lead to electric circuits which emulate the problem-solving capabilities found in mammalian brains leading to low power dissipation, small silicon footprint and high parallelism. Modern sequential binary computers are not efficient at reproducing these types of brain behaviour which are useful in applications such as image processing [48-50]. As an example, a supercomputer today can emulate a rat's or cat's brain but only with enormous power consumption compared to the one of the brain being emulated.

In the literature are many devices that can gradually change their resistance state in response to short programming voltage pulses [49, 50]. It is often stressed that devices should be able to enter as many states as possible to ensure accurate

emulation of synapses [52]. An example of this is the $Pt/HfO_x/TiO_x/HfO_x/TiO_x/TiN$ device proposed in [52]. This device can be set to thousands of different states using repeated identical stimuli with durations of 10 nanoseconds in a potentiation process. This is thanks to the final resistance of the device being the sum of those of the many active layers; each active layer may be in a different resistance state. This allows for very simple programming of circuits that emulate neuronal behaviour involving synapse potentiation and depression. This device [52] process variation is discussed and it is concluded that it is within tolerance for a functional neural network. The device is therefore deemed reliable at the simulation level. Also, other devices use sub-100 nanosecond pulses and low voltages (~±3 V, ±1.2 V) to gradually change state [49, 50].

In the literature, one also finds devices that may be used to emulate the forgetting process found in the mammalian brain. The $Pt/WO_{3-x}/Pt$ device found in [51] 'forgets' the state it had entered much more quickly when it is programmed using short rather than long pulses. The effect of a 0.1 millisecond pulse is 'forgotten' within tens of minutes, while that of a 0.5 second pulse is 'forgotten' within a few hours. Devices with poor retention, as opposed to long retention devices used in digital memory applications, may be very useful in brain-emulation application involving memory emulation. Short term memory emulation is needed in order to free storage of information that is no longer important [106]. Having devices that forget within a wide range of time scales is sought after in neural networks to achieve more realistic emulation of the mammalian brain [51].

ReRAM two terminal devices are easy to interconnect in a crossbar 2D array geometry in an effort to assemble useful networks. [50]. A circuit that relies on

ReRAM to emulate the dynamics of real synapses is proposed with 256x256 devices [48]. Ref. [107] talks about a 12 by 12 Pt/Al_2O_3/TiO_{2-x}/Ti-based ReRAM array that emulates synapses and is trained using a complex algorithm to recognise 3 by 3 input patterns. Ref. [108], instead focuses on a network that relies on memristors, transistors and resistors and is able to learn input patterns.

Simply described, mammalian brains are neural networks that can remap themselves. This may happen as the consequence of activity such as learning or recovering from brain damage [109, 110]. The basis of this self-change phenomenon is found in the neuronal synapse which is what allows the interconnection of neurons into networks. This appendix of the neuron behaves differently depending on its stimulation and behavioural histories. Synaptical behaviour depends on that the synapsis in question has previously 'learned'.

Since ReRAM devices are devices that are capable of certain synaptical behaviours, it is important that they also emulate plasticity through a learning capability. These devices do not return to their initial state but rather enter states that depend of the history of their previous states [50]. This allows for networks involving these devices to self-exclude defective components or regions (a form of seal-healing) and learn future behaviour from the past. This is the case in the pattern recognising array found in Ref. [107]. Unfortunately, through the literature, one can see that it is still not possible to assemble large functioning networks due to device interconnection challenges. While discussed examples make sure of bipolar devices, nothing seems to prevent the use of unipolar devices for the emulation of synapses (potentiation, depression, plasticity) and memory behaviour. Unipolar devices,

however, have higher power consumption on reset and may be more difficult to control due to all programming voltages having the same polarity.

2.4.6 Labelled Oxygen Exchange

The oxide-based ReRAM family of devices has been described, without positive proof, to rely on oxygen movement. It appears that such movements lead to unwanted surface deformation. It is interesting to study oxides to understand under what thermal and timing conditions chemically bound oxygen escapes these materials. Similarly, it is interesting to consider under what conditions atmospheric oxygen is incorporated into oxides.

The above processes fall under the practice of oxygen surface exchange. This is generally carried out in labelled atmosphere containing a high percentage of $^{18}O_2$ and $^{16}O^{18}O$. Diffusion within the oxide lattice then allows for a concentration gradient to form in the oxide [111]. Regrettably, literature in this field discussing amorphous silica appears non-existent as focus is given to crystalline oxides and perovskites [112-116]. Temperature, as an exchange parameter appears to be very important.

Temperatures as low as 450 °C appear to be used with perovskites [113, 114]. Binomial oxides such as Bi_2O_3 [112], ZnO, TiO_2, and ZrO_2 [115] also appear to undergo exchanging with low temperatures (500 °C). These same materials, along with a 40 nanometre-thick layer of silica [116], however also appear to undergo exchange with temperatures as high as 1000 °C.

Due to probable saturation, the thin silica layer found in [116] acquires an amount of oxygen that does not change whether the exchange is carried out for 2 or 8 hours. Only the depth of oxygen acquired into the silicon bulk is affected. The

binomial oxides found in [112, 115] exhibit differ exchange time (between 4 and 30 minutes) depending which phase the oxide was in. Perovskites [113, 114] undergo exchange to a depth of 250 micrometres from the surface over 5 to 9 minutes.

It is hard to draw precise conclusions from the literature about the exchange time of amorphous silica. However, a temperature between 450 and 1000 °C for a few minutes might well exchange most of the oxygen present in a silica thin film with the one present in air. This may be particularly true in the case of amorphous silica as oxygen-silicon bonds should be weaker than in the crystalline structure.

2.5 CONCLUSIONS

Recent literature shows that a high degree of interest exists in ReRAM devices. Most devices are studied using voltage sweeps to determine their switching voltages and in some cases their endurance. Pulsed programming is used mainly to quickly test endurance after many cycles. Pulsed reading at a lower voltage is used in endurance as well as retention tests. Long pulses usually involve the use of current compliance when electroforming and setting in a way similar to voltage sweeps. Otherwise, once resistance drops due to filament creation, current will increase, destroying the filament altogether. Current compliance is not used in those few devices that have a high built-in resistance, which acts as a current limiter. Short pulses also seem to be commonly used to test devices. These pulses often do not make use of current compliance in the programmer. If a pulse is sufficiently short, it should only lead to a field-driven phenomenon and not induce more Joule heating than necessary. The duration of the pulse is critical, as it should be long enough to electroform a filament but short enough not to melt it. However, in some cases it

also appears that current might be limited by high device resistance or unreported, but used, series resistors.

The stochastic properties found in some of today's ReRAM devices are not necessarily a downfall. Devices that gradually 'forget' their programmed state emulate the behaviour of biological neurons. In addition, if devices do not end up returning to their initial, pre-programing resistance state, they may instead show neuronal plasticity. This important phenomenon allows for reconfigurability or exclusion of malfunctioning devices. This translates into the possibility of creating reliable neuronal networks using unreliable ReRAM devices characterised by the right amount of stochastic noise. For certain classes of problems (such as those involving high parallelism), these will lead to useful non-deterministic machines using a few Watts of power as opposed to deterministic computer-based implementations that consume megawatts and are slow.

It is presumed in the literature that oxygen species move within oxide-based ReRAM devices and are released during operation. This is thought to lead to detrimental physical changes as well as depletion of oxygen, which hinders endurance. It is therefore important to further analyse such release to reduce or eliminate the phenomenon in future device iterations. Similarly, there is a case in which evidence of oxygen absorption emerges. However, this is not thoroughly investigated and only appears to occur in the first few nanometres of the active layer.

One also learns that oxides placed in an $^{18}O_2$-rich atmosphere at high temperature will start exchanging their $^{16}O_2$ atoms for $^{18}O_2$ ones. This creates an ^{18}O gradient into the oxide surface. Taking this process into consideration will be useful

when analysing devices cycled in $^{18}O_2$-rich atmosphere as these may very well heat up during reset and undergo a similar heat-driven oxygen exchange process. Given the alleged importance of oxygen in intrinsic ReRAM devices, this will help obtain clues to the physics of electroforming and switching.

3 EXPERIMENTAL METHOD AND RATIONALE

It was interesting to study and characterise filamentary unipolar ReRAM devices previously made at UCL using three approaches. Tests were started off by looking at device response to electrical stimuli. While most electrical tests were run in air, some were run in situ in a high vacuum, and some other ones in modified atmosphere. Additional experiments used in-vacuum tests to study the detection of oxygen emission. Isotopically labelled modified atmosphere tests were also used to study atmospheric oxygen absorption. Tests were always carried out starting on unformed (or pristine) devices, which could be recognised by their very high resistance.

3.1 SAMPLE PREPARATION

The unipolar device type that used throughout this work was based on a TiN/SiO_x/TiN structure. A cross section of this device is found in Figure 8, which is adapted from [2]. Many independent devices are present on each wafer as each device is simply made up of a small top electrode and active layer and bottom electrode which cover the entire wafer. Top electrodes are available in squares ranging from 400 micrometres per side down to 10 micrometres per side. Conductive filaments will electroform between electrodes within the SiO_x layer.

UCL's wafers were fabricated on P-type Si wafers with a 4 micrometre thermal SiO_2 layer. The bottom TiN electrode (100 nanometres thick) was deposited on top of this insulating layer using magnetron sputtering at room temperature. The active SiO_x (37 nanometres) layer was also co-sputtered from Si and SiO_2 confocal anode targets. This was carried out using an Ar beam with the substrate at 250°C, leading to an excess atomic silicon fraction of 11% in the SiO_x. This corresponds to a

stoichiometry of about $SiO_{1.3}$. The top contact layer was deposited using the same method used for the bottom electrode, making the device symmetric. Annealing was then carried out at 600 ˚C in Ar gas. The bottom TiN electrode was exposed in parts of the wafers to allow direct access from the top. When a wafer was cleaved and this contact lost, a new Ohmic ground contact was created by ion beam milling. For the purpose of this study devices fabricated under different conditions were not available.

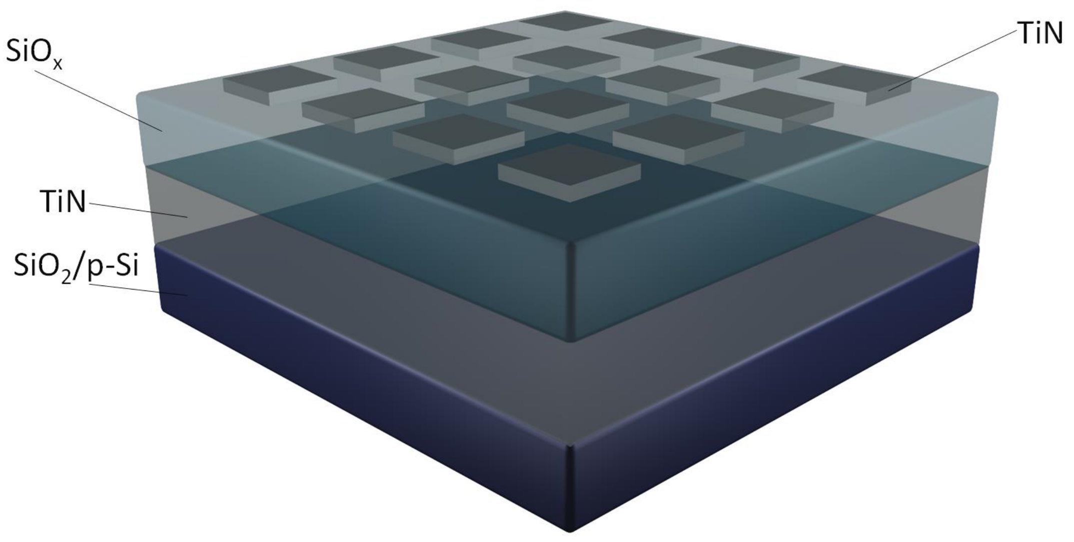


Figure 8: Schematic of a TiN/SiO_x/TiN wafer. The TiN layers have a thickness of 100 nanometres while the active SiO_x layer has a thickness of ~37 nanometres. The SiO_2 layer has a thickness of ~4 micrometres and is followed by an additional bottom layer of P-type:Si that is not shown. Adapted from [53].

3.2 Experimental Tools

3.2.1 Hardware

Characterisation of devices was mainly carried out electrically; device stimulation was always electrical but observations did not only come from electrical readings. A wide array of equipment was used, ranging from standard voltage sources with current measurement capabilities to high vacuum chambers with in

situ manipulators and SIMS detectors. Modified atmospheres, including isotopically labelled ones, were also used. Samples were kept clean using a sonicating bath and standard acetone, isopropyl alcohol (IPA), and deionised (DI) water cleaning cycles.

3.2.1.1 Electrical Measurement Equipment

A Keithley 4200-SCS (Semiconductor Characterization System) [98] was used for in-air tests. Source Measure Units (SMUs) installed in this instrument were used to carry out tests involving sweeps with current compliance and pulsing with long pulses (~14 microseconds) with current compliance. Figure 15 shows output traces of this device. 4225-PMUs (pulse measure units) were used to carry out faster pulse tests (tens of nanosecond range) without compliance. Figure 16 shows output traces of this module. This equipment was connected to a Signatone 1160 probe station, located in a shielded box, through triaxial connectors and cables when SMUs were used. This was not possible with PMUs as their output is in SMA format and required a single shield cable up to the probe station where BNC to Triax adaptors were used. Samples were left floating. All cables were of the same length to minimise issues involving transmission-line propagation delays. Vacuum held micropositioners (Signatone SP-150) were used inside the probe station. These controlled Signatone 'screw-lock' probes with 1 micron tungsten needles (Signatone SE-10T), which contacted devices under test. An optical microscope with a light source was used to aid the positioning of needles. The probe station stage was left electrically floating, or samples were otherwise placed on insulating glass. A Tektronix MDO4104-6 voltage oscilloscope [117], with a sampling rate of 5GS/S (200 picosecond resolution) was also used; the signal was tapped from one or two channels using BNC tees. With this instrument, it was also possible to measure current through a

DUT by measuring the voltage across a series 90 Ω resistor, which was inserted using a PCB protoboard with soldered BNC connectors and resistor. Signatone and Linkam heating stages were available for the above in-air measurement set-up.

In order to bias the sample for the in situ testing that follow, a Keithley 6430 Sub-Femtoamp Remote SourceMaster® [118] was used. This device outputs signals as short as a few tens of milliseconds and has current compliance capabilities. This source was chosen as it is much more portable than the 4200-SCS.

Except for regular sweeps on the 4200-SCS, all Keithley equipment was controlled using a Windows PC through a National Instruments 'GPIB-USB-HS' interface.

Given that current compliance on both the 4200-SCS and 6430 appears to protect devices from melting down, it probably kicks in within a few tens of nanoseconds (these details have been withheld by Keithley). Please refer to the fast pulsing results section (4.1.3) for confirmation of this.

3.2.1.2 High Vacuum chamber with Detectors and in situ Probes

A different set-up was used for in situ vacuum experiments: a standard Phillips vacuum chamber on which many devices are installed (part of this configuration is found in Ref. [119]) . The inside of this chamber is made of aluminium which never gets baked at high temperature and is about 15 years old. This means that organic contaminants are to be expected. Adsorption and desorption of gases are also to be expected in this chamber [120, 121]. Within the chamber, one could find present two Kleindiek micromanipulators (rated at 100 mA and 100 V) [53] with 1 micrometer tungsten whiskers (Picoprobe PT-14-6705-B [122]) and BNC connectors, which allow for electrical biasing of samples. A Ga ion

gun [119] is found in the top part of the chamber and normal to the sample, which it bombards with Ga^+ ions. This produces surface ions and electrons that are used for imaging and analysis. A Hiden Analytical model EQS secondary ion mass spectroscopy (SIMS [123, 124]) detector with ionizing filament for residual gas analyser (RGA) mode is also present and allows the detection of ionic and neutral emissions from samples as well as to carry out elemental depth profiles (when used in conjunction with the Ga^+ ion beam). An FEI quadrupole charged particle detector operating as a SIMS [119] allows to produce elemental depth profiles and elemental maps (when used in conjunction with the ion beam). A platinum gas source (50% Pt, 50% organic) is also to be found and it allows to deposit thin layers of Pt on samples through ion beam assisted deposition. A channeltron is also present; it detects charged particles and is mainly used for electron and positive ion imaging. Figure 9 shows a (a) photo and a (b) diagram of the set-up.

This setup made it possible to carry out experiments involving the detection of charged and non-charged oxygen species as samples were being biased. This same setup also acted as an ion beam microscope, a SIMS, an in situ RGA and a Pt deposition tool. More details on its use for surface characterisation follow in section 3.2.1.5.

a)

b)

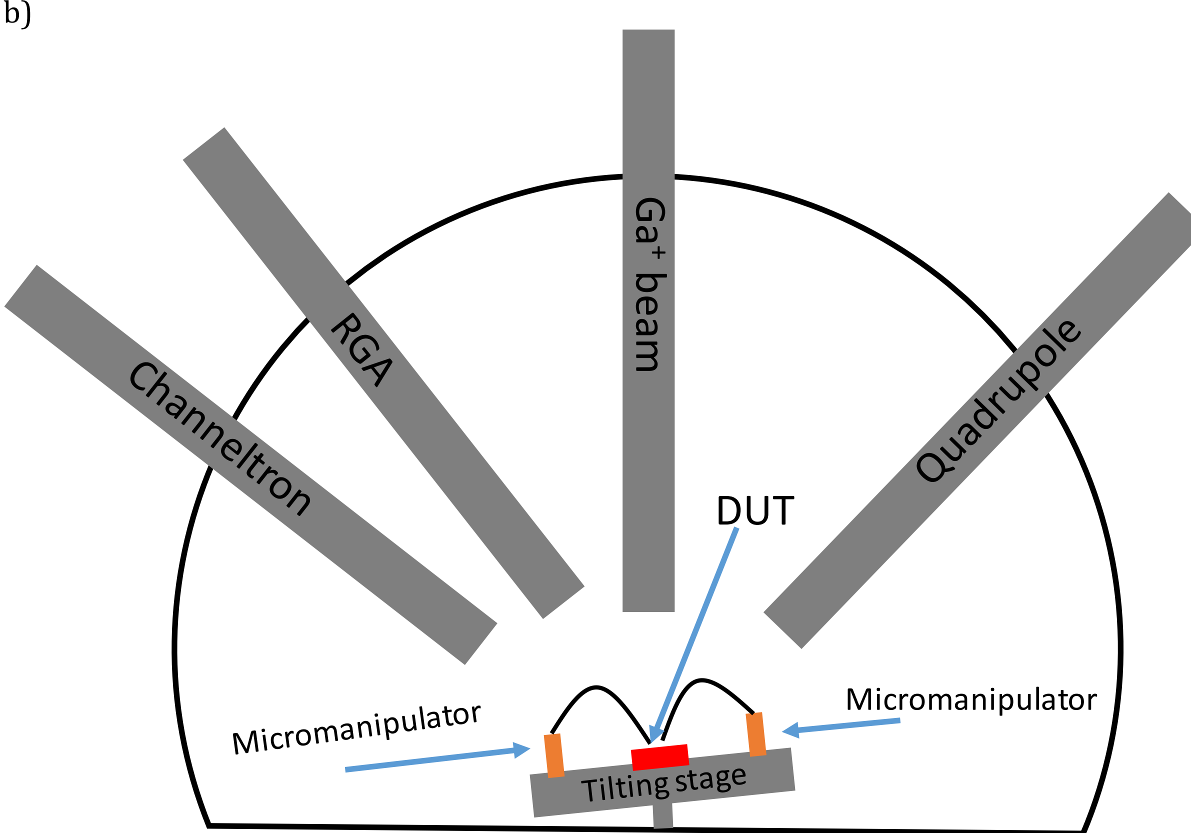


Figure 9: Vacuum chamber: a) Photo of the vacuum in situ testing setup. One of the $TiN/SiO_x/TiN$ wafers can be seen. b) Schematic diagram of the vacuum chamber set-up.

3.2.1.3 Oxygen Exchange Rig

A different setup was used to cycle devices in $^{18}O_2$-rich atmosphere. Equipment commonly used to carry out high temperature $^{18}O_2$ exchange of

previously oxidised materials (Figure 10) [125] was used. This rig was originally designed to help carry out slow high temperature, non-electrically biased $^{18}O_2$ exchange in oxides and it was used in a very different way.

a)

$^{18}O_2$ reservoir
Leak Valve
RGA
$^{18}O_2$ valve
Test tube
Main Chamber
Turbo pump
Keithly 6430 and PC
Sample
Main vacuum valve
Roughing pump

b)

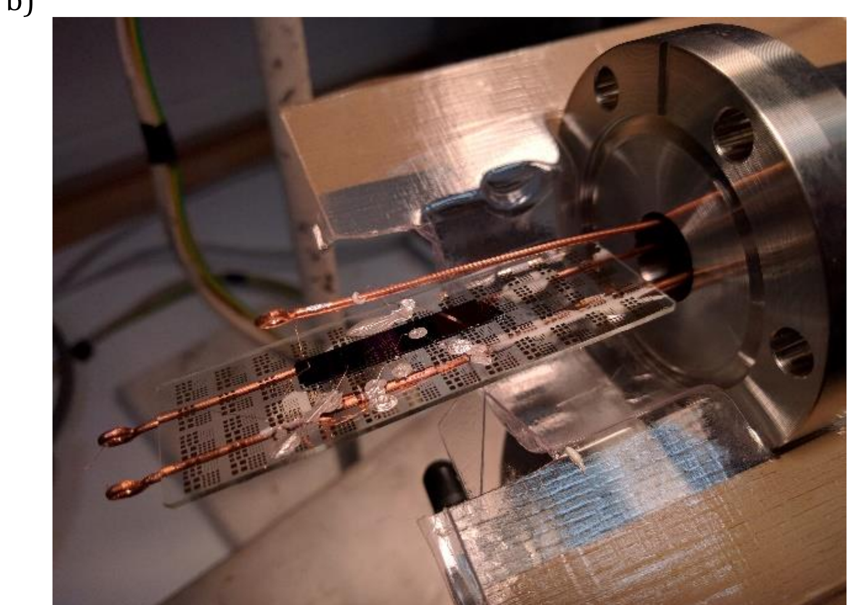


Figure 10 : Exchange rig used: a) diagram, b) image of the feedthrough setup used to carry out switching in $^{18}O_2$-rich atmosphere. This exchange rig was only used for $^{18}O_2$ based experiments. A different set up was used to carry out switching in different atmospheres.

As shown in Figure 10, this rig has both a roughing and a turbo-molecular vacuum pumps, which allows for a high vacuum to be achieved in a main chamber and a connected test tube. The test tube provides electrical connections to the outside world through an electrical feedthrough with four connectors. The main chamber and test tube can be isolated through a valve but some of the atmosphere

present in the test tube can be reinjected into the main chamber through a leak valve. This allows the installed Hiden RGA to characterise the test atmosphere present in the tube. This atmosphere is injected into the test tube from molecular sieve canisters containing a virtually 100%-oxygen mixture rich in $^{18}O_2$. Liquid nitrogen was available to recover the mixture into the reservoir canisters by chilling them and reducing their internal pressure; this difference in pressure allowed for the oxygen mixture to be sucked back into storage. The mixture oxygen purity is guaranteed by the molecular sieve found in the canisters, which will absorb most non-oxygen molecules. An $^{14}N_2$ cylinder (not shown in Figure 10) was available for the flooding of the high vacuum sections of the rig; this could be done using valves. Vacuum chambers are often flooded this way to speed up the process and avoid unwanted moisture generally found in air.

To make contact to the sample the following was used: a microscope glass slide, silver paint, high vacuum bicomponent glue, and thin copper wires harvested from an ultra-flexible cable. The glass slide provided a non-conductive stage of the correct size for the test tube, silver paint acted as conductive glue (e.g. to connect wires), thin wires were used to contact top electrodes thanks to their convenient spring constant and narrow size. Finally, high vacuum glue was used in large amount to attach the microscope glass slide to the vacuum feedthrough. A photograph of the setup may be found in Figure 10b. Voltage characterisation was carried out by connecting the Keithley 6430 to the feedthrough input using ~15 centimetre unshielded copper wires.

3.2.1.4 Sealed Gas Chamber with In Situ Electrical Biasing

This setup permitted testing devices in a modified atmosphere not containing expensive $^{18}O_2$, but rather cheap gases that could be left flowing for many hours. Here, samples were placed in a sealed Linkam probe station (Figure 11 and Figure 12) with tungsten needles (the same as those used in the SIMS chamber) and gas was forced through this small chamber's gas inlet. The output of the small chamber was connected to a gas bubbler, which contained oil submerging the gas input glass pipe by several millimetres. This ensured positive chamber pressure and no atmosphere return. After a certain amount of time the used gas would displace most air present in the chamber and create a controlled switching environment. $^{16}O_2$ and $^{14}N_2$ gas tanks were available along with standard pressure regulators.

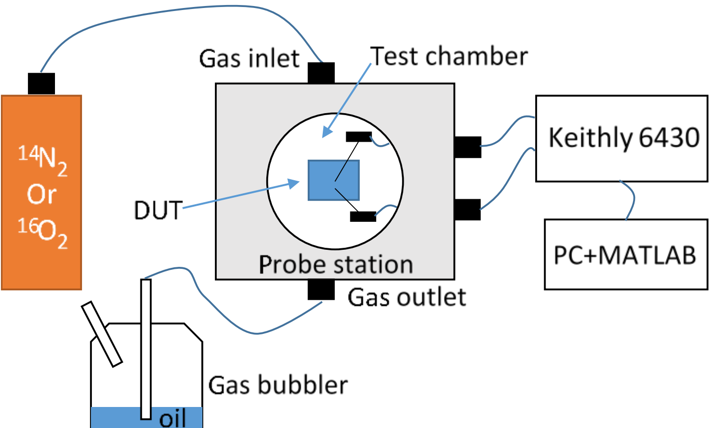
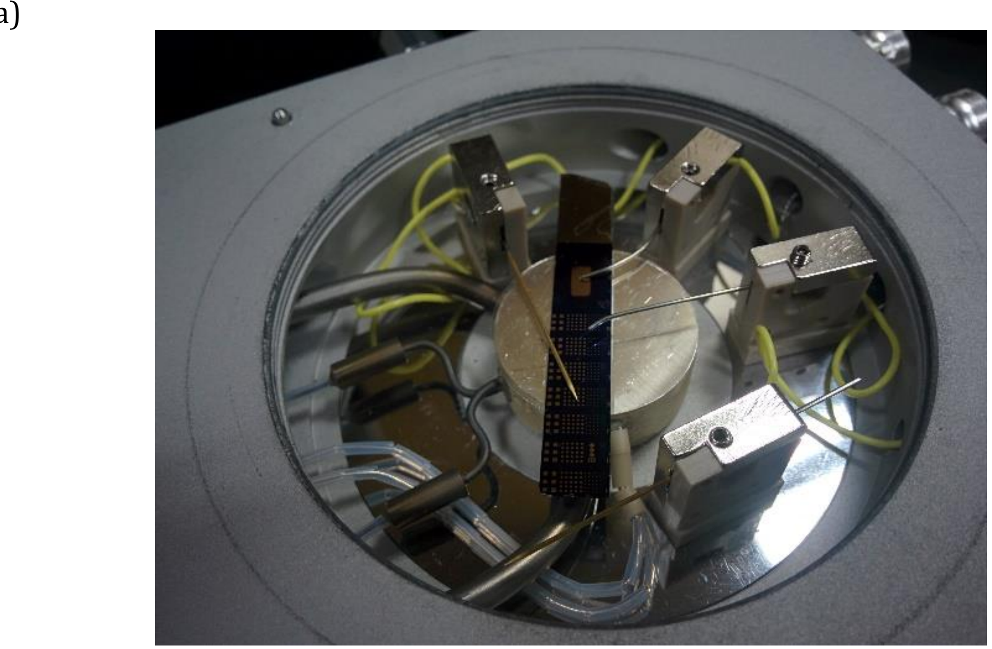


Figure 11: Schematic of the steady gas flow probe chamber setup: a sealed probe station (lid not shown) containing a sample which is contacted using two tungsten needles. These needles are biased through a Keithley 6430 SMU. The input gas comes from either a $^{14}N_2$ or $^{16}O_2$ gas cylinder with regulator. The exhaust gas exits the probe station to be fed into a bubbler, which guarantees a positive gas pressure in the chamber with respect to atmosphere.

a)



b)

Figure 12: Linkam chamber photos: a) inside of Linkam chamber with sample loaded. b) Linkam chamber setup with gas bubbler and gas inlet from high pressure gas cylinder (either $^{16}O_2$ or $^{14}N_2$).

Biasing was provided through the usual Keithley 6430 SMU. This was connected to the probe station using shielded BNC cables. Sweeps were carried out until switching was deemed to have occurred. Devices stuck in high resistance states were assumed to be either disconnected or defective.

Overall, this simple sealed probe station, led to electrical tests in atmosphere whose amount of oxygen had been augmented or reduced compared to air. The injection of $^{16}O_2$ and $^{14}N_2$ is what made this possible. This led to further observations concerning the importance of oxygen when it comes to device operation.

3.2.1.5 Ion Beam Microscopy and Quadruple-based SIMS analysis

SIMS analysis was carried out using the high vacuum setup described in 3.2.1.2 and shown in Figure 9.

SIMS analysis would start with the identification of an area of interest on a loaded sample. The kit would generally be used as an ion microscope and set the Ga^+ ion gun to a relatively low current mode (usually about 40 pA) and scan over the sample surface while translating the sample. This process leads to electrons and ions (secondary ions) being emitted from the sample surface thanks to beam interaction. The channeltron detector present in the vacuum chamber enables either a positive ion or electron (plus some negative ions) image of the surface under analysis to be constructed. The type of image can be chosen by changing the direction of the voltage bias applied at the entrance of detector (usually ±80 V). Electron images generally help understand the conductivity of samples whereas positive ion images provide information on topography.

Once a region of interest has been identified, the Ga^+ beam current can be increased up to 1000 pA (but more often 450 pA was used). A higher beam current allows a greater number of ions to come off the surface but also leads to significant surface milling. A higher count of secondary electrons allows the FEI and Hiden Quadrupoles to be used for SIMS elemental analysis of the surface.

The above detectors work very similarly and can identify either positive or negative ions of a chosen mass-to-charge-ratio. Given that two detectors were available, it was possible to identify both positive and negative ions simultaneously. However, only one ion-mass-to-charge-ratio could be sampled at a time on each detector.

Continuously scanning a region of interest with an ion beam mills the surface enabling depth profiles to be obtained. Different materials have different milling rates and secondary ion yields, leading to difficulty in converting from milling time to depth with absolute certainty.

It was possible to increase the dwell time of the scanning beam at each point within the region of interest. This leads to a slower scan rate per frame with consequent higher mill rate along with increased secondary ion counts per 'pixel'. This would made it possible to build an elemental map of what was being scanned using the SIMS detector, which is not as sensitive to ions as the channeltron. Carrying out subsequent maps helped generate mapped depth profiles, which can be referred to as elemental tomographies. This requires the Ga^+ beam to be in sync with the quadrupole SIMS detector. Unfortunately, such synchronization is only available with the FEI detector and not with the Hiden one.

SIMS analysis leads to the production of elemental depth profiles of sample as well as 2D elemental maps. When repeated, these 2D maps produced 3D elemental maps or tomographies of the samples.

3.2.1.6 Time of Flight (ToF) based SIMS

SIMS analysis was carried out using an IonToF time of flight SIMS (Tof.SIMS5) [123, 126]. This instrument, whose diagram is show in Figure 13, is similar to the

previously described SIMS in the way it excites the surface under analysis. It still relies on the use of an ion beam to generate secondary electrons. However, it has some substantial differences.

This figure is removed from the original version due to copyright restrictions.

Figure 13: Schematic of the iONTOF TOF.SIMS5. Image obtained from Ref. [126] by iONTOF.

Firstly, one now has two beams that can be used for different purposes. A Cs^{+} beam is used for milling and a Bi^{3+} beam for analysis. This means that heavy ions can be used for milling and lighter ones for analysis. This makes it possible to use lower accelerating energies, so analysis is better focused on the sample surface. The instrument is geared towards elemental map profiles.

Another important characteristic of this instrument is that it can analyse a very wide range of mass-to-charge-ratios all at the same time. This means that it is no longer necessary to only select one or a few elements per analysis. This, however, is accomplished by using a unipolar extraction field meaning that one must choose whether to look at positive or negative ions.

All data acquired during a depth profile are saved. This means that mass-to-charge-ratios of interest can be chosen an analysed post data acquisition. It is also possible to select regions of interest from which virtual elemental depth profiles may be generated. This allows for much greater flexibility in data analysis compared to the simpler SIMS.

The time of flight detector works by first accelerating all collected secondary ions to an equal energy. Smaller ions will travel faster than heavier ones and hit the detector first. It is therefore the time that has elapsed since the surface was excited with the analysis beam that determines the mass of the detected ion. This allows for a large wealth of information to be saved for each sampled sample slice.

3.2.2 Software

Certain instruments could not be used without adaptation and modification. The Keithley 4200-SCS is a complex and controlled via a built-in Windows XP computer. Unfortunately, the provided software does not allow for easy programmability of the SMUs and PMUs. Through the GUI, only simple sweeps (allowing “measured” or “programmed” voltage to be recorded) are possible. Programming the execution of pulses (whether via the SMUs or PMUs) was only possible through external control scripts. Scripts do not run in real-time and most probably real-time execution is not a feature found in the 4200-SCS except for PMU waveform execution and duration of each PMU pulse.

Fortunately, Keithley allows measurement commands to be issued remotely to the 420-SCS. Once the remote-control software KCIX has been launched, the SMUs may be controlled using proprietary commands over the available GPIB interface. This allowed the writing of MATLAB scripts to automate test execution. These 4200-

SCS SMU scripts were written in MATLAB in two versions. The first forces 0 V on the DUT between pulses and the second leaves the device terminals floating. Descriptions may be found in section A.2.1.

While the SMU could be directly accessed using commands built into KCIX, this was not possible with the PMU; it was necessary to write C drivers. PMUs are only accessible over proprietary 'LPT' C functions provided by Keithley. These are pre-installed on the Windows XP computer that is inside the Keithley 4200-SCS. It was necessary to write C functions/drivers which would receive GPIB commands/calls via KCIX, call the needed LPT functions, retrieve data, return it to KCIX and therefore to MATLAB over GPIB. Coding test scripts directly in C was certainly an option but the lack of a debugger did not make this appealing. Three C 'drivers' were written whose descriptions are in Section A.2.2.

All the drivers output the generated signal on channel 1 of PMU1 and the inverse of this signal on channel 2 of the same PMU. This is needed to double the programming voltage applied to the device under test (DUT). All devices being tested were left floating with respect to ground.

Several MATLAB scripts were written to control the above drivers. Most of these scripts provide automated data retrieving and saving to comma delimited files. They are described in section A.2.3.

When it comes to the Keithley 6430 Sub-Femtoamp Remote SourceMeter®, the device was only controlled over GPIB via MATLAB. Controlling this device required a different set of commands to those used above. Unlike the 4200-SCS, these are standardised GPIB commands and not proprietary. Simple scripts were

written to be executed while using the instrument for in situ switching. These scripts drove two different kinds of sweeps and are described in section A.2.4.

While running in situ tests, the RGA found in the previously described vacuum chamber was controlled using Hiden Analytical software running on a Windows 7 PC. Data could be exported into comma delimited files. The channeltron charged particle detector was instead connected directly to the Windows NT 4 PC controlling the rest of the microscope and FEI quadruple SIMS. Ion detection with this older detector was correlated with the appearance of white streaks on the screen. This was because the screen was left in synch with the blanked Ga^+ ion beam (the beam was not turned on or emitting). Normally, as the ion beam scans a surface, a pixel is drawn on the screen. The intensity of this pixel corresponds to the ion count obtained when scanning a point on the sample surface. This process is known as secondary ion detection imagining. Here, the detector was used for primary ion detection.

The IonTof ToF.SIMS5 is controlled through proprietary software on a Windows 7 PC. This same software can also be used to carry out data analysis and extraction on the large binary files generated during capture.

MATLAB was also used to process elemental maps generated by the FIB-SIMS as well as the IonToF-SIMS. Custom MATLAB scripts were used to reduce noise though binning (or averaging) 3D tomography data sets. Similar scripts were also used to normalise data to 255. This means that the full dynamic range of the 8bit greyscale output TIFF files is always used. Other scripts were used to create maps representing the ratio of two elements.

OriginPro 2016 and 2017 were generally used to plot all 2D plots.

3.3 EXPERIMENTAL TECHNIQUES

3.3.1 Electrical Measurements in Air

This section covers the electrical input parameters used to carry out general and pulsed characterisation in air. Electrical measurements in atmospheres other than air are described in section 3.3.2.

3.3.1.1 Current Compliance Voltage Sweeps

This first kind of sweeps were carried out to gain a general understanding of device behaviour. Since these sweeps are meant to change the resistance state of a device, they were usually carried out forward and backwards to ensure that the device retains the new state.

It was important to look at behaviour in air as this is the target environment for most memory devices. Pristine TiN/SiO_x/TiN devices were usually swept up to ~10 V to make sure an electroform would occur. Electroforming was always carried out with current compliance enabled. Resetting of TiN/SiO_x/TiN devices seemed to occur with sweeps reaching 3 V without current compliance. Setting of the same devices occurred with sweeps reaching ~ 5 V with current compliance. All sweeps occurred in the same positive polarity as devices were unipolar.

To carry out the above sweeps, the device under test (DUT) was often placed on an Al plate which was then placed on glass insulators placed on the Signatone stage. This grounding technique, appeared to lead to fewer unexpected events such as undesirable device meltdown (probably due to unwanted parasitic capacitance discharging).

3.3.1.2 Current Compliance Long Pulsing

After initial device characterisation of the device family using sweeps, it was chosen to use pulsing with current compliance provided by the Keithley 4200-SCS high output impedance SMUs to test reliability and endurance. Switching voltages determined from sweeps were used to have devices switch quickly between states. SMU pulses, but with small amplitude (<1 V), were used to read resistance states of devices.

The Tektronix MDO4104-6 oscilloscope was put in parallel with the DUT to track how current compliance reduces voltage over time as a device electroforms or sets. An equivalent circuit is shown in Figure 14 (top). Figure 14 (bottom) shows instead how the same scope was used to look at how current changes during resetting. To do this, a 90 Ω resistor was inserted in series with the DUT while resetting devices. The voltage measured across the fixed resistor could therefore be converted to the current going through the DUT by Ohm's law. This led to the tracing of how device resistance increases during reset (Figure 35).

Endurance testing involved the use of the non-floating version of Script #3 from section 3.2.2. After changing the switching parameters (voltages and currents) in the script, it would be run until switching had stopped occurring for at least a few cycles even after having moved the top needle to a different part of the device top electrode.

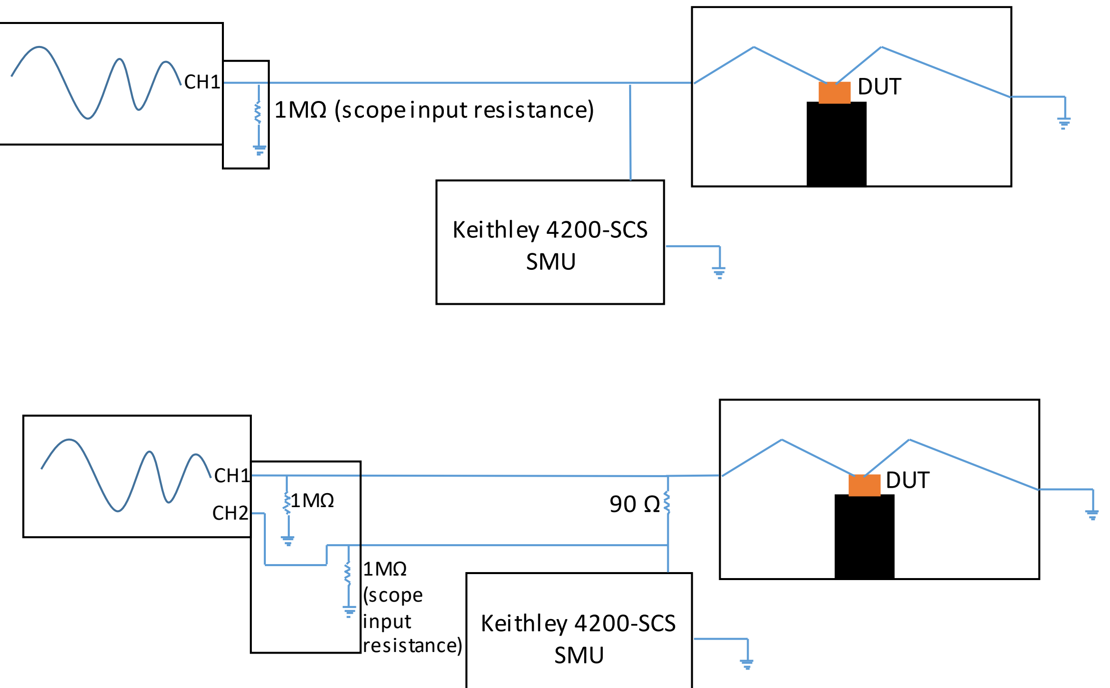


Figure 14: Equivalent circuits showing how the oscilloscope was connected to the SMU output while a device was being tested. The top circuit was used to track voltage across a device as it was being electroformed or set. The bottom circuit was used to track current across a device as the device was reset. Parasitic capacitances are not shown. The DUT changes its impedance dynamically.

Reliability testing was done using the floating version of Script #4 (section A.2.1), which allows switching parameters to be specified. Here the device would be electroformed and manually cycled a few times using the same script before sampling the device state over time. Sampling was always carried out 100 times and every 100 seconds. This was to achieve a reliability plot over a standard 10,000 second period, which corresponds to about 4.3 hours.

When devices did not switch as expected, it was possible to attempt switching them using manually specified single pulses defined in script #2 (section A.2.1). This way it was possible to test different voltages. When even this failed, KITE sweeps could be a last resort before considering a device to be defective.

It is interesting to observe the shape of a pulse generated by one of the SMUs. To do so, MATLAB and the GPIB interface were used to have the instrument

generate its fastest 5 V pulse. The SMU output was directly connected to the fast Tektronix oscilloscope with an input impedance of 1 MΩ. In another case, a parallel resistance of 556 Ω was added to emulate a ReRAM device in its LRS (Figure 15). As expected, shortened time constants can be observed when using the lower load resistance. No significant overshoot is to be observed within the 200 picosecond resolution of the oscilloscope. Pulse length is in the order of tens of milliseconds and appears to change with each pulse. This is most probably due to the MATLAB scripts and the Keithley SMU drivers not running in real time. Due to the use of current compliance and resetting being a self-limiting process, programmed pulsed length always exceeds required programming time rendering such length variations irrelevant. Curves appear relatively smooth and continuous despite the fact that three different voltage ranges must have been automatically used by the SMU: one covering from 0 V to 210 mV, another one covering from 210 mV to 2.1 V and a third one covering from 2.1 V to 21 V [98]. These ranges probably correspond to different physical circuits that are connected to the instrument output one at a time as needed.

It is also interesting to look at the output of the 6430 when using it to generate a programming pulse. Once again, this was tested under the same two loading conditions described above. The system time constant appears to be affected by the load conditions more than in the case of the 4200-SCS SMU. With the 6430, results shown in Figure 16, significant overshoot (about 0.7 V) is observed with the 1 MΩ load condition. This overshoot disappeared when the load resistance is reduced to 565 Ω. An interesting aspect of the 6430 output is that curves appear to be piece-wise functions. As an example, in Figure 16d, between 0 and 0.2 milliseconds, a one observes a slope that is different from the one found at the

beginning of the pulse of past 0.2 milliseconds. According to [118], the instrument's first voltage range goes up to 2 V. This means that two different internal voltage ranges are used to reach 5 V and in fact the curve shape changes at about 2 V as voltage ramps up. These changes in curve shape do not occur at 2 V when the voltage is being reduced. One could assume that some hysteresis is implemented in the range selection algorithm and ranges overlap [118].

Long pulsing with current compliance is a useful addition to sweeps with current compliance. This method quickly provides important device metrics such as endurance and reliability by avoiding the need for slower full sweeps. However, this method still uses SMUs whose shortest pulse length is in the range of tens of microseconds. The next section provides a solution.

3.3.1.3 Fast Nanosecond Pulsing

The PMU card found in the Keithley 4200-SCS was used. Given that this card can only output up to ±10 V, both its channels were used, to double the voltage applied to the DUT. This brought the maximum voltage that could be applied across a device to ±20 V.

Tests were mainly run by defining a pulse-sequence file in Microsoft® Excel. This file was read by a MATLAB script and sent to the Keithley 4200-SCS over USB and GPIB. Transferring large files takes minutes and so does retrieving collected data. This did not allow for quick test execution. Figure 17 shows the output of a sample input pulse-sequence file (PMU directly connected to oscilloscope), which appear to be quite well defined and continuous. Using single pulse C drivers allows for the output of similar (but single) pulses.

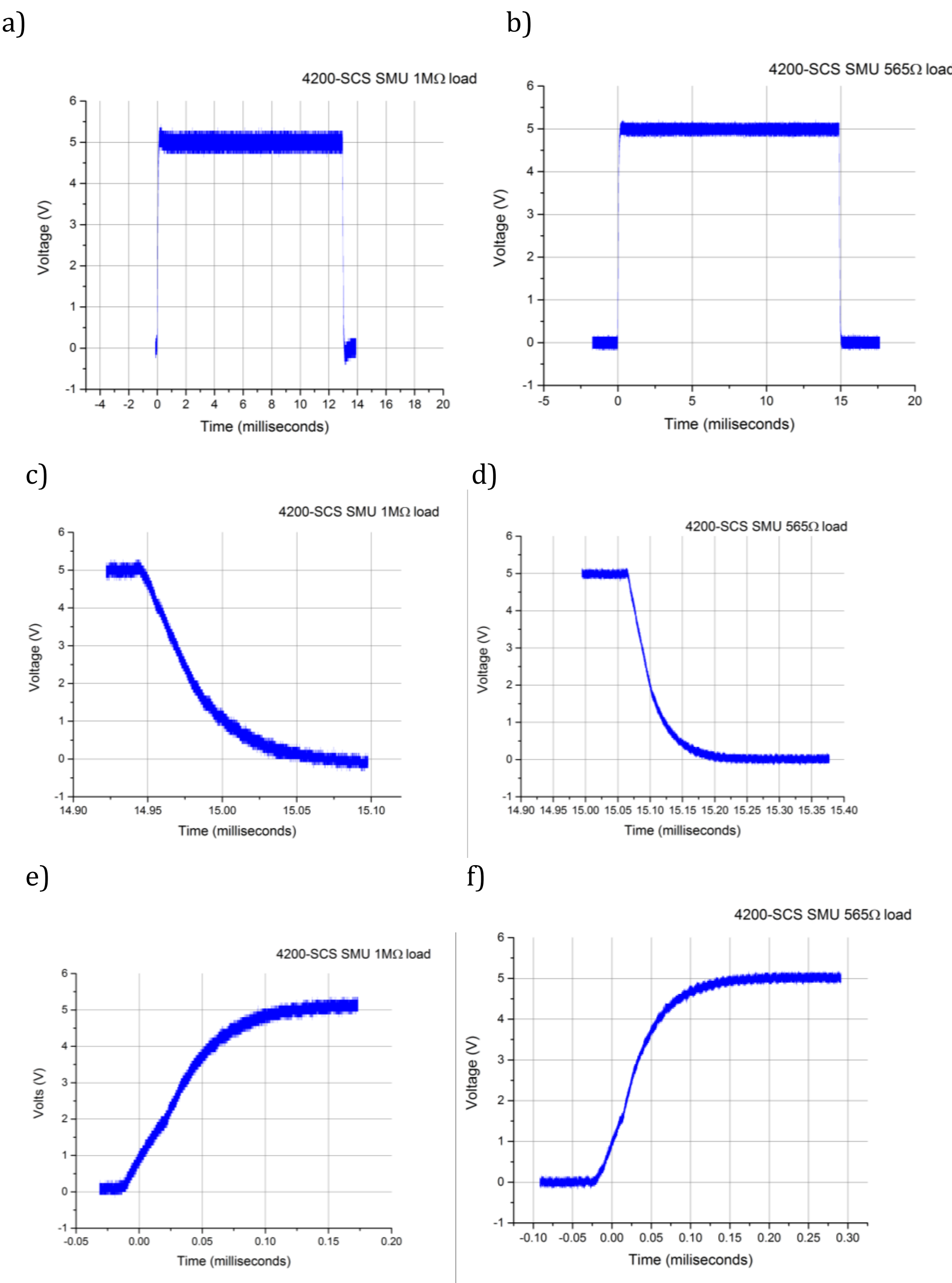


Figure 15: V/T plots of output behaviour of a 4200-SCS SMU. Two different loading conditions were used to emulate different ReRAM states. This helps understand the output timing of the instrument. Pulse length changes depending on computer software, which does not run in real-time. Rising and falling edges were sampled with 200 picosecond resolution whereas 'pulses' were sampled with 1000 picosecond resolution. Each plot comes from a different pulse.

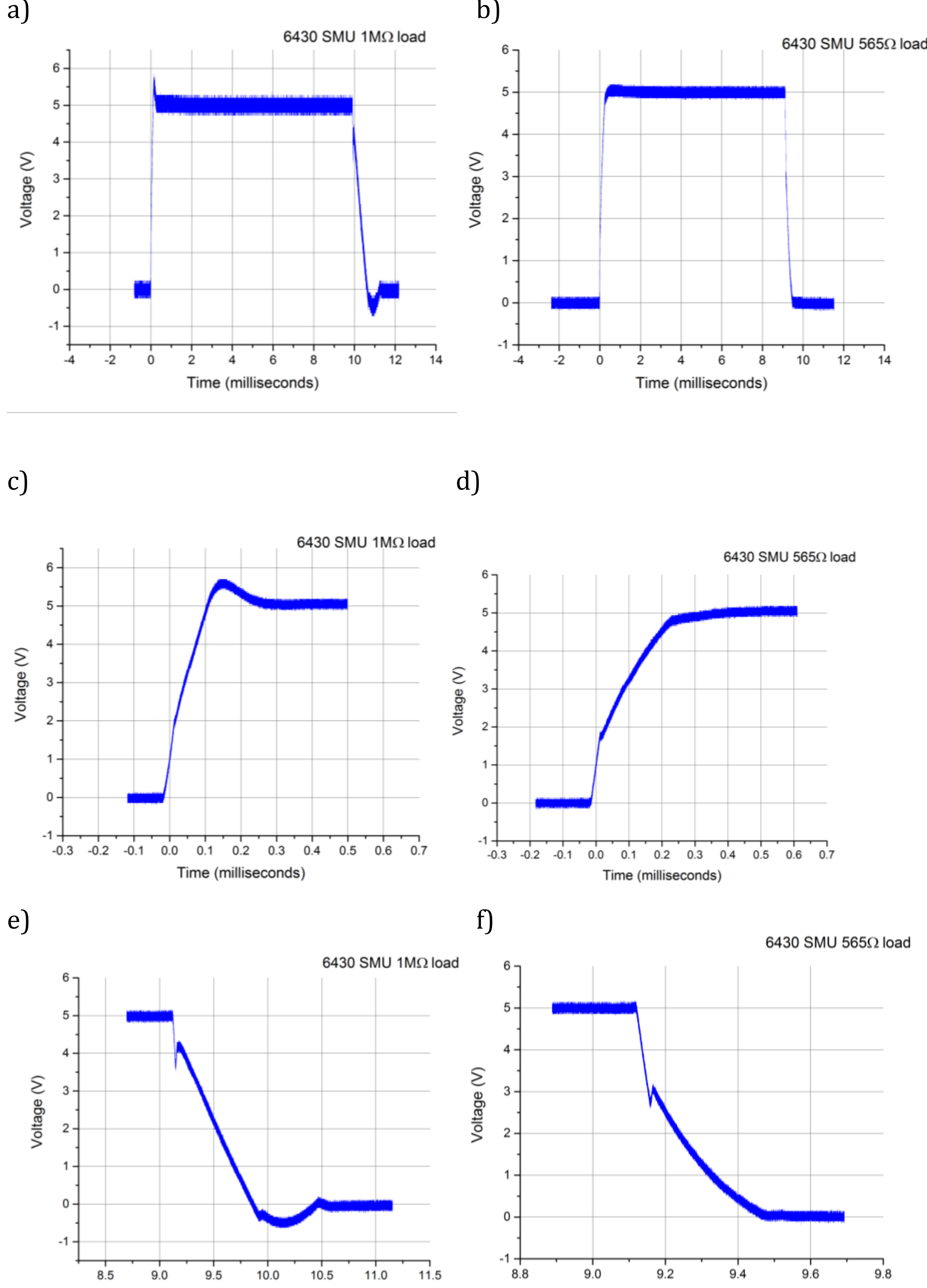


Figure 16: V/T plots of output behaviour of the 6430 SMU. Two different loading conditions were used. Pulse length changes depending on computer software, which does not run in real-time. Rising and falling edges were sampled with 200 picosecond resolution whereas 'pulses' were sampled with 1000 picosecond resolution. Each plot comes from a different pulse.

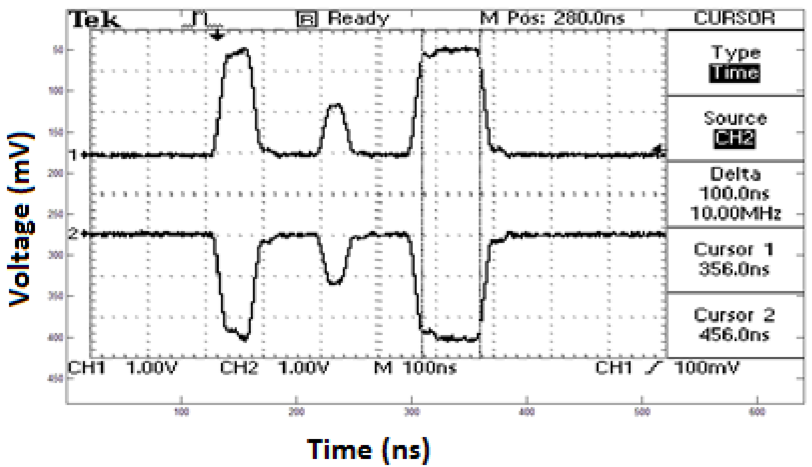


Figure 17: Oscilloscope plot of a sample PMU output from a short pulse-sequence defined in a comma delimited file. 40 ns at 5 V, 20 ns at 2.5 V and 100 ns at 5 V pulses are shown. All pulses have 20 ns rise and fall times. Both channels are shown .

In the case of slower SMU pulsing with current compliance, voltage is dropped as soon as a filament has been created. Current increases due to the filament being electroformed, compliance is reached and voltage is dropped. It will be shown that a long pulse only lasts for a few tens of nanoseconds. When a filament is thermally broken no current compliance is needed as the filament will stop conducting electricity. Using the PMU card, current compliance is no longer available and it must be emulated with knowledge of device behaviour.

The PMU channels have an output impedance of 50 Ω and use 50 Ω transmission lines (coaxial cable). This means that mismatch in impedance is greatest when a device is pristine, large when a device is in its HRS and smallest when a device is in it LRS. Given a constant specified voltage (or pulse amplitude), the device will receive the largest pulse when it is pristine.

All Keithley custom drivers were designed assuming a resistive load >1 MΩ. This means that when a pulse of V_i volts is specified, really V_o volts are delivered to a small load. This way, a maximum of ±10 V will only be delivered on a >1 MΩ load. As the resistance of the load shrinks (e.g. as soon as the filament has electroformed),

the delivered voltage will drop. The delivered voltage will be smallest when the DUI resistance approaches 0 Ω. The relationship between V_i and V_o is defined by Eq.3.1, which keeps in mind that two channels are used at once leading to doubled output impedance:

$$V_O = \frac{V_i}{100+R_{DUT}} R_{DUT} \quad \text{(Eq. 3.1)}$$

In practical terms, this means that identical pulses may be used to set and reset (or even gradually set and reset) a unipolar device such as a TiN/SiO_x/TiN one. In the previous sections, it was shown that reset voltages need to be lower than set voltages. Fortunately, a device that needs to be reset will also be in a significantly lower resistance state than that of a device that needs to be set. This, per Eq. 3.1, means that a smaller voltage than defined will reach any device that needs to be reset. Therefore, a pulse whose magnitude was previously used to set the device, can now be used to safely reset it.

The resistance of a DUT can be accurately measured independently of mismatch. This is because the PMU returns sampled values for both current and voltage. Sampling never occurred during programming pulses as these are expected to dynamically change the device resistance and would lead to unreliable results. As with SMU pulsed tests, resistance was sampled using small voltage reading pulses following programming pulses.

Tracking device behaviour when it was connected to the PMU was possible by keeping an oscilloscope in parallel to the two channels and looking at the sum of the two traces (4.1.2.2). The equivalent circuit used for these measurements is shown in Figure 18.

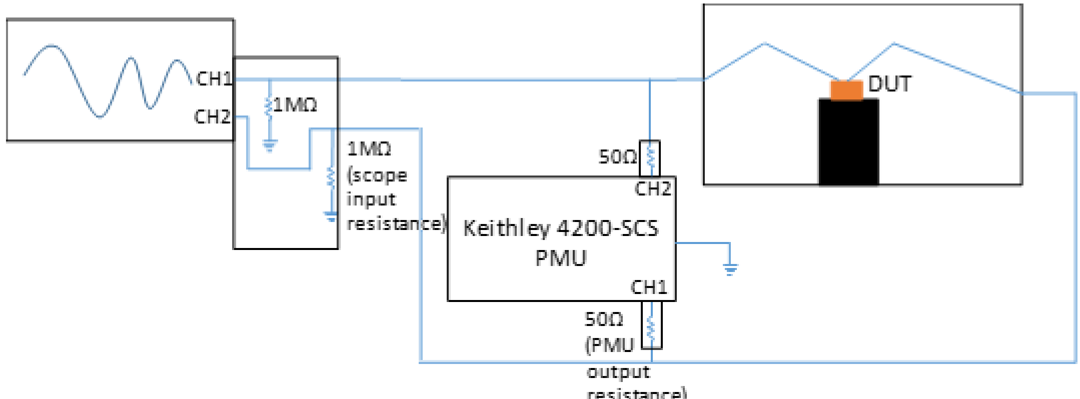


Figure 18: Equivalent circuit used to keep track of device behaviour while using PMUs. Parasitic capacitances are not shown. The DUT changes its impedance dynamically.

This programming technique will translate into programmers that will not need large current compliance circuits or the ability to produce different voltages. These circuits will only need to have programmable pulse width. This last requirement can be easily implemented when a fast-enough system clock is available.

3.3.1.4 Neuromorphic Behaviour

Once again, short PMU-generated pulses were used to stimulate devices. However, now more sampling, or reading, pulses are used following each programming pulse to fully observe the transient behaviour of devices. It is interesting to create volatile states using pulses that are shorter than those leading to non-volatile states and therefore see device behaviour over time.

This form of programming is not possible using SMUs and current compliance, as the objective is now to stop filament electroformation before it is fully created and stabilised. SMUs also have sampling rates that are too slow to observe fast state relaxation.

3.3.2 In situ Switching

In situ switching experiments were carried out in three different environments: the FIB high vacuum chamber with biasing needles and SIMS/RGA detectors, the $^{18}O_2$ exchange rig and custom made electrical contacts to the DUT, and a Linkam sealed probe station that allowed gases to be flowed through.

In all tests, electrical biasing was carried out using the simpler Keithley 6430 Sub-Femtoamp Remote SourceMeter® controlled through a MATLAB script and a USB to GPIB interface. It was connected to two Kleindiek micromanipulators in the vacuum chamber using BNCs, to a vacuum feedthrough on the exchange rig using unshielded wires, and to the Linkam probe station using BNC cables. In the case of the vacuum chamber, the sample was grounded either using one of the two manipulators or the stage itself. A non-regular kind of sweep was used for these tests. These pseudo-sweeps were made up of pulses, each of increasing magnitude, interleaved with low voltage reading pulses. The source output was left floating between all pulses. Devices loaded in the exchange rig were attached to insulating microscope slides using silver paint leading to a floating silicon wafer bulk. Devices had a similarly floating bulk in the Linkam probe station thanks to a layer of insulating glass between the device bulk and the stage

3.3.2.1 Oxygen and Ion Detection

Detection of emitted species must be carried out in vacuum. Devices were stuck to an Al stage using silver paint or a metallic clip and placed in the custom Phillips microscope vacuum chamber. The pressure of the chamber was ~$2x10^{-6}$ mbar. Imaging of the tungsten needles was done using secondary electrons, or ions, while the microscope ion beam scanned the surface with a beam current of 30 pA.

Beam power and imaging time had to be minimised to limit unwanted ion beam milling.

The Hiden RGA attached to the vacuum chamber was used to detect emitted oxygen. Parameters, or internal acceleration voltages, were optimized for O_2 detection for all experiments. This was done using auto-tuning scripts. For some experiments, a wide spectrum of mass to charge ratios was scanned to observe which gases are released into the chamber during device biasing. For other experiments, only a few ratios were scanned at a higher refresh rate to obtain a high time resolution transient curve of those gases that were previously detected.

The RGA was first left running overnight to establish a background (Figure 19). As the RGA filament heats up, the number of counts increases, then decreases as the pressure reduces. Several spikes are detected at regular intervals. According to Hiden, the manufacturer, this can be due to unwanted charging/discharging phenomena occurring within the RGA. These spikes only correspond to single data points. Multiple data points spikes were never observed while sampling overnight without biasing. Therefore, during device biasing, any spike involving several data points will be considered as emission from the device.

Firstly, the Hiden Residual Gas Analyser (RGA) was left running while a TiN/SiO_x/TiN wafer was present in the vacuum chamber. The RGA was left overnight sampling once every second several mass-to-charge-ratios of interest corresponding to $C^{16}O$ (28), $C^{18}O$ (30), $^{16}O_2$ (32). $^{18}O^{16}O$ (34), $^{18}O_2$ (36). This was done to set a baseline of species present in the FIB chamber and to see whether the loaded sample would outgas (results in section 5.2).

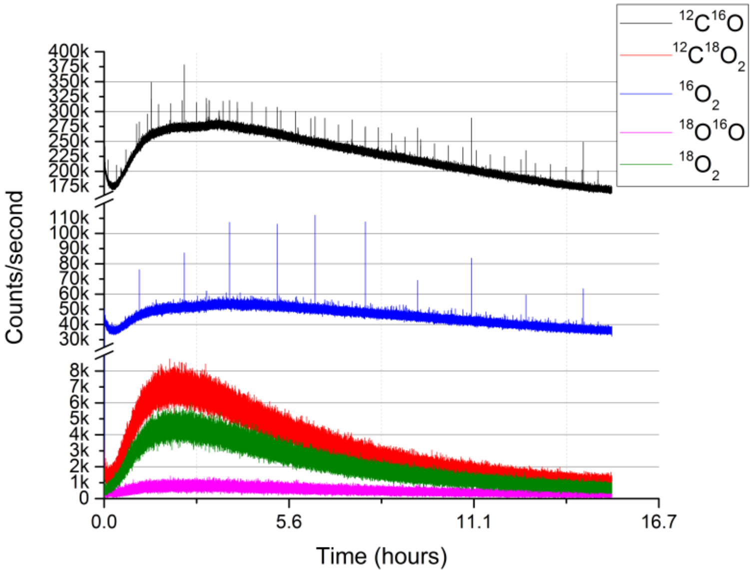


Figure 19: Overnight RGA profile looking for oxygen species and organic contaminants. The initial increase in counts is due to heating of the ionizing filament. The subsequent decrease is due to reducing pressure. All visible spikes are single data points.

Survey scans were run from mass-to-charge-ratios of 0.4 to 40 with an increment of 0.1 and a dwell time of 10 milliseconds. Scans studying the transient behaviour of emitted oxygen only attempted detection of mass-to-charge-ratios 28 and 32. These represent ionized O_2 and CO molecules. The latter are expected to form due to a non-negligible background count of carbon species. The RGA was left running while voltage sweeps were carried out on TiN/SiO_x/TiN devices. Resistance was measured using pulses of 0.8V following each sweep pulse. Device voltages were increased until a significant change in resistance occurred ($\log(|\Delta R|) > 1$).

It must be noted that the RGA ionising filament produces positive ions out of non-charged molecules. It is these positively charged molecules that are detected. Therefore, negatively charged molecules should not be accidentally detected as part of the neutral species; this would require double ionisation.

The detection of oxygen species emitted in ionic form was carried out using the channeltron detector. This detector has an outside front-end biased at +80 V,

which is used to attract negative ions. Ions attracted by this grid are accelerated into the detector thanks to an internal grid biased at +180 V. This detector does not provide mass information but only a white dot on a screen when a negative particle is detected.

Further ion emission analysis was carried out using the SIMS quadrupole. Here, the channeltron was biased with a negative voltage of about -80 V to repel and push negative ions towards the entrance of the SIMS detector. While the scanning beam was blanked, the detector was used to sample mass-to-charge ratios 16 and 32 corresponding to $^{16}O^{-}$ and $^{16}O_2^{-}$. A variety of sampling times were used while only focussing on mass-to-charge-ratio 32.

3.3.2.2 Cycling devices in an $^{18}O_2$ rich atmosphere

The rig described in 3.2.1.3 was used to cycle devices in an $^{18}O_2$ atmosphere. Starting with the fully assembled hardware, the rig was first pumped to a high vacuum. The main valve was then shut down, isolating the exchange tube (where the sample was loaded) from the pump and RGA. $^{18}O_2$ mixture was then injected into this section. Electrical cycling could then be carried out, followed by recovery of the $^{18}O_2$ mixture and flooding the vacuum by shutting down pumps and injecting N_2 gas (i.e. 'venting').

Fragments of sample wafers were attached to microscope glass slides using silver paste. These were therefore glued to a high vacuum electrical feedthrough using high vacuum glue. Copper wires were then extracted from a high flexibility cable to contact the TiN top electrodes and the feedthrough contacts. These wires were secured using silver paste on the glass slide and their spring constant allowed contact. If wires moved off contact, this could easily be observed electrically.

A modified oxygen atmosphere (containing about 33% $^{18}O_2$, 33% $^{16}O_2$ and 33% $^{(16+18)}O_2$; more details follow) could be injected from molecular sieve canisters up to a pressure below the atmospheric one. 400 mbars was always chosen as a higher pressurisation could have rendered vacuum gaskets leaky due to too small a difference in pressure. A small amount of this mixture was leaked into the turbo-molecular pump stage so that the installed Hiden RGA could be used to characterise it.

The usual Keithley 6430 was then used to sweep devices using a 0.05 V increment sweep whose voltages were interleaved with read pulses. This meant that every increment in voltage was interleaved with a read pulse whose voltage would not change. Sweeps were stopped once devices showed a significant change in resistance. Reading was mostly carried out at 0.1 V or 0.5 V; the exact read voltage is specified when presenting results. Sweeps would start from the read voltage. Devices that showed good switching through sweeps would be cycled using setting and resetting pulses with voltages learned from sweeps. 1 mA or 3 mA current compliance was used to electroform and set. These current compliance values led to a good compromise between switching ratio and endurance. High currents led to lower resistance LRSs but shorter endurance. The SMU was controlled as previously described and sweeps were therefore used. Switching appeared to have occurred by observing a change in resistance of about one order of magnitude. Devices were stuck in a very high resistance (>10 MΩ) were deemed to be either disconnected or defective. This was the case with about 80% of devices.

Before carrying out exchange experiments, it was important to establish the contents of the $^{18}O_2$-rich atmosphere. Figure 20 shows an RGA survey of the

exchange rig high vacuum chamber before and during a leak from the exchange chamber, which contained ~400 mbar of the $^{18}O_2$-rich atmosphere. At mass-to-charge-ratio 32, corresponding to $^{16}O_2$ gas, a partial pressure of 5.37×10^{-7} $torr$ was observed. At mass-to-chrage-ratio 36, corresponding to $^{18}O_2$, a partial pressure of 6.45×10^{-7} $torr$ was observed. At mass-to-charge-ratio 34, corresponding to $^{18}O^{16}O$, a partial pressure of $1.075\cdot10^{-6}$ $torr$ was observed. From these values it could be determined that the total oxygen pressure was: $(5.37+6.45+10.75)\times10^{-7} torr = 22.57\times10^{-7}$ $torr$ and the $^{18}O_2$ partial pressure plus the $^{16}O^{18}O$ partial pressure was: $\left(6.45+\frac{10.75}{2}\right)\times10^{-7}$ $torr = 11.83\times10^{-7}$ $torr$. Therefore, the $^{18}O_2$ fraction was: $\frac{11.83\cdot10^{-7} torr}{22.57\cdot10^{-7} torr}\cdot 100 = 52.41\%$.

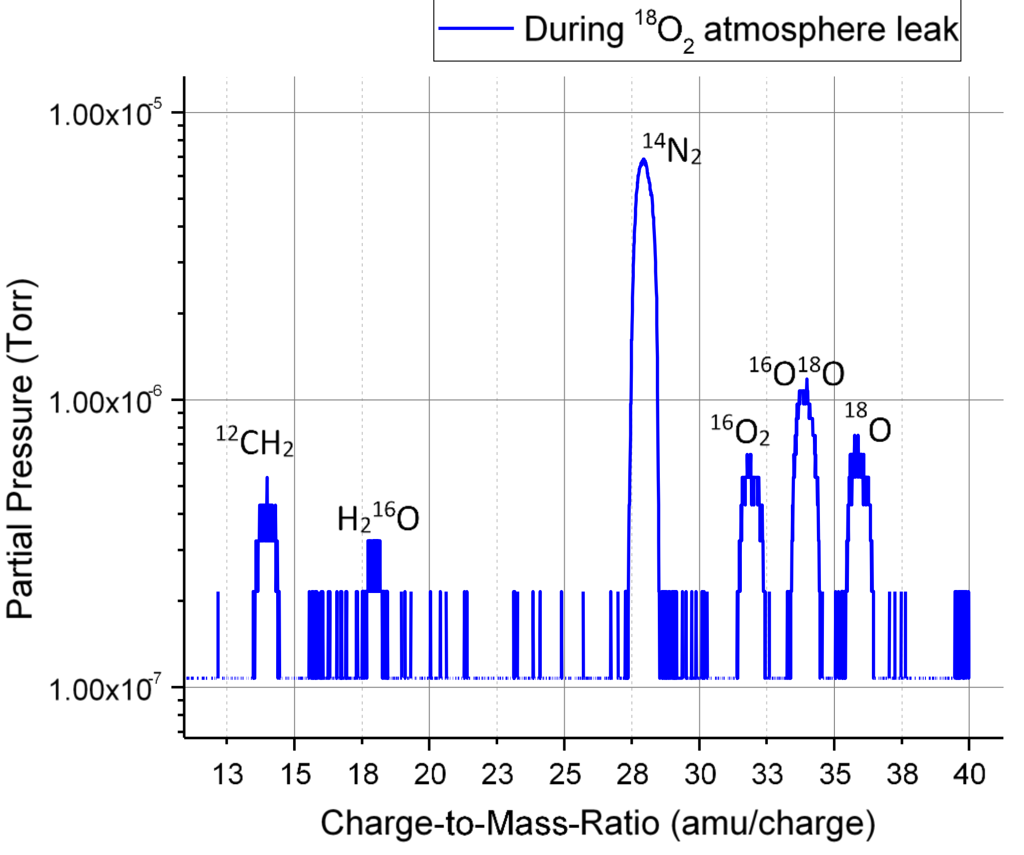


Figure 20: RGA survey scan of the exchange rig high-vacuum chamber during $^{18}O_2$-atmosphere leak. Labels placed on peaks indicate which molecule each peak may be representing.

Due to instrument error, the peak for mass-to-charge-ratio 34 was found at exactly 33.99, mass-to-charge-ratio 32 plateaued between 31.95 and 32.01, and mass-to-charge-ratio 36 had its peak at exactly 36. Other peaks could be attributed to contamination present in the pump section of the rig partially due to organic compounds introduced by the pump itself. Mass-to-charge-ratio 14 was probably due to $^{12}CH_2$. Mass-to-charge-ratio 18 most probably corresponded to water and mass-to-charge-ratio 28 to N_2.

3.3.2.3 Cycling devices in $^{16}O_2$ and N_2 rich atmosphere

The Linkam sealed probe station was connected to either a $^{16}O_2$ or N_2 tank depending on the type of atmosphere devices were to be cycled in. The station output was always connected to a bubbler containing about 1 cm of fluid to ensure positive injected gas pressure in the chamber. The tank regulators were adjusted to have vigorous bubbles in the bubbler.

When adding $^{16}O_2$ to the chamber's environment, the flow was left running for a few minutes (between 5 and 15 min) before starting electrical measurements. As $^{16}O_2$ was added to an atmosphere that already contained some, it was clear that the partial pressure of $^{16}O_2$ would quickly increase. When injecting N_2 gas, the objective was to displace as much O_2 as possible. In this second case the gas flow through the chamber was left running for a longer time (between 30 min and 1 hr) before starting electrical measurements.

Electrical measurements were carried out by connecting the sealed probe station to the Keithly 6430 SMU. Electroforming, resetting and setting of devices was carried out using a 0.05 V step voltage sweep in which sweep pulses are interleaved with low voltage read pulses. 1 mA and 3 mA were used for current compliance

during electroforming and setting sweeps. 100 mA current compliance was used to reset. Sweeps were generally manually stopped following a significant change in measured resistance and started at read voltage. Read pulses had an amplitude of 0.1 V.

3.3.3 Surface Analysis and Tomography

3.3.3.1 Quadrupole SIMS Measurements

Samples cycled in $^{18}O_2$ rich atmosphere were placed in the FIB SIMS chamber. Using the ion beam-microscope feature of the instruments, it was possible to identify areas of interest, which showed typical surface deformations (more on this in section 5). These images were generally obtained using a 30 keV, 42 pA beam and a magnification of about 10 kX using 6 second scans. Faster scanning was used to locate regions of interest with minimal beam damage. Both electron and positive ion imaging were used by changing the direction of the bias applied at the entrance of the channeltron.

Elemental analysis was carried out using a higher beam current of about 350 pA so that the quadrupole would have sufficient secondary ion counts. It was important to understand the exact mass number for each mass-to-charge-ratio of interest. This was done by running a mass survey on a non-interesting region of the sample. Due to instrument error, mass-to-charge-ratios are often off by a small percentage.

Following this step, it was possible to carry out different types of analysis. ^{18}O and ^{16}O depth profiles could be carried out by leaving the beam scanning with the now higher current on regions of interest while $^{18}O^-$ and $^{16}O^-$ ions were being

sampled. These profiles were carried out on deformed regions, and pristine regions of the same TiN contact. This allowed to observe whether changes in ^{18}O occurred as functions of depth. It can be assumed that the yield ratio of the two isotopes is the same in a given material and material properties may well change between an unbiased region and a cycled/biased region. Therefore, it is important to look at isotopic fraction rather than absolute counts of any of the two isotopes [111]. Alternatingly sampling for both isotopes allows for this to be done.

Similar beam settings could also be used to capture ^{18}O and ^{16}O elemental maps. In this case, given the thin nature of the active layer and high beam current needed, it was chosen not to alternate between isotope mass-to-charge-ratios but rather to acquire a ^{16}O map every two ^{18}O ones.

Depth profiles could be directly exported and processed in OriginLab for plotting. Elemental maps had to be manually saved as TIFF files following each scan. These files were therefore processed in MATLAB; they were normalized and therefore saved as $^{18}O_2$ percentage maps. Given that no binning was used, the images kept their original 477 by 477-pixel resolution.

3.3.3.2 ToF-SIMS Measurments

The iONToF time of flight SIMS requires calibration of mass-to-charge-ratio numbers. Following this step, which is carried out on a specific target, data is captured by alternating the milling and analysis beams. It is therefore important to use appropriate beam currents and accelerating voltages for each beam. Low resolution ion imaging is available to identify regions of interest.

The Cs^+ beam is used for milling and an energy of 2 keV was used along with a beam current of 80 nA. The Bi^+ beam, which is used for data analysis was set up with

an energy of 20 keV, a current of 0.3 pA and a settle time of 100 microseconds per pixel. The scanning area of the first beam was of 300 by 300 micrometres and the one of the second one of 50 by 50 micrometres. 256x256 pixel maps with one shot per pixel were chosen to reduce sampling time. A one second delay was used between the end of milling and the start of analysis.

The pressure of the main chamber could easily reach $6.4\times10^{-9}\,mbar$. An electron flood gun was kept on during sampling to avoid charging. This had a filament current of 2.45 A and an extraction bias of -40 V.

Collected data was processed using IonToF's proprietary software which allows for the extraction of depth profiles by charge-to-mass-ratio as well as ASCII files containing elemental map bitmaps. Depth profile data was processed and plotted in OriginLab. Elemental map topographies were processed in MATLAB and, following 2x2x2 binning (to reduce noise), exported into lossless TIFF files. These exported maps represent $^{18}O_2$ percentage maps and are normalized to 255, which corresponds to the dynamic range of 8 bit black and white TIFF files.

Bi has a higher atomic mass than that of Ga. Therefore, despite similar energies between the FIB and the ToFSIMS, in the case on Bi analysis ions penetrate less into surfaces. the ToFSIMS looks more strictly just at the surface being analysed than the FIB.

4 RESULTS FROM ELECTRICAL MEASUREMENTS

This chapter covers electrical behaviour results of TiN/SiOx/TiN devices in different atmospheres: air, high vacuum, 400 mbar $^{18}O_2$-rich 100% O_2 atmosphere, $^{16}O_2$-rich atmosphere and $^{14}N_2$-rich atmosphere.

4.1 ELECTRICAL BEHAVIOUR IN AIR

4.1.1 Current Compliance Voltage Sweeps in Air

4.1.1.1 Switching Voltages

Figure 21 shows that typical current compliance voltage sweeps in air on TiN/SiO_x/TiN wafers. One has an electroforming voltage of about 8.3 V with 3 mA current compliance, a reset voltage of about 2.7 V (current reaching 20 mA) and a set voltage of about 5.5 V, again with a 3 mA current compliance. As explained, a 3 mA current compliance appears to be optimal for UCL experimental TiN/SiO_x/TiN devices as it leads to the best compromise between endurance and separation between LRS and HRS.

The set curve shows some typical competing processes before set is achieved [55]. These could be thermal reset processes, which are eventually fully overcome by electric field leading to device set. Curves shown in Figure 21 only focus on the first three resistance switching events; these data may be affected by some variations discussed in section 4.1.2.1 which are due to contact surface oxidation.

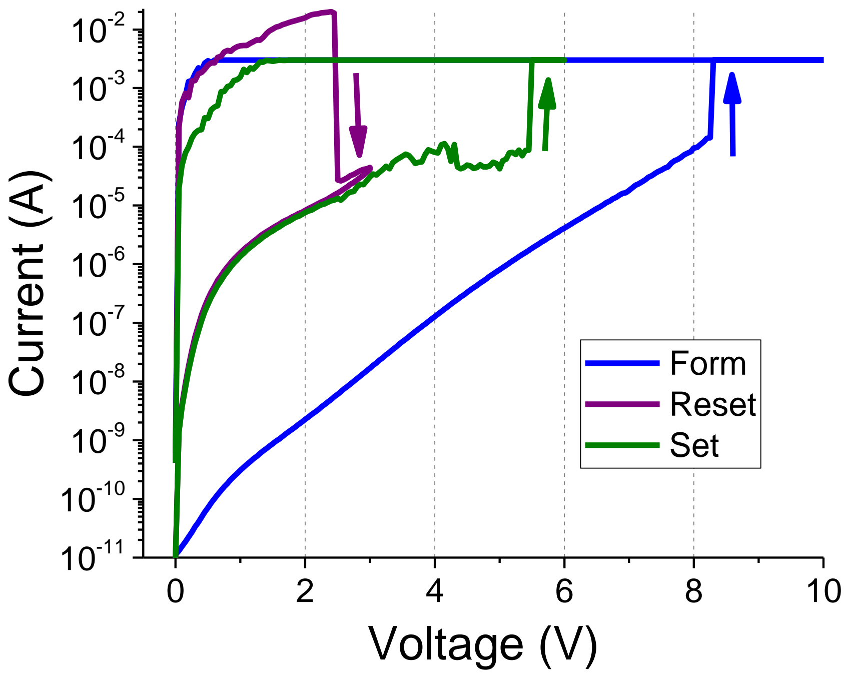


Figure 21: Typical electroform, reset and set IV curves of a TiN/SiO$_x$/TiN wafer device. This device electroformed at ~8.27 V with 3 mA current compliance, reset at 2.5 V without current compliance (current reaches 20mA) and reset at ~5.5 V with 3 mA compliance.

Figure 21 data were collected using the Keithley 4200-SCS to record the programmed voltage and not the real device voltage. Current compliance works by reducing voltage to ensure that the current limit is not exceeded. The actual voltage can be collected by selecting the "measured" option. These curves are shown in Figure 22 using the same input parameters discussed above, and plotted as a connected scatter. Here, upon electroforming, voltage is dramatically reduced to ensure 3 mA is not exceeded. During reset, current compliance does not kick in and no jumps in voltage are observed. During set, one observes a sudden voltage reduction as soon as setting has occurred. Lastly, ~5 mA is reached during reset; this is much lower than in the case of Figure 21.

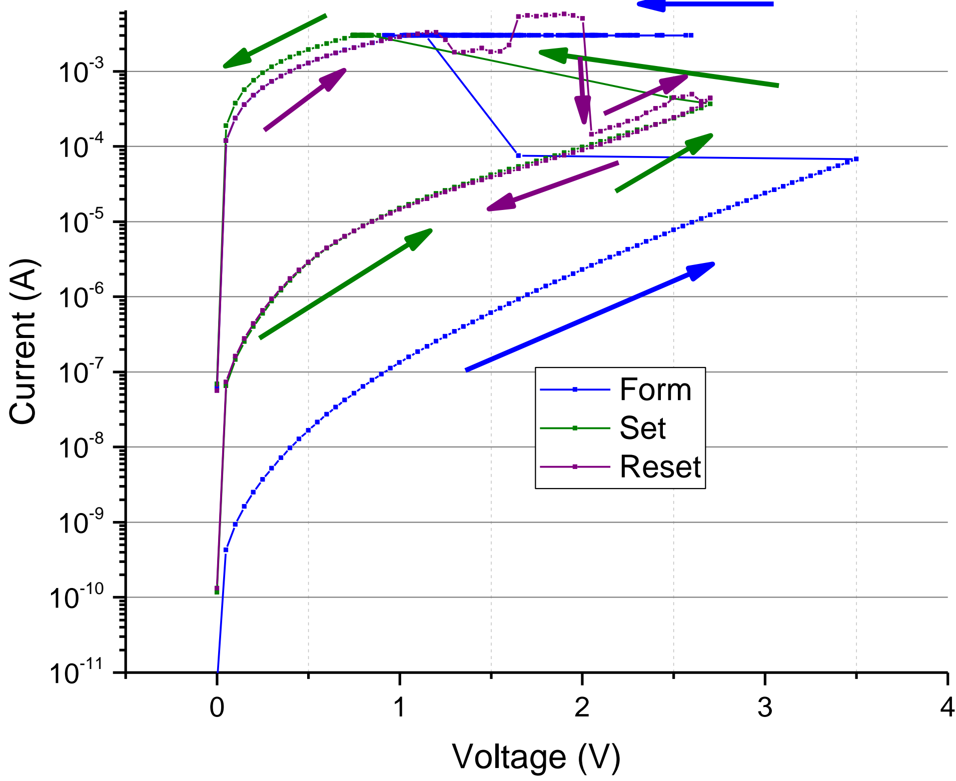


Figure 22: IV sweeps showing electroform, reset and set of a device. Both voltage and current were measured at each point and this helps visualize how current compliance (3mA, in this case) works by limiting current through the adjustment of applied voltage. The low electroforming voltage (~3.5 V) of this device is unusual. Resetting reaches about 5 mA.

When it comes to electroforming voltages, data (partly using Mark Backwell's data) were plotted without taking electrode size into account. Results are shown in Figure 23. Here 3 mA data were collected using Keithley 4200-SCS regular sweeps and 1 mA data using pseudo-sweeps with the Keithley 6430. Out of 12 random electroforming sweeps, none had ended with devices stuck in an unrecoverable LRS, which is signalled by very low resistance (e.g. < 200 Ω). This shows that electroforming in air is a reliable process. A lower current compliance leads to lower electroforming voltages because this lower current is reached earlier during the forward electroforming sweep leading to a probably 'weaker' filament [25]. The 'weaker' a filament the easier it should be to reset it (i.e. lower reset voltage). 'Weak' filaments may also be unstable over time and useful in neuromorphic applications (section 3.3.1.4).

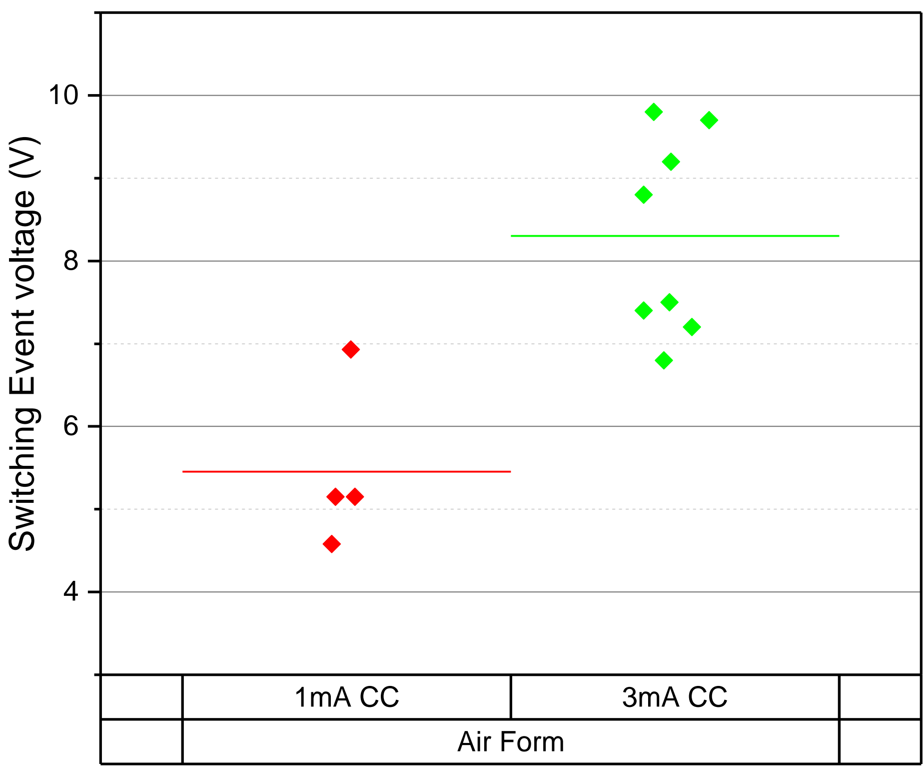


Figure 23: Collected in-air electroforming voltage data. Each marker represents a voltage at which devices successfully electroformed. Horizontal lines represent average values for each data group.

While this study did not involve the collection of large amounts of data to substantiate the average switching voltages mentioned above for set and reset, a former UCL student (Leon Garnett) did. These results are based upon about 3000 sweeps, and are shown in Figure 24. These results show that electrode size does not seem to influence voltages in a significant way and were run with 3 mA current compliance for electroform and set. ~2.7 V ensures reset about 80% of the time. ~4.2 V ensures set about 80% of the time.

Overall, the switching voltages found are comparable to those found in the literature for filamentary ReRAM devices [41, 42, 55, 81]. A few devices found in literature seem to switch at voltages that are about an order of magnitude smaller

[60, 97]. Nonetheless the switching voltages of UCL's devices remain competitive as they allow for modern CMOS programmers to be implemented with ease [22].

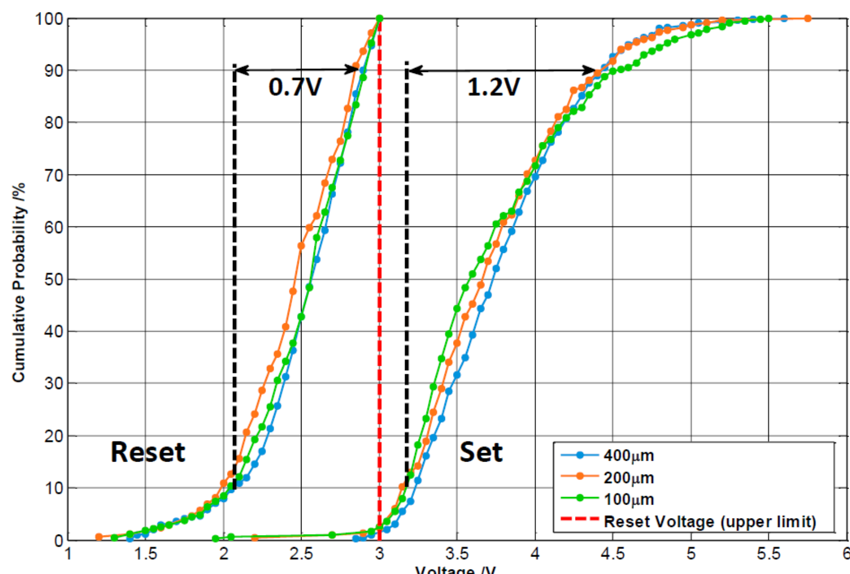


Figure 24: Probability spread of reset and set voltages of TiN/SiOx/TiN device. 2.7 V ensures reset about 80% of the time. 4.2 V ensures set about 80% of the time. Courtesy of Leon Garnett.

4.1.1.2 Conduction Methods

It is interesting to use the data in Figure 21 to obtain information on what conduction methods occur in different states of the switching cycles. While it might seem appealing to use figure Figure 22 as a source of data, Figure 21 appears to provide more stable/reliable data with fewer jumps.

These methods were described in 2.2 and summarized in Table 1. It is interesting to carry out analysis to find out whether trap assisted tunnelling (TAT), Ohmic conduction, Pool-Frenkel hopping, Thermionic Emission, Fowler-Nordheim tunnelling, and quantum point conductance occur. These conduction methods have been previously shown to occur in oxide based devices [2, 36, 55, 59].

Figure 25 shows no evidence of Fowler-Nordheim tunnelling occurring as the $\log(I/V^2)$ v. $1/V$ plot in Figure 25 has no obvious linear regions. This reconciles with statements found in [2] which states that Fowler-Nordheim would only occur at

breakdown voltages in devices similar to ours. Data points involving current compliance were not plotted as they do not represent actual device readings. Similarly, low voltage ranges were not plotted as Fowler-Nordheim involves tunnelling [85] and therefore requires a high electric field to occur [86]. Only the forward electroforming sweep was plotted as this is the only section where the existence of Fowler-Nordheim conduction would not be overwhelmed by other conduction mechanisms.

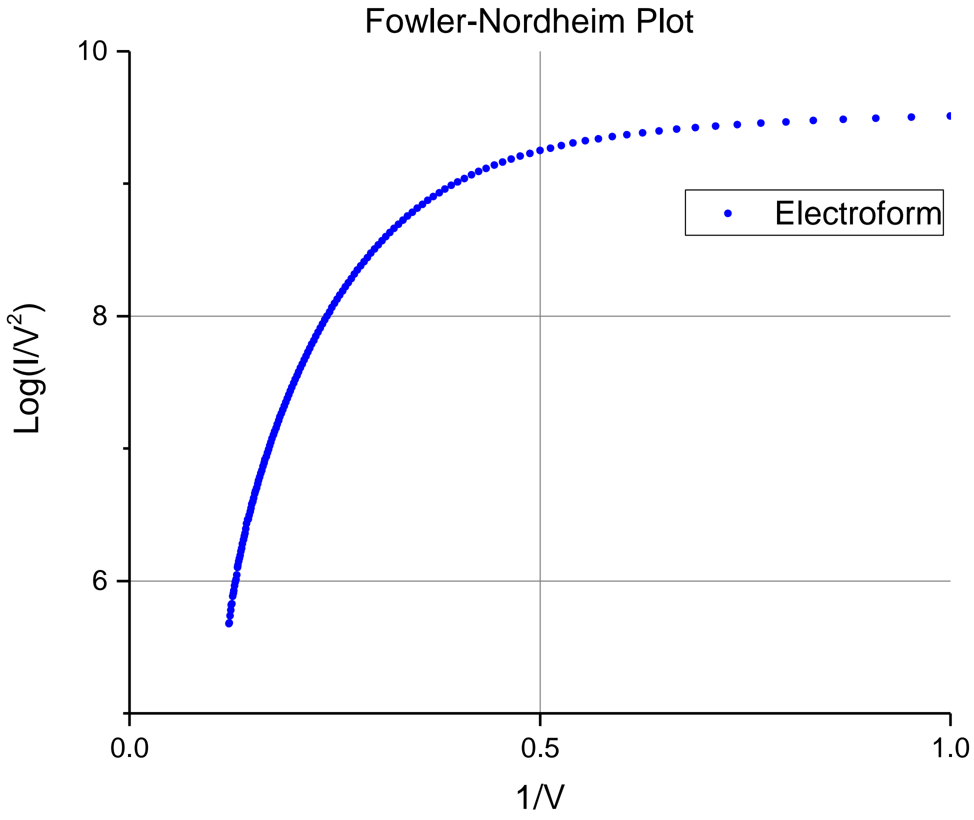


Figure 25: Fowler-Nordheim plot on log(I/V²) v. 1/V axes of the high voltage section of the forward electroforming sweep. No linear region indicating Fowler-Nordheim conduction is found. Original data found in Figure 21.

In Figure 26, one attempts to find linear regions on a $\log(I/V)$ v. $V^{1/2}$ plot, which would indicate probable Poole-Frenkel Emission conduction. Here one finds a linear region with a gradient of 2.95 made up of 66 point. This region corresponds to the electroforming process (forward sweep) between ~3.3 V and ~7 V with a current in the tens of microamperes (therefore far from compliance) and appears

reliable due to a high number of data points. The extracted slope can be used to extract physical parameters of devices using the Poole-Frenkel equation found in 2.2. Using a thickness of 37 nanometres, one calculates an absolute permittivity of 2.47×10^{-10}, which corresponds to a relative one of 27.91 . This result is comparable to the one of 23 found in Ref. [2] on a similar silicon oxide. The expected value for silica should be 4 and for silicon 12. As stated in Ref. [2], this might confirm that Poole-Frenkel conduction occurs within devices but with difficulty or along with some other conduction mechanism.

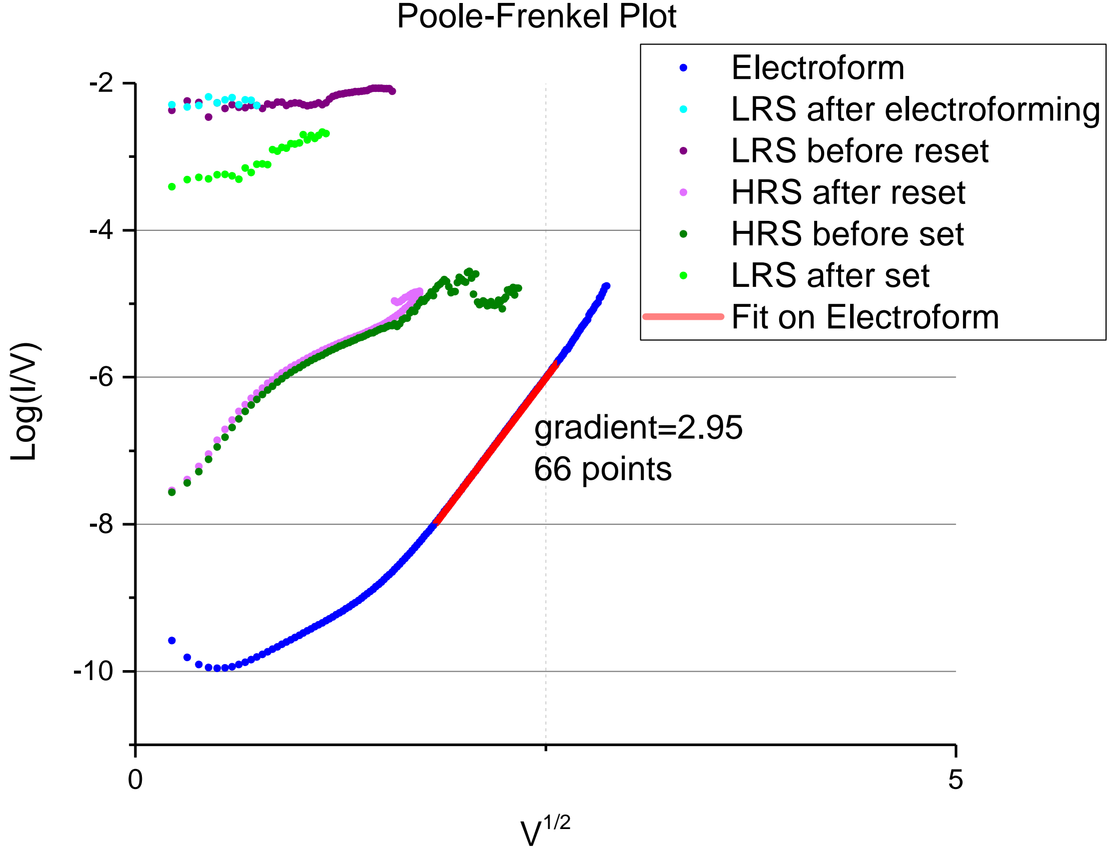


Figure 26: Poole-Frenkel plot. Log(I/V) v. $V^{1/2}$ plot of data found in Figure 21. Straight lines should appear when Pool-Frenkel hopping conduction occurs. One region is found during electroforming between ~3.3 V and ~7 V when current was in the tens of microamperes.

Figure 27 shows a plot using $\log(V)$ v. $\log(I)$ axes. This type of plot enhances linearity in linear plots. Linear regions with a gradient of 1 represent linear regions in the original data. Two regions have gradients extracted from 39 data points

whose values are 2.61 and 2.65. These linear regions occur between 0.37 V and 1.35 V with the device in HRS following reset and similarly with the device in the same HRS before setting; it in not surprising that the two gradients are very similar. Given that the gradients are of almost 3, these regions show that an IV relationship between quadratic and cubic took place in the device. This could be the result of multiple condition methods occurring concurrently. In both cases current is well below compliance.

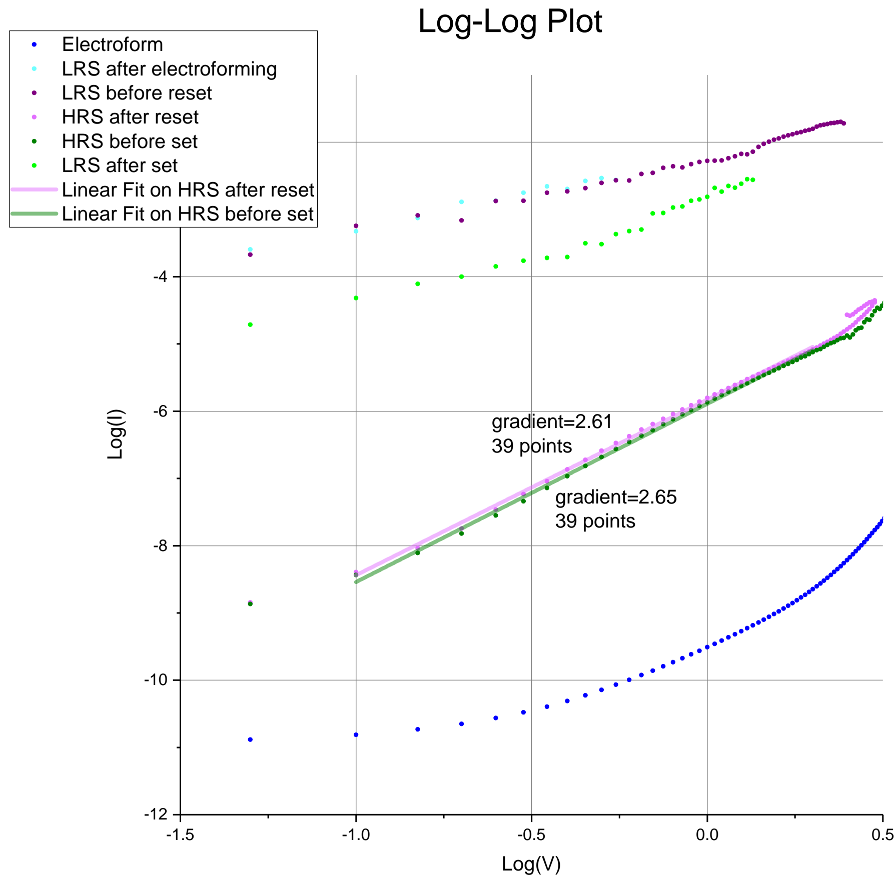
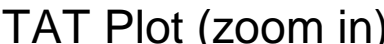


Figure 27: Log-Log plot of IV sweeps of data found in Figure 21. Straight lines with a slope of about 1 should appear when Ohmic conduction occurs. Higher slopes may indicate other processes and are found between 0.37 and 1.35 V with the device in HRS following reset and similarly with the device in the same HRS before set.

Figure 28 shows a $1/V$ v. $\log(I)$ plot of the original data except for electroforming on the pristine device. This is because TAT tunnelling cannot occur on pristine devices where tunnelling conduction mechanisms would dominate. On

reset devices (HRS), however, only a small part of the filament is insulating and therefore TAT may still occur. These plots should show linear regions wherever TAT conduction occurs. Two linear regions appear and can be used to extract TAT barrier height. These two regions have gradients of -6.7 and -1.2 and rely on 22 and 12 data points respectively. The first one occurred while the device was being set and the voltage was between ~2.55 and ~3.55 V. At that time, the device was in a HRS and current was in the tens of microamperes. The second linear segment occurred when the device was being reset and voltage was increasing from ~1.4 V to ~1.95 V. In this second case, the device was in a LRS and current was high increasing from 8 mA to 15 mA. While temperature must have been elevated during this high current phase, TAT should not have temperature dependency. Additionally, current compliance is not used during reset and therefore it has not affected this result.

Starting from the TAT equation, -1.2 leads to a barrier height of 0.038 eV and -6.7 to a barrier height of 0.12 eV. The first is obtained with a device in a LRS while the second with a device in its HRS. It makes sense that the barrier height appears to be higher in the HRS as electron clearly find it harder to flow. These results are comparable to those found in Ref. [55] on equivalent silicon rich silica. Here 0.086 eV (LRS) and 0.52 eV (HRS) were extracted respectively. The HRS value of 0.12 eV also compares to the one of 0.40 eV found in Ref. [2] when it comes to positive biasing and a HRS. However, no values near 0.038 eV are found in Ref. [2]. A study carried out on n-Si/SiO_2 leads to a barrier height of 0.879 eV [88], which given the different material could support the one of 0.12 eV here extracted. Overall, it can be stated that low barrier TAT probably occurs in the devices.



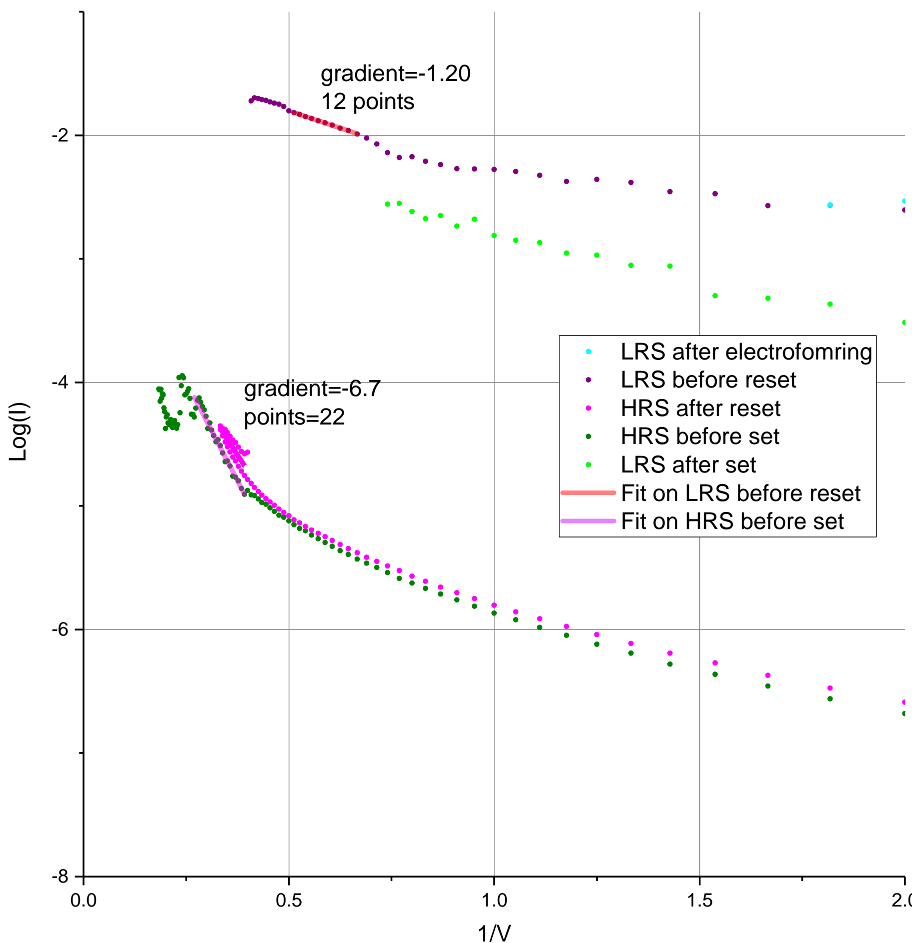


Figure 28: TAT plot (zoom in). Straight lines should represent barrier TAT conduction. Two reliable linear looking regions are found with gradients of -1.20 and -6.7. These are based on 12 and 22 data points respectively. The first region occurred in the HRS before set between ~2.55 V and ~3.55 V. The second region occurred in the LRS before reset between ~1.4 V and ~1.95 V.

Data shown in Figure 21 can also be used to check for thermionic emission conductance. This can be done using the $V^{1/2}$ v. $\log(I)$ plot found in Figure 29. Here one observes three linear regions with gradients of 3.34, 2.67 and 1.68, respectively. Linear fittings used to extract these values relied on 83, 21 and 12 data points respectively and are therefore to be considered reliable.

The 3.34 gradient corresponds to the electroforming forward sweep between ~2.65 V and ~3.55 V. Using the thermionic equation, it was possible to extract a permittivity of $1.61{\times}10^{-10}$, which corresponds to a relative permittivity of 18.18.

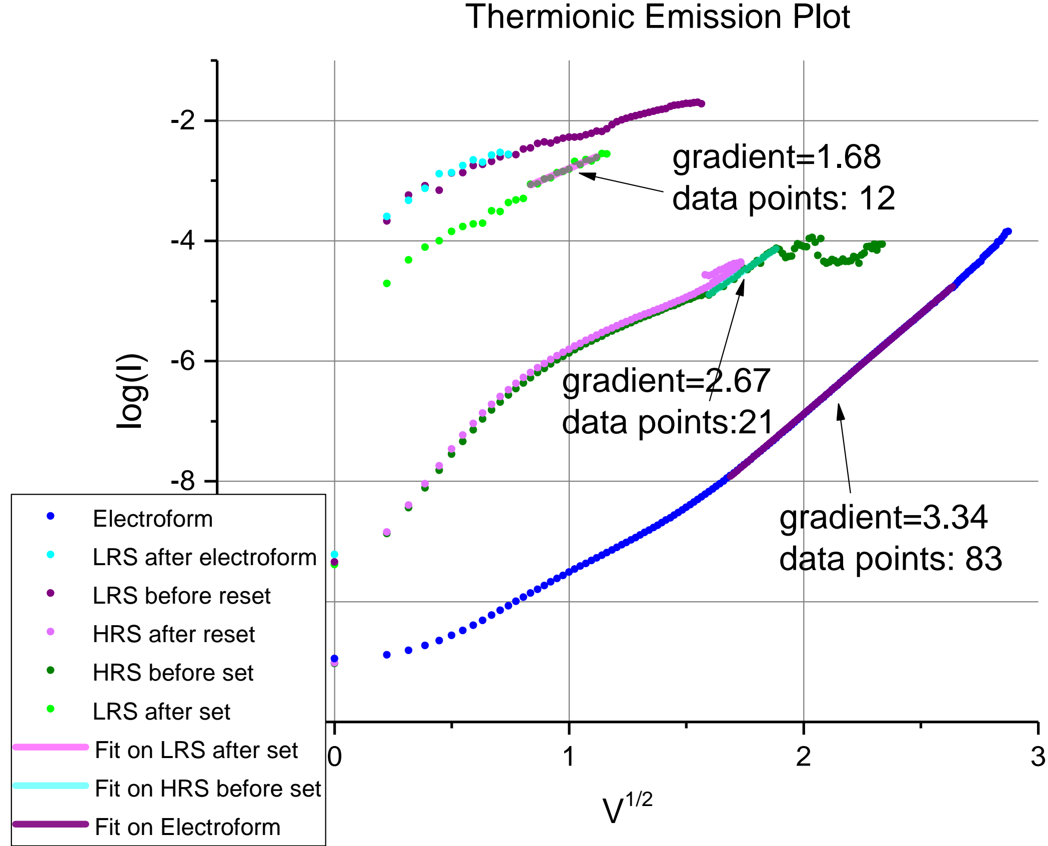


Figure 29 :Thermionic emission plot of the IV curves found in Figure 21. The three linear regions indicate possible thermionic emission conductance.

The 2.67 gradient extracted from the forward set sweep in a HRS, leads to an extracted permittivity of 2.02×10^{-10}, which corresponds to a relative one of 22.81.

Finally, the 1.68 gradient was extracted following set in the LRS. It leads to an extracted permittivity of 3.21×10^{-10}, which corresponds to a relative one of 36.25.

The above values are quite distant from the expected one of 4 for silica and 12 for silicon. As extrapolated from ref. [2], this might confirm that thermionic conduction occurs within devices but with difficulty or along with some other conduction mechanism. The value closest to the expected values, 18.18, was extracted from the electroforming phase. This is not surprising as this is the device state where one would expect thermionic emission to occur.

Data shown in Figure 21 may also be used to analyse whether quantum contact point conduction occurs in the UCL devices. This can be done plotting the same data on a G/G_0 v. voltage plots. From these plots, it is possible to see that whenever a filament did not exist (HRS) conductance was less that $1 \cdot G_0$. This data was discarded as noisy. Data appeared of interest in the HRS found in Figure 21. Figure 30 shows these results. Figure 30a shows conductance by the G_0 during the backwards set sweep. Here zooming in, or looking at raw data, one can see plausible jumps of: 1, 1.5, 3, 6, 5, and 4.5. Figure 30b shows conductance by the G_0 during the forward reset sweep. Here one sees plausible jumps of: 10.5, 13.5, 7.5, 5.5, 5, 7, 6, and 2.5. Figure 30c shows the backwards electroform sweep by the G_0. Here one does not see any jumps by the $G_0/2$ or G_0. All presented jumps are rounded up to the first decimal digit to remove possible instrument error. These results are highly comparable to those found in SiO_x in Ref. [37] showing that also in the $TiN/SiO_x/TiN$ devices in question quantum conductance occurs. As discussed in the same paper, the presence of jumps involving $G_0/2$ reinforces the idea that filaments are of a semi conductive nature.

Overall, the data shown in Figure 21 could be used to extract different types of conduction occurring in the devices. This led to Figure 31, which summarises all results. Quantum conductance is not shown in this figure as it hard to precisely locate where it occurs within the forward reset curve and the backwards set curve.

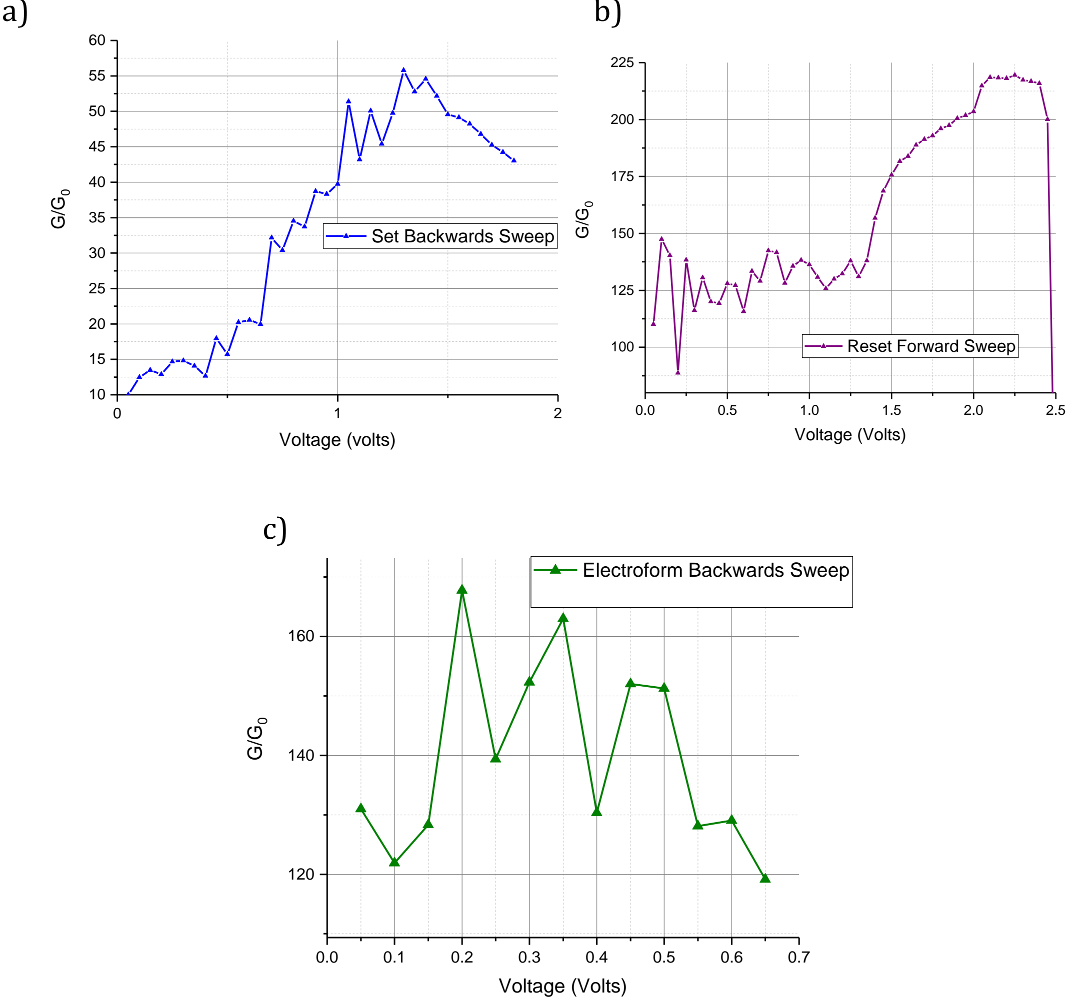


Figure 30: Plots used to detect the existence of quantum point contact conductance in the tested devices.a) shows the set backwards sweep, b) shows the reset forward sweep c) shows the electroform backwards sweep. Evidence of quantum contact conductance emerges; more details may be found in the main text.

Poole-Frankel conduction and possibly thermionic emission appear to be dominant phenomena during initial electroforming. This leads to the idea that, despite current being low, electrons are sufficiently thermally excited by such current to start hopping over triangular energy barriers. This reconciles with an observed inability to electroform devices when using very low current compliance as discussed in Ref. [102].

Trap assisted tunnelling, the conduction method expected to occur in conductive filaments, appears, following initial electroforming, when field is quite elevated. For this process to occur a filament must have been previously created.

It is important to notice that the extracted conduction methods are the consequence of linear region observed on manipulated and collected data whose density changed as a function of voltage. Rerunning data collection using smaller voltage steps would lead to high data density that might well lead to new observable linear regions. This would lead to further insight on conduction methods. Additionally, having multiple processes occurring at the same time has certainly affected extracted constants as well as results. It would be interesting to rerun tests in near absolute zero conditions to remove all conduction methods that rely on the thermal excitation of electrons.

4.1.2 Current Compliance Long Pulsing

Following the statistical determination of switching voltages for $TiN/SiO_x/TiN$ devices in air, it was possible to have these change state using pulses with current compliance. This also helped characterise reset time with this type of pulses by tracing the current going through a device using an oscilloscope.

4.1.2.1 Endurance and Reliability

Pulses with current compliance were applied to the device under test (DUT) with the setup and shapes shown in 3.2.1.1 and 3.3.1.2 respectively. Devices were cycled with variable results.

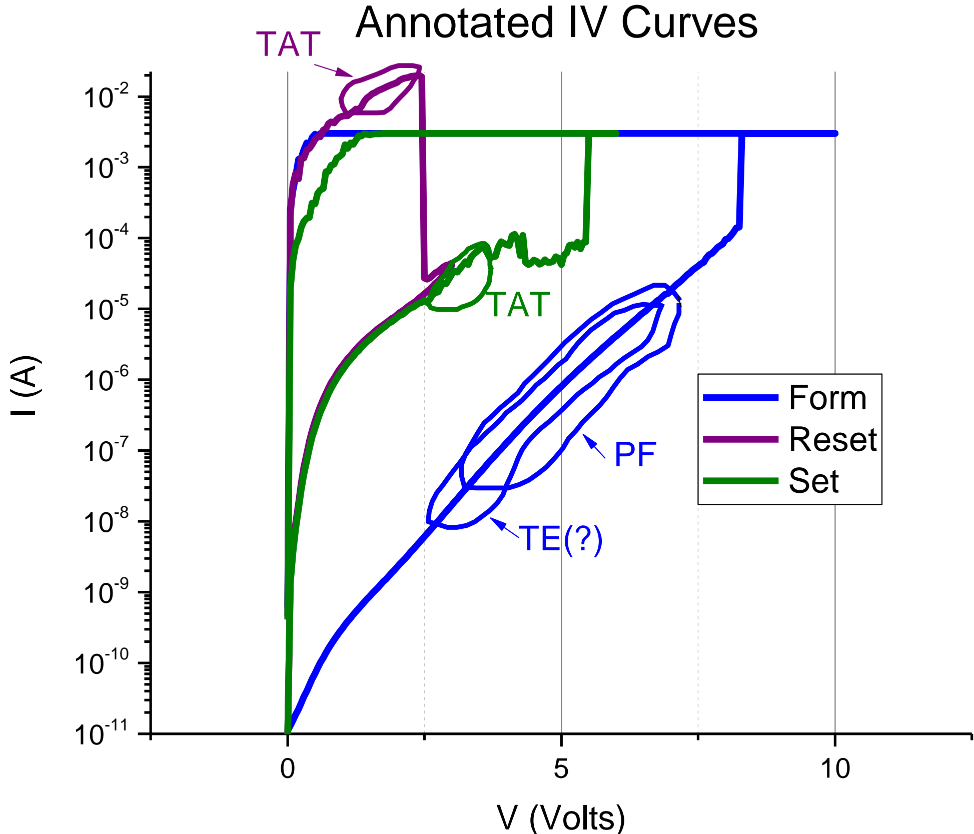


Figure 31: Annotated version of IV curves found in Figure 21. Annotations show the different conduction methods that have been extracted and make physical sense. TAT, possible thermionic emission, and Poole-Frenkel conductions are observed. Some of these processes seem to overlap.

Results from pulsing with compliance show that TiN/SiO_x/TiN devices have performance comparable to devices reported in the current literature. During about 160 cycles (or 320 switching events) a resistance contrast of about one order of magnitude is maintained. Following these cycles, the devices appear to get stuck in an unrecoverable state, which can be a LRS, an intermediate resistance state or a HRS. Cycling was obtained by electroforming at 15 V with 3 mA current compliance, resetting at 2.7 V with no compliance and setting at 7 V with 3 mA current compliance. Pulses with current compliance never reached the specified voltage; voltages above those observed in IV curves should all lead to an identical result. Electroforming and setting voltages were kept higher than necessary because current compliance offers protection from irreversibly damaging a device. Reading

the device state following a pulse was carried out at 1 V as this does not appear to affect the device thanks to their high non-linearity. Information on the length of these pulses is in section 3.3.1.2 (Figure 15); pulses are longer than the time it takes a device to switch state and limited by current compliance or self-limiting resetting. Results are shown in Figure 32 and are comparable with those found in Ref. [44, 60]. It appears that it takes measurements of about 8 cycles for device behaviour to stabilise; at first, HRS readings do not appear very stable. This could be due to an initially non-Ohmic contact present between the probing needle and the device top contact, which slowly breaks down. The first few atomic layers of the top electrode might be oxidised by atmospheric oxygen. In the same figure, the LRS appears to have greater fluctuation than the HRS; this is due to the logarithmic y-axis of the plot.

The most common form of device failure led to devices stuck in a very low resistance state (about 150 Ω). In this case, it seemed impossible to break a filament. It may be that the region around the filament is sufficiently depleted of oxygen to prevent oxygen vacancies from being re-oxidised. Similarly, a filament could become unrecoverable following hard electric avalanche breakdown, which would follow a sufficiently high current and consequent heating. This kind of insulator breakdown follows the fast acceleration of electrons and electron holes towards the two electrodes. This fast process knocks out bound electrons free leading to more charge carriers and therefore an even higher current. This further foster the phenomenon leading to an avalanche [100].

On other occasions, devices would get stuck in a temporarily or permanently unrecoverable intermediate resistance state. This might correspond to a filament of

an intermediate strength, which would require a higher current to reset or a higher electric field to set. In physical terms, a strong filament could correspond to a wider one. In these tests, programming pulses are non-adaptive and do not change per the state of the DUT.

Devices stuck in high or medium resistance states often appeared to have insulating deformed regions underneath the probing needle (Figure 48); this was the case with the device whose cycling is shown in Figure 32. Moving the needle to a clean TiN region in such cases would recover operation. Similarly, this kind of failure could be avoided by frequently moving the needle on the contact surface to avoid the expanding deformed regions. These features are clearly insulating, or disrupt electrical continuity to the filament. Hence, it is not oxide or filament failure, rather a problem with the contact.

Low voltage read pulses were used to verify the state of devices over time at room temperature. A device was forced to enter either a LRS or a HRS and its resistance was sampled over time at 85°C. This elevated temperature was used to emulate the temperature of an integrated circuit during regular operation as well as to provide the device with accelerated aging. Devices were sampled with a 0.5 V pulse (a voltage that does not affect device state and does not show high noise) 100 times every 100 seconds or 10,000 times in total. Over 10,000 seconds, TiN/SiO_x/TiN devices appear to be stable as shown in Figure 33. The two chosen resistance states happened to have resistance separation larger than usual. Exceptions of this kind were not uncommon in the first cycles. This is comparable to what is found in the literature in Ref. [44].

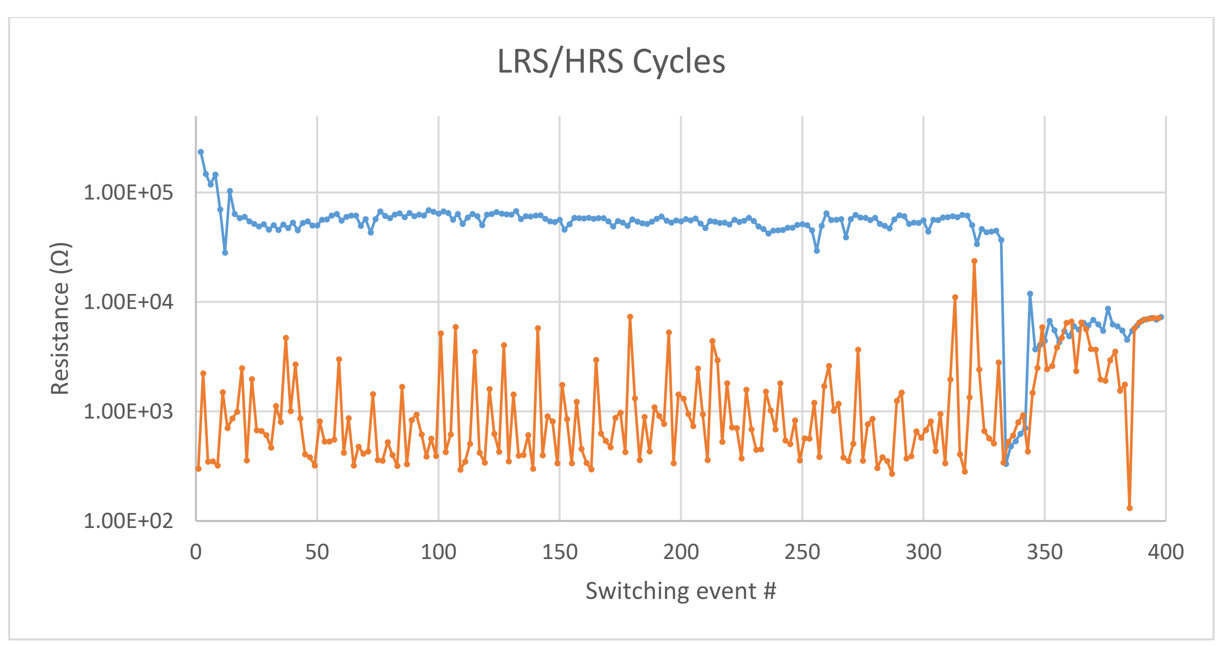


Figure 32: Cycling results from a TiN/SiO$_x$/TiN device showing consistent separation of at least one order of magnitude between states for about 160 cycles. This device was cycled 200 times using pulses with current compliance to test endurance. Electroforming: 15 V with 3 mA of current compliance; resetting: 2.7 V with no compliance; setting: 7 V with 3 mA of current compliance; reading 1 V.

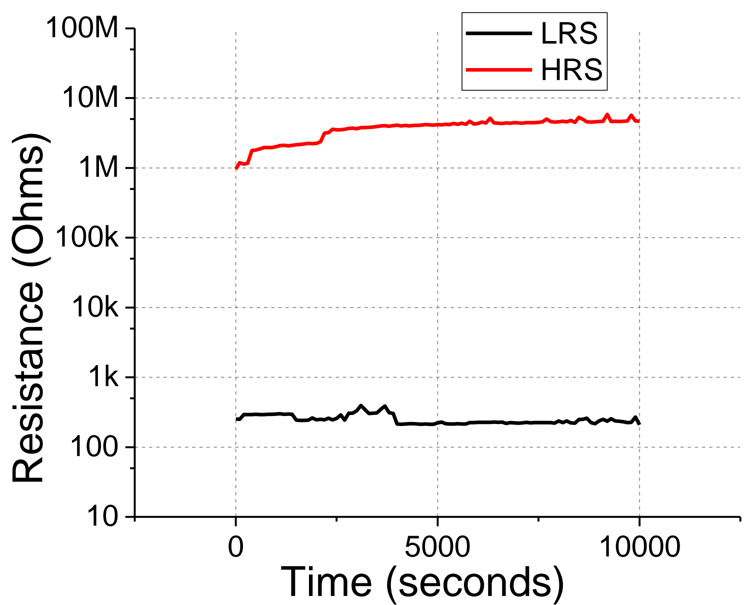


Figure 33: Reliability over time of a TiN/SiO$_x$/TiN- device tested at 85°C. The two resistance states were measured at different times. Devices appear to be stable in both LRS and HRS states over at least 10,000 seconds.

4.1.2.2 Oscilloscope Measurements

Using the setup show in Figure 14, it was possible to carry out studies of device behaviour when it comes to timing using SMUs. Figure 34 shows how voltage drops across a device as it sets. Here a 7 V pulse was used and a current compliance limit of 3 mA (more information in section 3.3.1.2). The device resistance starts dropping at about 3.3 V, rather than 7 V. This shows that, when setting or electroforming using an SMU, the specified voltage is not important if it is sufficiently high. The voltage therefore abruptly drops to about 2 V in about 150 nanoseconds. While this may well represent the main setting phase, resistance keeps decreasing over another 2 microseconds while the voltage is kept relatively low. This slower phase, it can be assumed, leads to a very strong non-volatile set. At around 25.25 microseconds, one sees a temporary increase in voltage, and therefore in resistance. This may well correspond to a competing reset phenomenon occurring due to filament heating. This phenomenon is however temporary as the strong field effect prevails and the electroforming process continues.

Current compliance appears to track filament setting (or electroformation) by quickly reducing voltage. This avoids hard meltdown while delivering an optimal voltage during the process. This quick process appears to lead to filaments whose strength is proportional to current compliance. This is because the amount of time a voltage is applied across the device is proportional to current compliance.

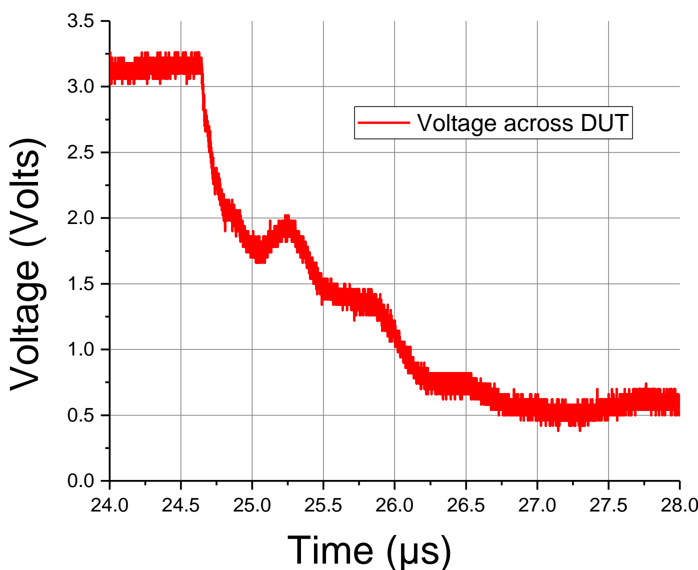


Figure 34: V/t plot of set. As a device electroforms or sets, its resistance gets smaller and the current compliance circuit present in the SMU adjust its voltage output to make sure current stays within the pre-set limit. Here one has a successful set using a nominally 7 V pulse with 3 mA compliance. 7 V is never reached across the device.

Figure 35 shows what happens while resetting (from 5k Ω to 30k Ω) a device with a filament that was previously electroformed or set using an SMU and current compliance. The plot tracks current by measuring voltage across a 90 Ω series resistor as well as the voltage applied by the Keithly 4200-SCS. This directly measured voltage should always be 2.7 V but clearly the SMU negative feedback system struggles at ensuring this. The strong filament requires an extended amount of time to fully reset: about 2.5 milliseconds. Some competition between setting and resetting phenomena may also be observed where the device partially sets and may explain this. The applied electric field accelerates oxygen ions away from silicon bonds and towards the device anode. This leads to a more conductive filament through trap assisted tunnelling. However, at the same time, current heats up this filament as well as the anode providing oxygen ions with enough energy to move and bind back with silicon dangling bonds. It is also interesting to observe that no resistance change occurs during the first 1 millisecond of the programming pulse.

This may mean that it surprisingly might take time for the filament to heat up. The input voltage trace should be stable at 2.7 V but oscillated probably due to noise or SMU instability.

Devices electroformed with current compliance have very stable filaments that naturally require long reset times. As explained, the applied bias is reduced as the filament grows in length and width allowing for a reliable filament to electroform without risking hard breakdown. It will be shown in section 4.1.3.3 that devices can also be electroformed much more quickly and therefore also reset more quickly. These devices will have weaker but nonetheless stable filaments that can be used in data storage applications.

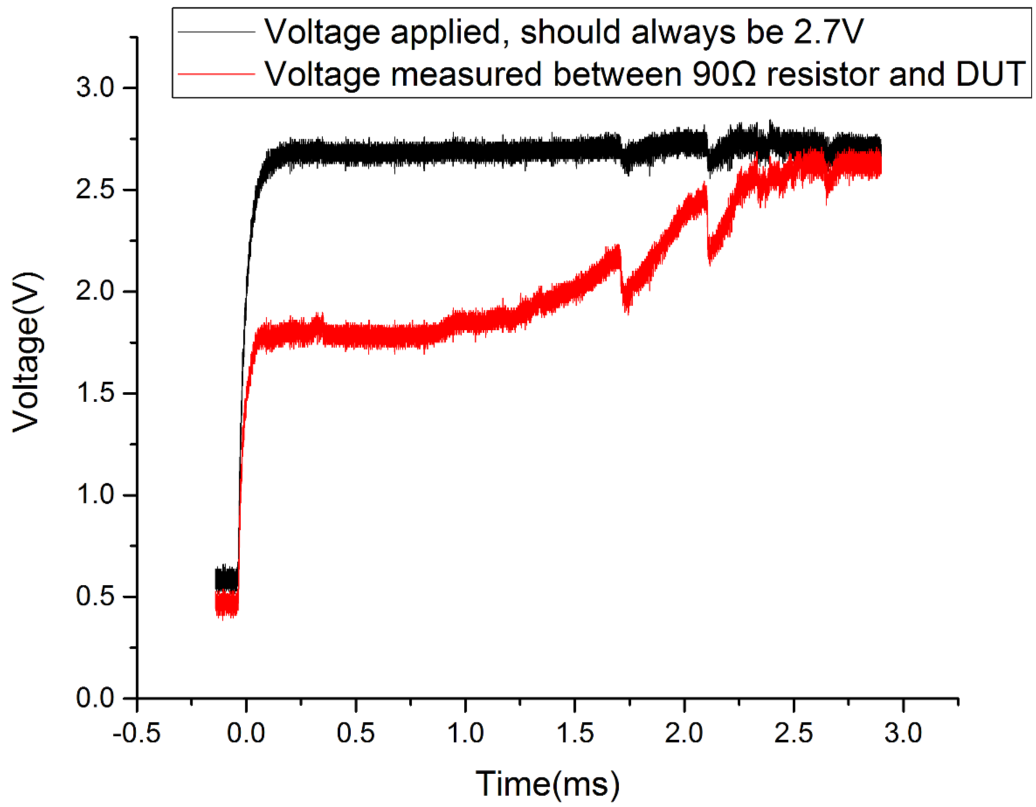


Figure 35: Voltage measured across a 90 Ω resistor in series with a DUT being reset by SMU. This device was previously electroformed by SMU. The reset process appears to be very slow. The device resistance went from 5k Ω to 30k Ω. The black trace is the voltage applied by the instrument. This voltage should be at 2.7 V through the plot thanks to negative feedback. However, the instrument appears to struggle in ensuring this.

It must be noted that these voltage traces are affected by all parasitic capacitances and resistances present in the system. These dynamically change as the device electroforms and may introduce response delay.

4.1.3 Fast Nanosecond Pulsing

No deformed regions were observed to develop while switching with the switching methods covered in this section. Conversely, the previously described use of current compliance, led to deformed regions. A possible explanation to this change in behaviour is the way that current compliance is implemented during sweeps or long pulse programming. Current is limited by gradual voltage reduction (as shown in Figure 34) to ensure that the specified current limit is not exceeded. This might lead to the electroformation of a very strong filament, which would imply a large oxygen ion migration toward the anode. Instead, fast pulse programming might only allow for the electroformation of weaker (or physically smaller) filaments as the programming voltage is abruptly reduced once the filament is electroformed or set. This type of filament would lead to a smaller amount of oxygen ions migrating toward the anode.

4.1.3.1 Regular Switching

Regular switching involved user interaction after each pulse; a pulse magnitude and length would be specified. This pulse would be applied, a read pulse would be applied, and the sampled resistance would be read. This information was used to decide which pulse to apply next. The rise and fall times of all pulses was 20 nanoseconds. As shown in Figure 39, PMU pulses do not seem to show significant overshoot. Reading was done using 500 nanosecond, 1 V pulses. Given that human interaction was needed, tens of seconds passed between pulses. Given that

programmed states did not change between manually issued commands, it can be concluded that devices showed data retention at least in the tens of seconds.

It was possible to electroform with a 75 nanosecond pulse at 10 V. This brought the resistance of a $TiN/SiO_x/TiN$ device from $1.5x10^5$ Ω to $2.44x10^2$ Ω. Resetting was then possible with a much longer 5 microsecond pulse at 6.5 V. This increased resistance from $3.76x10^2$ Ω to $7.71x10^4$ Ω. The device voltages of these pulses were, in reality, much lower while the device was in a low resistance state due to impedance mismatch. It was then possible to keep on switching the device between similar states using 10 V and 6.5 V pulses of about 96 nanosecond and 5 microsecond durations, respectively. Electroforming appeared to be the most critical step to allow for proper switching operation. Pulses too short in length or small in amplitude would not change the state of the device (or did so only temporarily). Pulses of excessive duration or voltage would put devices into an unrecoverable LRS due to hard breakdown. Traces of pulses of the above type may be found in section 4.1.3.3. No overshoot is visible within the oscilloscope 200 picosecond resolution.

Some device states appeared to be volatile, as subsequent readings produced different resistances. This would imply that, following certain pulses, data retention was less than a few tens of seconds, or the time taken to issue the next command. This phenomenon is further discussed in the next subsections.

4.1.3.2 Shaped Pulse Switching

Switching using waveforms was carried out to understand whether devices can be quickly switched between states without using current compliance. These shaped pulses were defined using start and stop voltages and times. The method described in the previous section required user interaction following each

pulse; this added unwanted delay. Circuit programmers without current compliance would be favourable thanks to a smaller number of transistors, which would lead to smaller footprint and power consumption.

Switching was indeed possible with some devices, as shown in Figure 36. Two distinct devices will be described. In the first case a device could be electroformed with a 200 nanosecond pulse of 15 V and subsequently switched between states with 75 nanosecond pulses of 15 V. This was probably possible because of the drop in applied voltage while the device is in its LRS. This drop in voltage occurs due to impedance mismatch. The voltage specified on the PMU is only delivered on high resistance loads. Once the load resistance drops, so does the applied voltage, which goes well below the specified one of 15 V as per Eq. 3.1. The fact that it appears possible to electroform, set and reset using an identical voltage would allow for further simplification of programming circuits. In the literature, there are many reports of successful switching using pulses with lengths of tens of nanoseconds [42, 43, 64, 99]. However, there are no reports of techniques that allow the use of identical voltages to carry out all three switching operations: electroforming, resetting and setting. In the second case, devices could be set and reset using identical pulses. However, in this case electroforming was carried out with a pulse different in magnitude from the subsequent ones. Electroforming was accomplished with a 100 nanosecond pulse of 15 V and setting and resetting by 65 nanosecond pulses of 6 V. As explained, this is thanks to change in impedance mismatch. Filaments electroformed without current compliance appear to be resettable with much faster pulses than those electroformed with current compliance. Current compliance, by reducing voltage as electroforming occurs, leads to ‘strong’ filaments.

Other devices presented in Figure 36 show that it is possible to use identical pulses to gradually change a device state and toggle it between states. A 13 V, 110 nanosecond pulse electroforms a device, which then gradually increases and decreases in resistance with identical pulses of 10 V, 110 nanoseconds. Another 'noisier' device toggles between states thanks to 15 V, 65 nanosecond pulses following electroforming with a pulse at 15 V, 220 nanoseconds. Another device was electroformed with 10 V, 110 nanoseconds and subsequently reset and gradually set with identical 6 V, 65 nanosecond pulses.

Figure 36a shows some exponential fits to voltage vs. time recorded data in an attempt to find patterns. Here, resistance tends to increase following the first read pulse that follows a programming pulse. This first reading and those that follow appear to fit Eq 4.2 with A being negative.

$$R = R_0 + Ae^{-t/\tau} \text{ Eq. 4.2}$$

In the case of "Fit #1" one has $R_0 = 296.53\Omega$, and $\tau = 3.49\times10^{-7}\ sec$. τ describes the first order system time constant, which, may be directly related to the time it takes parasitic capacitance (RC) to charge. As resistance increases, "Fit #2" can be used. Here, one has $R_0 = 1397.09\ \Omega$, and $\tau = 6.18\times10^{-7}\ sec$. Importantly, an increase in base resistance and a consequent increase in time constant τ occur. This makes sense as an increase in device resistance would slow down the time it takes parasitic capacitance to charge up. In the first case, one calculates a parasitic capacitance of $\tau = 3.49\times10^{-7} sec = RC = 296.53\Omega\times C$ so $C = 1.18\ nF$. In the second case, $C = 0.44\ nF$.

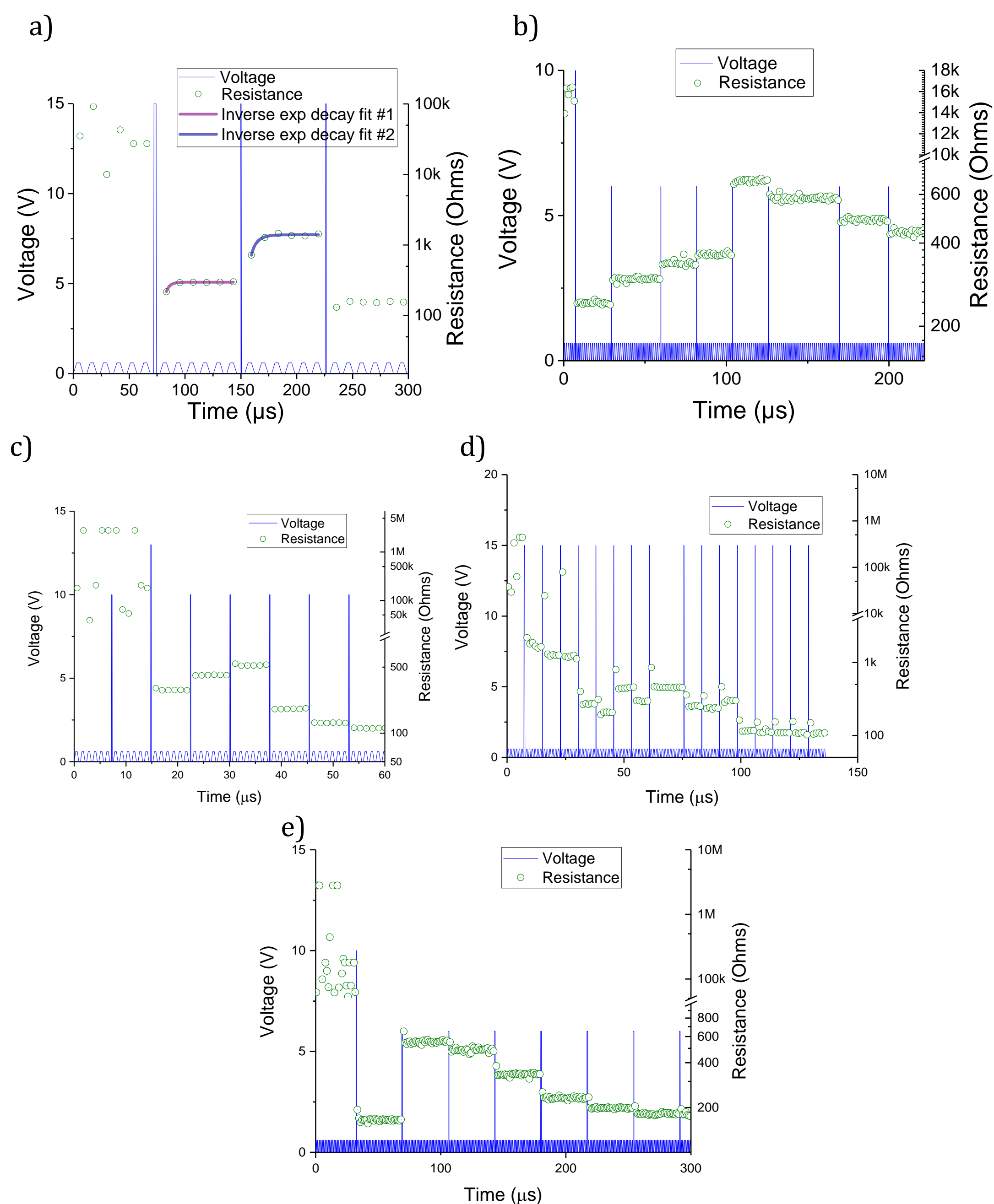


Figure 36: Switching between states using identical pulses. a) following electroforming with a 15 V, 200 ns pulse, the device is switched with 15 V, 75 ns pulses. Some exponential fits are shown. b) following a successful 10 V, 120 ns electroforming pulse, identical 6 V, 65 ns pulses were used to gradually bring the device into different resistance states. c) A 13 V, 110 ns pulse electroforms the device, which then gradually increases and decreases in resistance with identical pulses of 10 V, 110 ns. d) this noisier device toggles between states thanks to 15 V, 65 ns pulses following electroforming with a pulse at 15 V, 220 ns. e) This device was electroformed with 10 V, 110 ns and subsequently reset and gradually set with identical 6 V, 65 ns pulses.

Given the above extracted capacitances, it is interesting to see whether devices should have a capacitance in the nano-Farads. Capacitance is defined as $C = \epsilon_0 \epsilon_r \frac{A}{d}$. Using an area of 400 µm ×400 µm (the filament width is negligible), a distance

of 37 nanometres (see section 3.1) and an ϵ_r between 4 and 12 (for silica and silicon, respectively; see section 4.1.1), one expects a parasitic capacitance between 0.153 nF and 0.459 nF. The silicon extracted capacitance appears to be extremely close to the extracted capacitance of 0.44 nF in the previous paragraph. This leads to the idea that the change in resistance following the first read pulse might well be due to devices charging up rather than chemical or physical change.

It is also possible to use different voltage PMU pulses to cycle a device between states [53]. This was mainly carried out by a former UCL student Konstain Zarudny. Some results are shown in Figure 37. Here, devices can be set with pulses of 7 V and 70 nanoseconds and reset with pulses of 3.5 V and 100 nanoseconds.

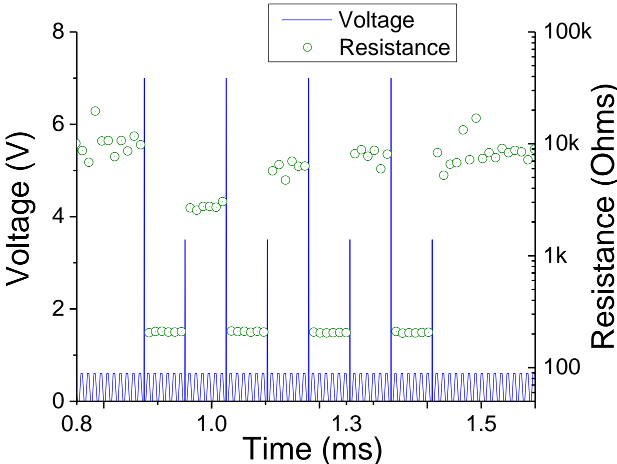


Figure 37: Device cycling using fast pulses without current compliance. Pulses of 7 V and 70 nanoseconds were used to set and pulses of 3.5 V and 100 nanoseconds were used to reset. Six read pulses were used between each programming pulse which occurred every ~30 milliseconds. These applied voltages are higher than those seen by the device due to impedance mismatch.

Unfortunately, the above results do not seem to apply to all devices, due to fabrication process variation. Some devices will only electroform into an unrecoverable LRS. The main challenge associated with this kind of device programming lies in the fact that prior knowledge of switching time is required to successfully switch a device and avoid unrecoverable LRS states. A pulse that is too

short might only lead to an unstable LRS and one that is too long to an unrecoverable LRS. It is now interesting to look at electroforming dynamics when using fast pulses and to also study the behaviour of pulses which lead to unstable or transient states as these states can find use in analogue computing and applications involving the analogue emulation of biological neuronal synapses.

4.1.3.3 Oscilloscope Tracking of Electroforming Behaviour

It is interesting to track the voltage across the DUT during a waveform as described in 4.1.3.2 [39]. Equation 3.5.1 in section 3.3.1.3 uses impedance mismatch to describe a drop in voltage as device resistance drops.

Figure 38 shows the full waveform that was sampled by the oscilloscope during this experiment. The first pulse (100 nanoseconds, 10 V) did not electroform the device but the second pulse (110 nanoseconds, 15 V) did. The third and subsequent pulses are programmed identically to the first. The apparent change in amplitude is due to impedance mismatch following electroforming. The recorded negative voltages are a sign of impedance mismatch between the DUT impedance and the PMU output impedance; this can be observed after electroforming, which occurred after the 2nd pulse. The equivalent transmission line coefficient must be negative. It is interesting to observe that the specified voltages are never actually reached across the DUT. This may well be due to the 1 MΩ/channel input impedance of the oscilloscope, which is inserted in parallel to the DUT and leads to impedance mismatch. The thick line along 0 V represents low voltage read pulses. Also these read pulses (1 V, programmed) reduce in measured amplitude following electroforming due to impedance mismatch. Once programming pulses end near the end of the plot, so do read pulses. Therefore, the thickness of the line at the very end

of the plot shows the amount of noise present in the system which may well be due to the use of BNC-terminated coaxial cables instead of triaxial cables and coaxial Ts.

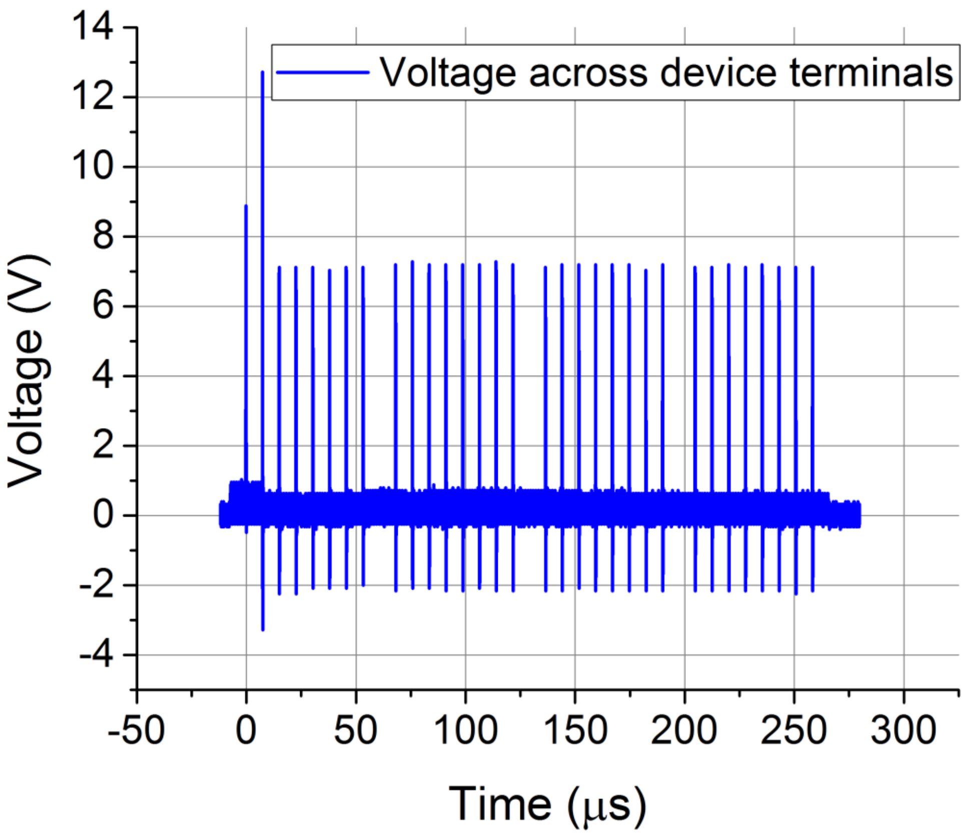


Figure 38: Voltage measured across the DUT as a positive pulse sequence was applied. The recorded negative voltages are a sign of impedance mismatch between the DUT impedance and the PMU output impedance; this can be observed after electroforming, which occurred after the 2nd pulse. The sequence uses 110 ns pulses programmed to be 10 V, 15 V and 10 V onwards. The 1 V read pulses along 0 V appear as a thick line, which reduces in amplitude following electroforming due to impedance mismatch.

Figure 39 shows a zoomed in and high resolution (200 picosecond resolution) voltage transient during part of the 2nd and electroforming pulse found in Figure 38. Resistance (and therefore voltage) does not simply decrease but it also temporarily increases. Such trend shows the presence of competing phenomena over about 25 nanoseconds; the device appears to undergo some thermally induced resetting (silicon dangling bonds reacquire oxygen ions) while the electric field effect electroforms the device (oxygen ions are taken away from silicon and moved

towards the anode). The two main electroforming steps have been highlighted in yellow; the competing reset phenomenon is in between these two.

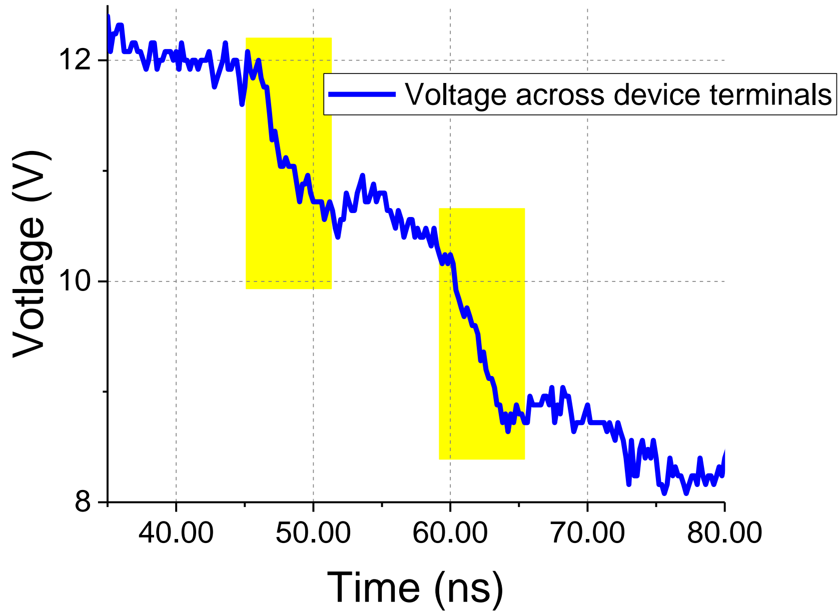


Figure 39: Voltage across terminals of a device being electroformed using a fast pulse without current compliance. The drop in voltage corresponds to filament electroformation and occurs due to impedance mismatch. In this case, it appears that the phenomenon is made up of two steps (highlighted) with competing resetting elements (such as a thermal effect from Joule heating) appearing between the two highlighted regions.

It must be noted that these voltage traces are affected by all parasitic capacitances and resistances present in the system. These dynamically change as the device electroforms and may introduce response delay. Therefore, an electroforming time of 25 nanoseconds must be carefully looked at from a worst-case scenario point of view. Parasitic capacitance probably make the device appear slower than it is.

4.1.3.4 *Transient Behaviour of the Resistance State*

Waveform tests were ran during which the a device state was sampled for a longer time following a programming pulse. This was achieved by using many quick read pulses. The observed transient behaviour appears similar to the one found in biological neuronal synapses. This opens doors to using ReRAM devices for the

analogue emulation of networks of these cells often used in implementing artificial intelligence.

In some cases, following a switching event, a device would gradually return to its previous state. Relaxation of both reset and electroform/set states occurred with a 10 V, 100 nanosecond electroforming pulse followed by 6 V, 65 nanosecond switching pulses. These electroforms relaxation seemed to occur over times of 10 microseconds. In some other specific cases, a device seemed to electroform only following many stimuli, which made its resistance gradually increase. This occurred by using 15 V pulses. The first electroforming pulse had a length of 200 nanoseconds and the pulses that followed of 75 nanoseconds. Figure 40 shows these behaviours.

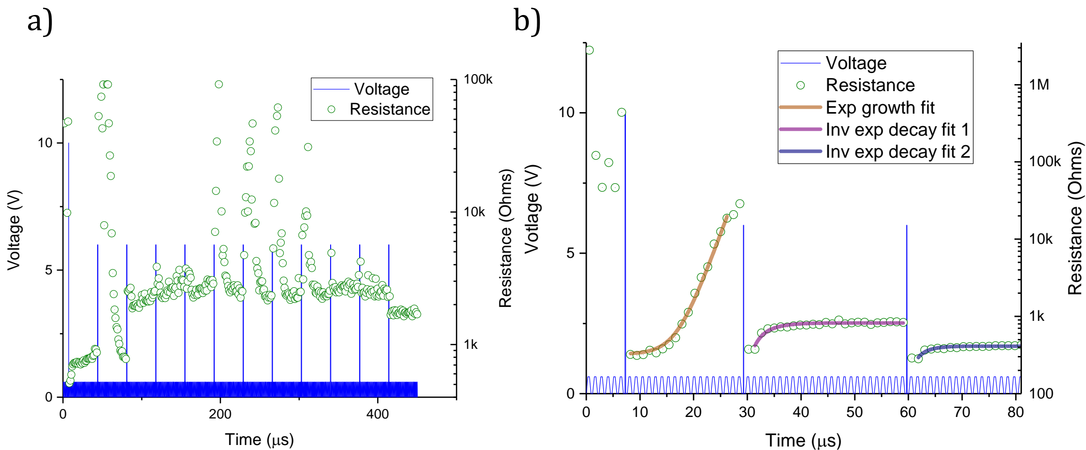


Figure 40: Temporary device states: a) temporary electroform with relaxation followed by more similar events (10 V, 100 ns followed by 6 V, 65 ns). b) temporary electroform followed by relaxation to pristine state (same pulse sequence as left). Equations were fitted to extract time constants.

Available resistance data found in Figure 40b was used to extract time constants; Figure 40a appeared to be too noisy. Eq. 4.2 was used for the second and third plot segments (with a negative A) and Eq. 4.3 for the first segment (positive A). The fit on the third segment leads to $R_0 = 322\ \Omega$, and $\tau = 2.29\times10^{-6}\ sec$. The fits on the first two segments lead to $R_0 = 374\Omega$, $\tau = 2.92\times10^{-6}$ and $R_0 = 410.47\Omega$, and $\tau = 2.12\times10^{-6}\ sec$, respectively.

$$R = R_0 + Ae^{t/\tau} \; Eq.\ 4.3$$

Therefore, the time constants for the exponential phenomena seem to remain quite consistent throughout the different device states. This probably means that during this type of behaviour, devices do not undergo significant physical, and therefore parasitic capacitance, changes compared to regular switching.

The above time constants and resistances, lead to the following equivalent capacitance: $C_{first} = \frac{2.92\times10^{-6}}{374} = 7.8\ nF$, $C_{second} = \frac{2.12\times10^{-6}}{410.47} = 5.16\ nF$, and $C_{third} = \frac{2.29\times10^{-6}}{322} = 7.11\ nF$. Given that the device parasitic capacitance was previous calculated to be between $0.153\ nF$ and $0.459\ nF$, it is possible to conclude that the observed phenomena are not due to device charging/discharging.

Literature exists on this type of behaviour in the emulation of brain memory. Here, the behaviour found in TiN/SiO_x/TiN devices is faster than the one found in some published material [49, 51] as these devices 'forget' their state within tens of microsecond compared to seconds. Having devices that forget within a wide range of time scales is sought after in neural networks to achieve more realistic emulation of the mammalian brain [51].

During the above experiments, different devices showed different behaviour. However, similar behaviour was observed in many devices. While this is unwanted in deterministic digital system, the same is not true when using ReRAM devices in neural-like networks which emulate aspects of biological brain behaviour (see section 3.3.1.4). Stochasticity (or stochastic noise) is needed in such networks in order not to get stuck high or low states. Similarly, plasticity found in neuronal synapses is also found in UCL ReRAM devices when one observes that device behaviour depends on its past. This, allows for the self-rewiring of networks leading

to the complete or partial exclusion of devices that are not operating sufficiently well. Therefore, if out of stochasticity, certain devices are not up to expected performance, they can simply be excluded. These networks do not attempt to produce exact results like deterministic binary system. The objective is to produce acceptable results very quickly and through extremely limited power consumption. Given the above, results produced by UCL devices are very appealing [48, 50, 108-110].

4.2 Electrical Behaviour in Modified Atmosphere

Electrical in situ experiments involving TiN/SiO_x/TiN devices were carried out in different atmospheres described in 3.3.2 and here summarised:

- High Vacuum
- $^{18}O_2$-rich atmosphere at 400 mbar, which translates into an O_2 partial pressure twice as high as in air.
- $^{16}O_2$-rich atmosphere at about 1 atmosphere.
- $^{14}N_2$-rich atmosphere at about 1 atmosphere

Different switching behaviours may now be looked at in the above atmospheres. Results are grouped by the type of switching event. Each subsection starts with a summary of results, which are then discussed.

4.2.1 Electroforming

Figure 41 shows electroforming IV curves in different atmospheres. Successful electroforming was possible in all atmospheres except the $^{14}N_2$-rich one. In this atmosphere, electroforming would not occur up to a tested voltage of 20 V.

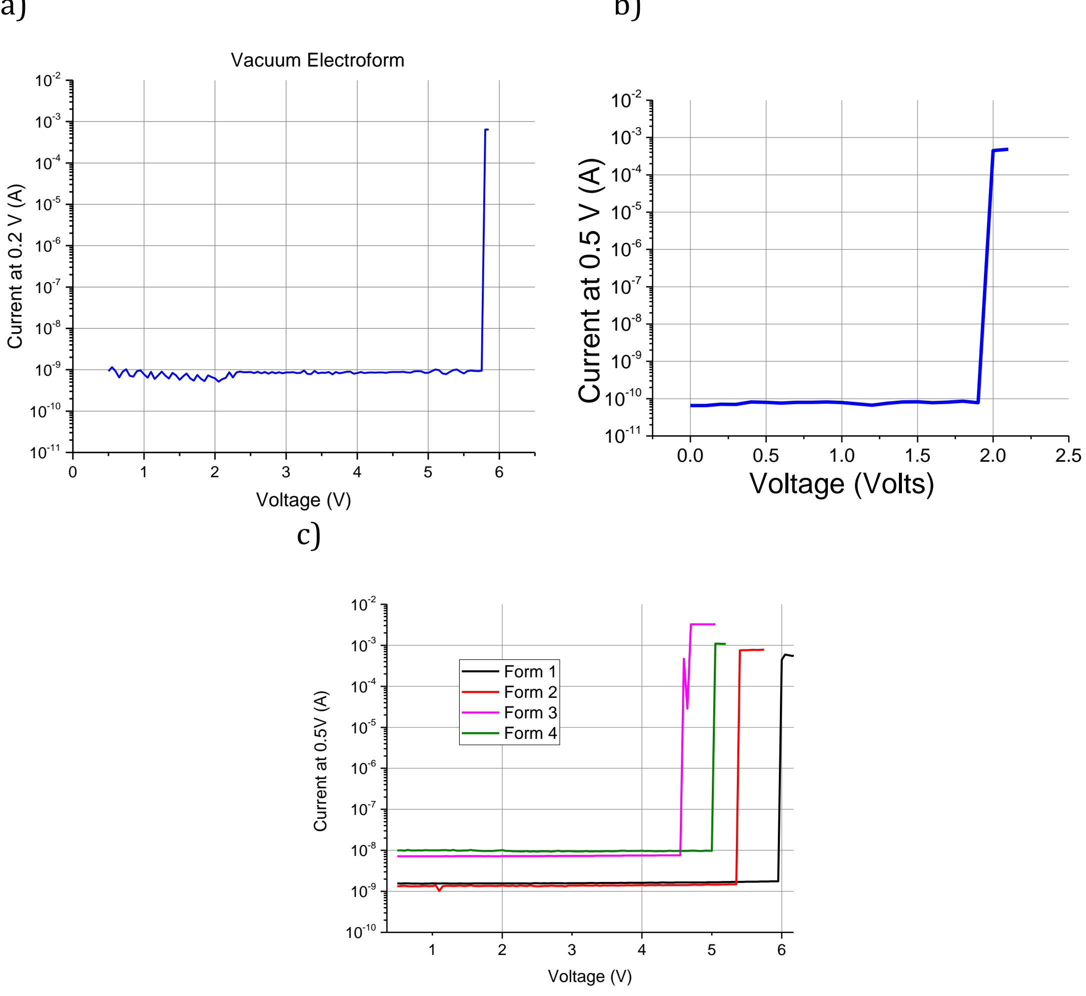


Figure 41: IV curves showing electroforming in different atmospheres. a) High Vacuum (5.8 V, 1 mA current compliance), b) $^{18}O_2$-rich (2 V, 1 mA), c) $^{16}O_2$-rich (between 4.7 and 6 V, 1 mA). Electroforming was suprisngly not possible in the $^{14}N_2$-rich atmosphere.

All collected electroforming data points can be plotted on a column scatter plot. For comparison, it is important to also include results from in-air switching previously presented in section 4.1.1. Results are shown in Figure 42 and may be directly compared. No significant difference is obvious between different atmospheres. However, it does appear that frequent device breakdown was occurring in high vacuum at relatively high voltages.

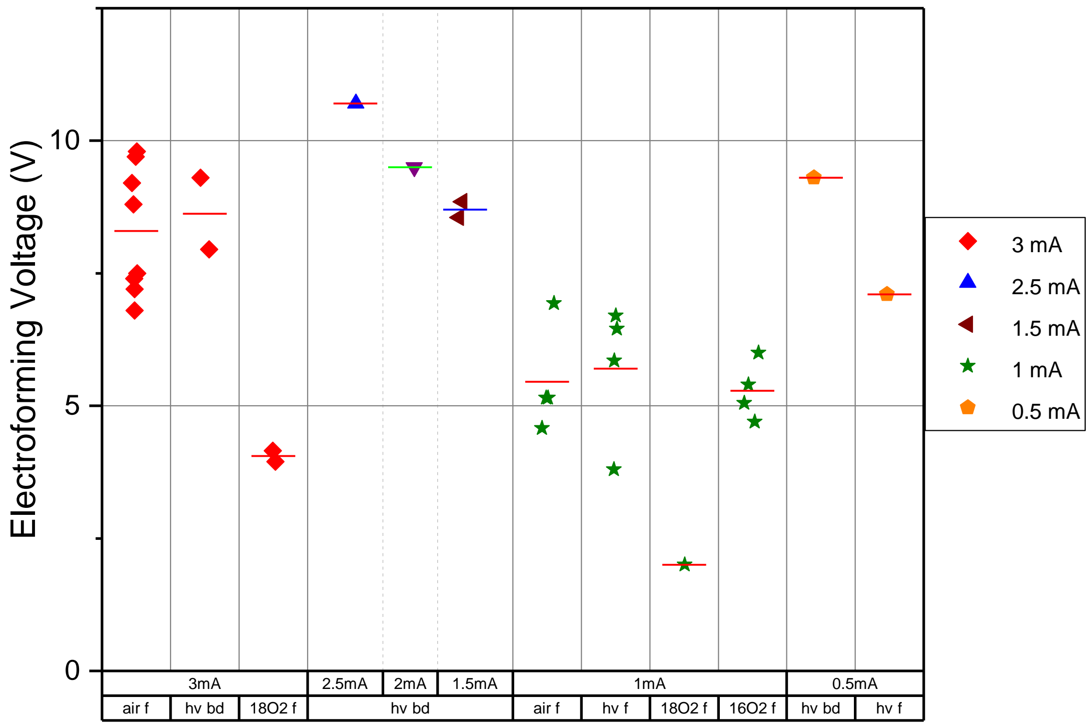


Figure 42: Column plot showing collected electroforming voltages. Different symbols represent different current compliances as shows in the legend. "air f"=air electroform, "hv bd"=high vacuum break down, "18O2 f"=$^{18}O_2$ electroform, "hv f"=high vacuum electroform, "16O2 f"=$^{16}O_2$ electroform.

4.2.1.1 High Vacuum

Figure 41a shows the pseudo IV curve (section 3.3.2.1) of a device which electroformed in high vacuum at ~7 V with 0.5 mA current compliance and could therefore be reset in the same atmosphere; this device did not enter an unrecoverable LRS. This was quite unusual as most devices on which electroforming was attempted in a high vacuum did (6 out of 10). Using a current compliance lower than the one generally used in air (3 mA) and smaller voltage steps seemed to only partially help. Stopping sweeps right after a switching event occurred generally helped prevent device hard breakdown. Keeping in mind the previously described oxygen migration model, it could well be that a low external pressure allows oxygen to escape devices more easily and therefore more quickly and in greater quantity compared to air. This is compatible with oxygen vacancy theory which relates oxygen movement to filament electroformation [55, 102] as faster oxygen

movement should lead to faster electrical changes. This could mean that, in high vacuum, device resistance drops more quickly than in air and that the negative feedback circuitry in the Keithely 6430 is not quick enough at reducing applied voltage. The resulting current being well above compliance would allow for hard breakdown [27] to occur often and unexpectedly. Meltdown risk was present at every programming pulse for which current compliance was necessary. However, this change in behaviour is also consistent with some literature [75], which concludes that lack of atmospheric oxygen negatively affects switching from HRS to LRS in nickel-oxide devices. Similarly to in air operation, hard breakdown could be detected by device resistance dropping below a few hundreds of ohms. Satisfactory electroformation would lead to a resistance of about 600 Ω, as in air. Given the small amount of devices that were successfully electroformed, voltage statistics were not computed.

Overall, vacuum electroforming occurs at voltages comparable to those seen in air but hard breakdown appears more likely than in air. Out of 12 randomly picked electroform sweeps in air, none showed evidence for device meltdown (section 4.1.1.1). In high vacuum, however, 6 out of 10 electroforming sweeps led to hard breakdown.

Some SiO_x devices appear in Ref. [43, 93-95] which seem to only switch in vacuum. Unfortunately, these results cannot be compared to those obtained with UCL devices. This is because UCL devices have filaments that develop within the active layer and not on its outer and exposed surface.

4.2.1.2 *$^{18}O_2$-rich atmosphere*

Successful electroforming, shown in Figure 41b, was easy to achieve in $^{18}O_2$-rich atmosphere as, unlike in high vacuum, no devices underwent hard breakdown unexpectedly. This behaviour is not unexpected as the 400 mbar total atmospheric pressure probably allowed for oxygen movement to occur at a rate only slightly faster to the one found during in-air switching. Lower electroforming voltages than in air were observed on the three devices that electroformed. However, due to the small sample size, no conclusions should be drawn out of this result. The small number of devices tested is due to difficulty in achieving successful electrical contact using copper whiskers and a shortage of TiN/SiO_x/TiN pristine devices to test.

4.2.1.3 *$^{16}O_2$-rich atmosphere*

Electroforming devices in $^{16}O_2$-rich atmosphere occurred, as shown in Figure 41c at voltages ranging from 4.7 to 6 V and without unexpected hard breakdown. With the few tested devices, one observes behaviour close to the one expected in air. Per the previously described oxygen-vacancy theory, this is reasonable as devices are exposed to a total pressure slightly higher to the one found in air. Therefore, oxygen movement should not occur at a faster rate as one would expect to have been the case in the previous two subsections. Contacting devices was not easy within the chamber and this, along with the cost of leaving a $^{16}O_2$ constantly flowing, led to few devices being tested.

4.2.1.4 *$^{14}N_2$-rich atmosphere*

Electroforming of devices in $^{14}N_2$-rich atmosphere appeared to be unexplainably impossible using previously used voltages and current compliance settings on 6 devices. It was necessary to do so in air before injecting the $^{14}N_2$ inert

gas and attempting to reset. This result does not reconcile with any of the previous ones involving modified atmosphere. Electroforming in high vacuum shows that this process does not require atmospheric oxygen. Similarly, electroforming in $^{16}O_2$-rich atmosphere shows that devices can electroform with a total pressure above the one found in air. This also leads to the conclusion that while a steady flow of gas may have acted as heat carrier and cooled down the device, formation hindering due to reduced temperature is to be ruled out.

4.2.2 Resetting

Figure 43b shows resetting IV curves in different atmospheres. Resetting was possible in all atmosphere except the $^{14}N_2$ one. Here devices would toggle between states but never achieve stable reset; reset was therefore unsuccessful.

Figure 44 compares the different reset voltages obtained. Reset in air results were shown using a distribution curve in section 4.1.1. These results should not be directly compared as filaments were electroformed in different atmospheres and with different current compliances; filament quality may well affect reset voltages. Figure 44 should shows a large spread of voltages with no obvious patterns other than low voltage resets in $^{18}O_2$-rich atmosphere.

4.2.2.1 High Vacuum

Resetting during high vacuum operation is shown in the pseudo IV curves found in Figure 43a (1.25 V with 0.5 mA current compliance). To reset, it was necessary to have successfully electroformed a device. As explained, this was not easy and it was only possible to attempt reset on 6 devices. Out of these devices 4 reset. According to former UCL student Leon Garnet, who obtained reset voltage statistics in air using sweeps, reset failure has also been observed in air. It is not

surprising that the reset rate is quite high on devices that had previously electroformed in vacuum. Successful vacuum electroforming means that devices did not release all oxygen and did not enter an unrecoverable state. These are devices that may well still have oxygen available for reset. Additionally, resetting does not rely on current compliance and therefore slow current compliance circuitry should not affect the process outcome as it was in the electroforming case.

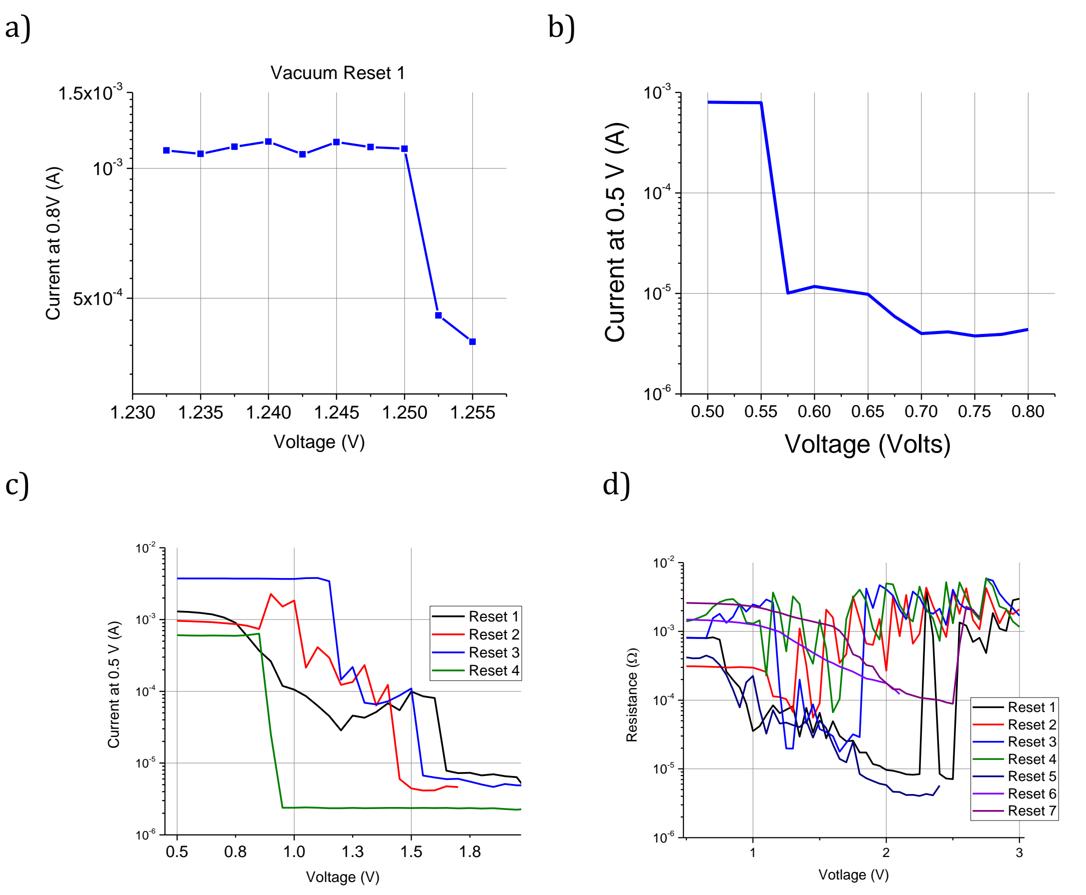


Figure 43: IV curves showing reset events in different atmospheres: a) high vacuum (1.25 V), b) $^{18}O_2$-rich atmosphere (0.55 V), c) $^{16}O_2$-rich atmosphere (between 0.95 V and 1.75 V), d) $^{14}N_2$-rich atmosphere (all unsuccessful).

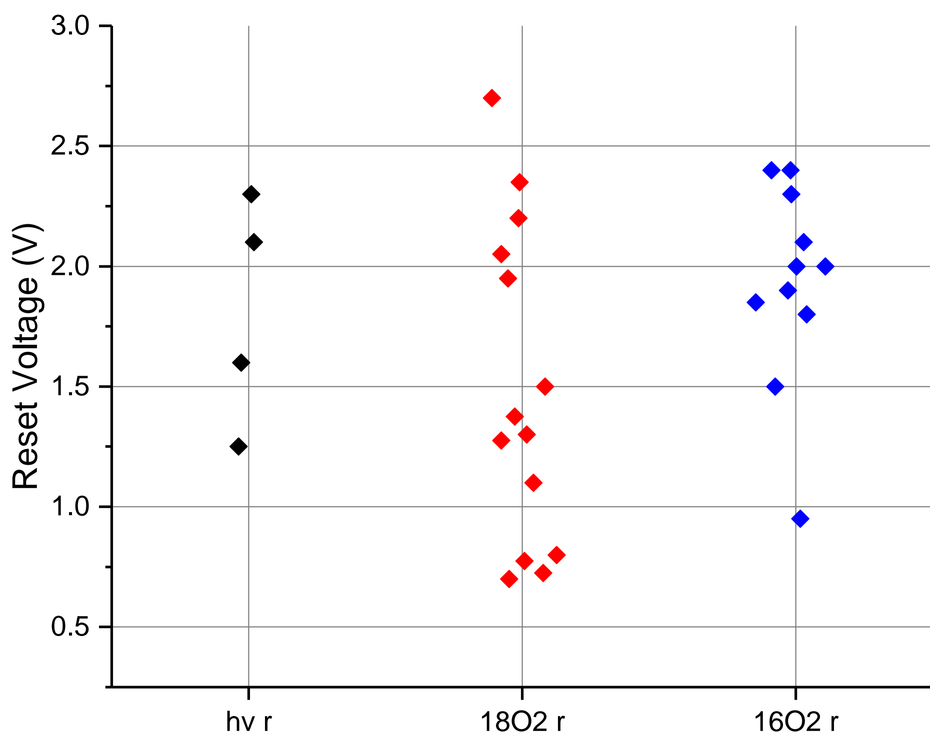


Figure 44: Column scatter of reset voltages. No clear patterns emerge other than low voltage resets in $^{18}O_2$-rich atmosphere.

Once again, it does not make sense to attempt computing voltage statistics due to a very small number of devices. Additionally, reset voltages also depend on filament electroformation process and strength. Therefore, a filament electroformed in vacuum may be very different from one electroformed in air leading to different voltage which cannot be directly compared to the one of a filament electroformed in air. It would be interesting to attempt high vacuum reset on devices previously electroformed in air.

4.2.2.2 $^{18}O_2$-rich atmosphere

Resetting devices was easy to achieve in $^{18}O_2$-rich atmosphere with sweeps and saw few failures during device cycling. Successful resetting is shown in Figure 43b which shows that resetting can occurs at less than 1 V 4 out of 14 times on devices electroformed at 1 and 3 mA (resetting in-air never occurred below ~1.2 V given 3 mA electroforming 4.1.1). One could conclude that voltages much lower than

in-air were sufficient to reset devices. While the small number of devices, once again, does not allow for the computation of statistical conclusions, it could well be that a greater oxygen partial pressure than in-air allowed such reduction in voltage. According to the previously introduced oxygen vacancy model, oxygen species are needed to render the top part of a filament insulating when resetting. It could well be that species present in the device atmosphere are used during this process in addition to those that get stored near the top part of the filament. A different interpretation, however, would see resetting occurring at lower voltage due to filaments that electroformed at lower voltage than in air. These filaments would therefore be weaker and require a lower reset voltage. It would be interesting to repeat this experiment on filaments electroformed in air to see which theory applies. Resetting was accomplished using pseudo-sweeps 13 times. On top of this, resetting was also carried out in cycling tests using 2.7 V long pulses without compliance and generated with a Keithley 6430.

4.2.2.3 $^{16}O_2$-rich atmosphere

$^{16}O_2$-rich atmosphere resetting results are shown in Figure 43c. Switching in an atmosphere with higher-than-air oxygen partial pressure led to successful reset 11 times at voltages that do not stand out from the statistics obtained in air. This is consistent with the previous described oxygen-vacancy theory which explains that filaments need to reacquire oxygen species from their surroundings to reset. The normality of the reset behaviour could also be found in the fact that during electroforming pressure was very close to 1 atmosphere, devices were still sealed to the outside world while pristine, and electroforming was reliable with voltages often seen in air.

4.2.2.4 $^{14}N_2$-rich atmosphere

Resetting on devices previously electroformed in air with 1 mA current compliance was attempted in $^{14}N_2$-rich atmosphere. However, as shown in Figure 43d, all six attempts led to temporary, partial ($\log|\Delta R| < 1$) or fully unsuccessful resets; proper resets could not be achieved. Devices stayed in a LRS. This shows that devices normally electroformed in air may need atmospheric oxygen to successfully reset. While this goes against the reset results obtained in high vacuum, it must be noted that reset in high vacuum was only attempted on selected devices that still had some embedded oxygen available through the use of low current compliance.

4.2.3 Setting

Figure 45 shows that setting is possible in high vacuum, $^{18}O_2$-rich atmosphere and in $^{16}O_2$-rich atmosphere. Setting in $^{14}N_2$ was not attempted as resetting did not work properly. In the case of $^{18}O_2$ pseudo IV-sweeps, toggling between states occurs before the setting process can be deemed completed. On some occasions toggling does not settle with increasing voltage.

Figure 46 shows set voltages in different atmospheres. Data is spread widely and does not show obvious patterns. Despite this, these values should not be directly compared as electroforming of the different devices occurred under different conditions in terms of atmosphere and current compliance.

4.2.3.1 High Vacuum

Those few devices that reset led to attempting high vacuum setting. Some results are shown in Figure 45a, which shows setting being possible at around 1.3 V with 0.5 mA current compliance. Here, like with electroforming, precautions had to be taken to avoid hard break down. The same reasons explained when talking about

electroforming apply. Overall, successful setting was observed four times. Such a small dataset is not enough to draw any statistical conclusions. These values, as it was in the case of the reset voltages, should not be directly compared to those obtained in air as a filament electroformed in high vacuum may lead to different reset and set voltages. It would be interesting to attempt high vacuum set on devices that were previously electroformed and reset in air. Nonetheless, it has been shown that high vacuum setting is possible.

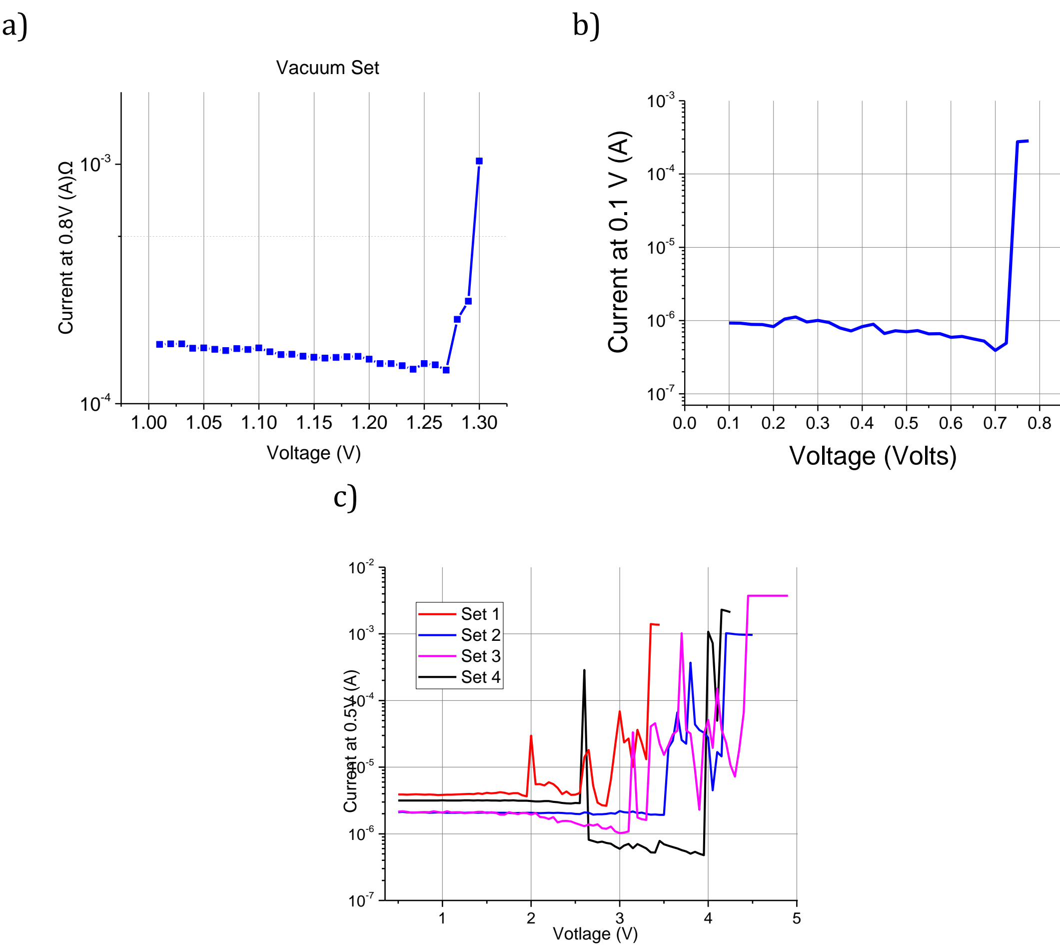


Figure 45: IV curves of set events that occurred in different atmospheres: a) high vacuum (1.3 V 0.5 mA current compliance), b) $^{18}O_2$-rich atmosphere(0.75 V, 1 mA), c) $^{16}O_2$-rich atmosphere (see section 4.2.3.3).

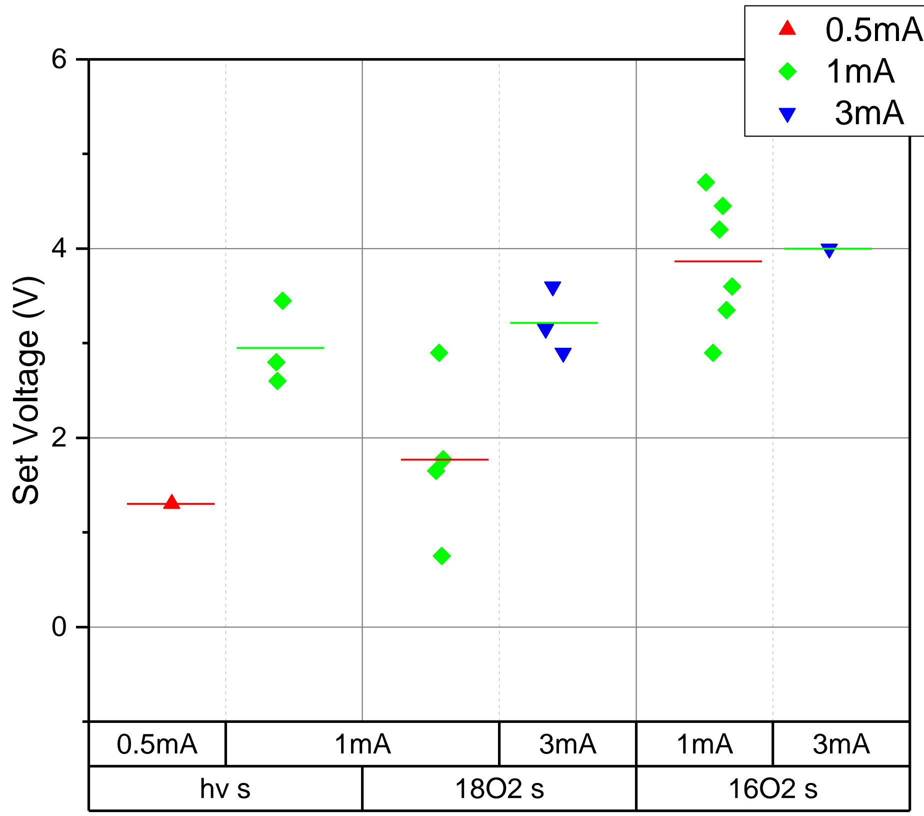


Figure 46: Column scatter plot of set voltages. Data is widespread and does not indicate any trend.

4.2.3.2 $^{18}O_2$-rich atmosphere

Setting of devices in $^{18}O_2$-rich atmosphere was also easy to accomplish 11 times without failures. As shown in Figure 45b, this could be done with voltages as low as ~0.75 V. Just like with electroforming, this process was observed to occur at voltages much lower than in-air (where setting needs about 3.6 V at 3 mA to have a 50% chance of success, Figure 24). Once again, few devices were used and meaningful statistics may not be calculated but this could be an indication of a similar situation as the one explained in the paragraph talking about electroforming. However, also in this case, it could be that reduced voltages are needed to set because filaments were electroformed at lower voltages than in air.

Looking at Figure 47 one can see an unstable set; this phenomenon was observed in two devices. In these cases, during set voltage sweeps, devices would set

but then quickly reset. This would happen repeatedly leading to oscillatory behaviour. This phenomenon was never observed in air to this extent. It may well be due a high oxygen partial pressure, which leads to device reset even under low current conditions. However, it could also be that a device whose filament was electroformed in $^{18}O_2$-rich atmosphere would display this behaviour while being set in air. Filament strength/width could be the root cause of such oscillations.

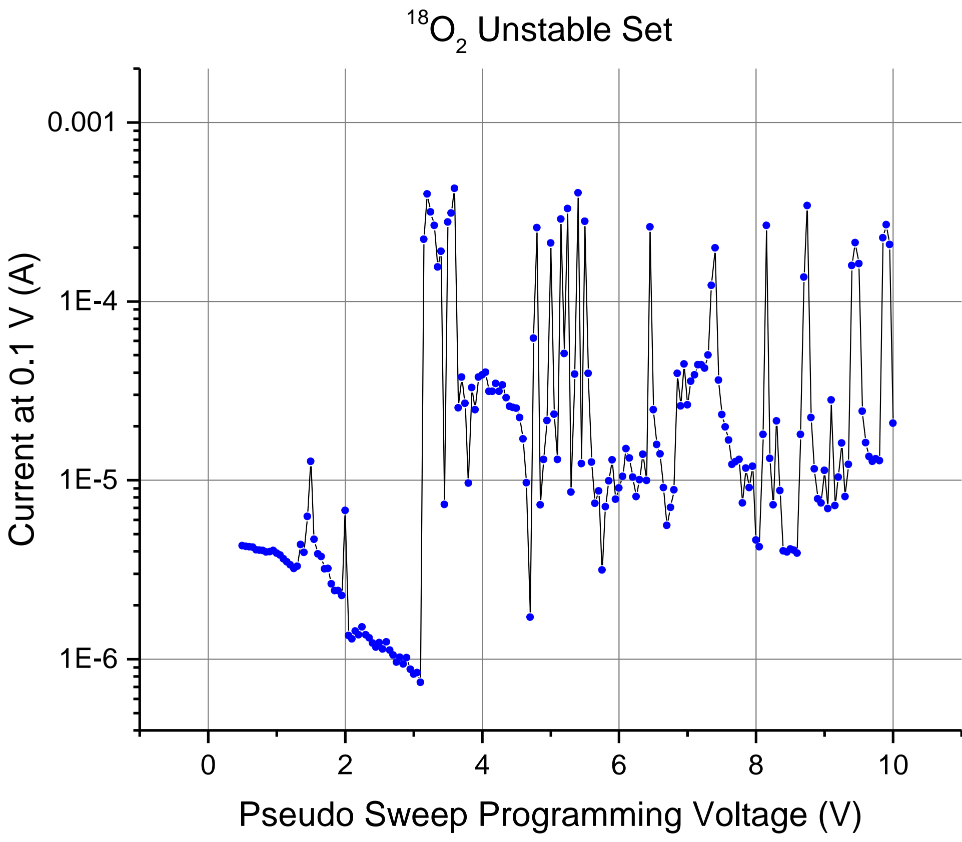


Figure 47: Unstable set in $^{18}O_2$ atmosphere. This may be due to a high oxygen partial pressure which leads to device reset even under low current.

4.2.3.3 $^{16}O_2$-rich atmosphere

Setting devices in $^{16}O_2$-rich atmosphere appears to be quite erratic in nature, as seen in Figure 45c; as shown, some devices toggled between states up to 4 times before entering a stable state. Devices would set, quickly reset and eventually set. The final setting voltage is between 3.5 and 4.7 V. Despite this voltage range being similar to the one seen in-air, the erratic behaviour is different from the one

generally observed in air and could be explained through the higher availability of oxygen molecules compared to air. This might well reset filaments following weak resets occurring during the initial reset phase. In other words, the presence of competing phenomena during set is enhanced by high oxygen availability compared to in-air experiments. The reason why this does not happen during electroforming could be that the extra oxygen available only manages to reach the first few nanometres of the active layer (a reset filament only increases in resistance in its very top part). Similarly, during electroforming, oxygen species moving towards the anode might have left pathways to the filament which atmospheric oxygen can use. One could also argue that set voltages are different from in-air experiments because the electroforming stage did not take place in air. However, this is unlikely as electroforming total pressure was very close to 1 atmosphere, devices were still sealed to the outside world while pristine, and electroforming was reliable with voltages often seen in air.

4.2.3.4 $^{14}N_2$-rich atmosphere

Setting in $^{14}N_2$-rich atmosphere was not possible due to erratic resetting as shown in section 4.2.2.4. It would be interesting to electroform and reset devices in air and therefore study whether $^{14}N_2$-rich atmosphere setting is possible.

4.2.4 Conclusions

Only a few conclusions can be drawn from results shown in this section. The main reason for this is that devices were formed in different atmospheres leading to different filaments and the impossibility of directly comparing reset and set between different atmospheres. Nonetheless, what follows has been shown on unipolar devices while operating them unipolarly:

- It is easier to melt down devices while attempting electroforming in high vacuum than it is in air.
- Electroforming in N_2-rich atmosphere is not possible.
- Resetting in N_2-rich atmosphere leads to erratic behaviour.
- $^{16}O_2$-rich atmosphere leads to erratic set (assuming that electroforming in this atmosphere is comparable to electroforming in air).

The above statements lead to the following conclusions:

- A certain total atmospheric pressure is needed for devices to electroform slowly enough not to meltdown.
- Resetting devices formed in air appears to require atmospheric oxygen.
- Setting is not favoured in atmospheres that have an oxygen partial pressure higher than air.

All above statements reconcile with oxygen vacancy theory and add to it that following electroforming these devices not only rely on their own oxygen content but also on oxygen present in their surrounding atmosphere.

5 DETECTION OF EMITTED OXYGEN

5.1 OBSERVATIONS ATTRIBUTABLE TO OXYGEN MOVEMENT

During device cycling, deformed regions developed on the surface of the top contact of the device under test (DUT). These optically visible regions appeared mainly when cycling using pulses with compliance but also when cycling through sweeps. Similar phenomena are thoroughly discussed in the literature [5, 25, 28, 60, 76, 82] and generally attributed to oxygen gas escape. Here, a plausible explanation is that negatively charged O^{2-} interstitial ions [102] diffuse towards the TiN anode and build up. Because of Coulomb repulsion these ions apply a large amount of force against this electrode, which deforms into a bubble-like shape. This deformation looks like a dark stain under the optical microscope. Once physical tension overcomes the physical strength of the TiN thin-film electrode, the ions are released as O_2^- and as neutralised $^{16}O_2$ oxygen gas molecules. The fact that the anode is positively charged when biased should help the neutralisation process.

The surface area of the deformed regions does not increase by a regular amount at each switching cycle. Rather, these features seem to develop and expand abruptly at random times. This, as mentioned above, could be justified by the electrode requiring a certain pressure before it will give in and let oxygen ions and gas escape. Therefore, pressurized pockets might explain the way these deformed regions develop. This oxygen-pocket model is supported by a different behaviour while sweeping with compliance. In this case, the phenomenon does not appear to be as significant. It appears reasonable that while sweeping, chemical changes occur gradually and the produced oxygen has time to find space to slowly diffuse into the SiO_x and TiN. Instead, when switching occurs by pulsing, many ions are generated at

once and find electrical acceleration towards the anode. These regions, as previously discussed in section 4.1.2, are insulating and therefore hinder device endurance. Deformation was observed to always originate on the edge of a device and expand towards the probing needle. This could be because some neutralised ions act as gas and travel along a least resistance path. The electrode edge should allow for easier gas escape than the middle of the same electrode.

Although beyond the scope of this research, Atomic Force Microscopy (AFM) was carried out by team member Mark Buckwell to better understand the topography of the above discussed deformed regions. Report will be on some of the results obtained on devices previously biased in this work. Mark Buckwell's publications may be referred to for further information on the technique [56]. Figure 48a shows an optical image of a deformed region following device cycling. Figure 48b, instead, shows the topography of the same deformed region as produced by a Bruker Icon AFM. Figure 48c helps understand heights in Figure 48b.

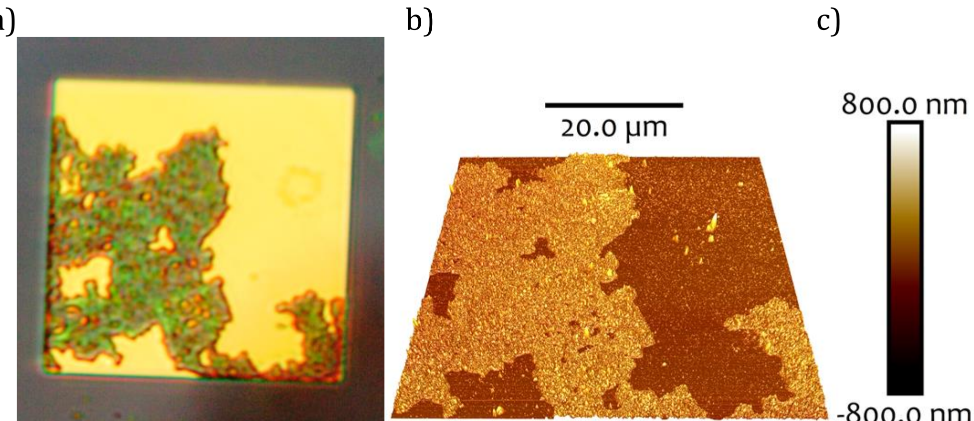
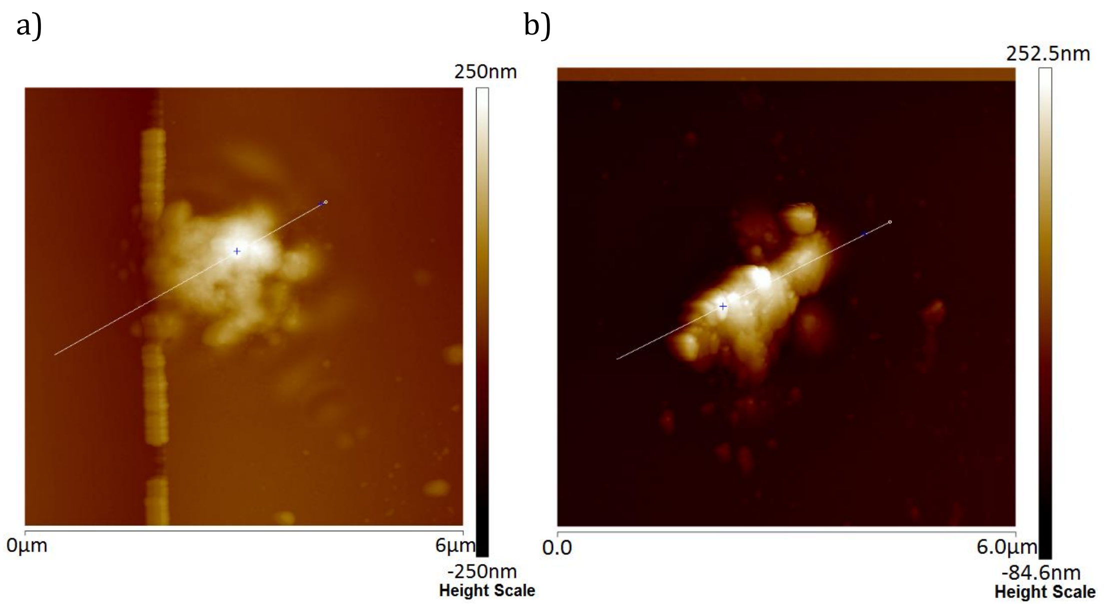


Figure 48: Surface deformations: a) optical microscope image of a deformed region appearing while cycling a TiN/SiOx/TiN device. b) AFM rendering of the same region[3]. c) height scale for the image found in b). b) and c) provided by Mark Buckwell.

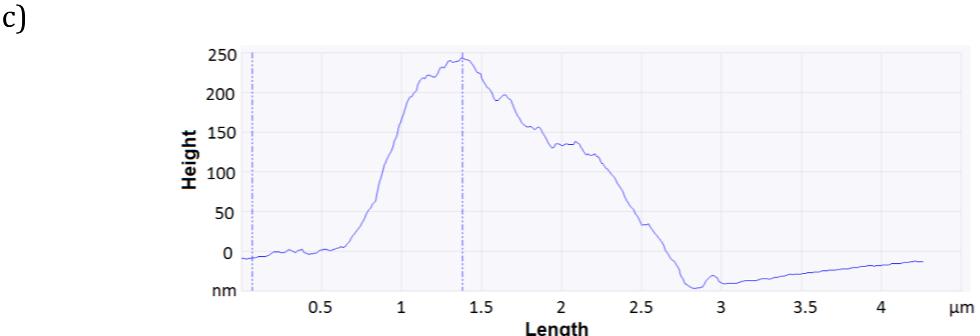
Figure 49a looks more in detail at deformed regions. Here the results of a single electroform carried out with a 15 V pulse with 3 mA current compliance. Reading resistance at 0.5 V, it was possible to see that the device went from 9.3×10^9 Ω to 1.06×10^3 Ω. Figure 49b shows the topography of a different device electroformed and read identically that went from 5.92×10^9 Ω to 1.99×10^2 Ω. Figure 49c and Figure 49d, are line scans through the AFM topography maps. Here, in both cases, that the height of the deformed region is about 250 nanometres and the width about 3 micrometres.

Figure 50a shows a deformed region created through 4 cycles carried out using 15 V and 3 mA current compliance to form, 2.7 V and 100 mA current compliance to reset and 7 V and 3 mA to set. This led to a region of ~3.5 micrometres with a height of 500 nanometres (Figure 50c). Figure 50b shows the deformed region created by -15 V (negative) form with 3 mA current compliance (resistance at 0.5 V went from 6.24×10^9Ω to 9.81×10^4Ω). This type of electroform could not be reset despite the expected symmetry of the TiN/SiO_x/TiN family of

devices. The pulse led to a region of interest with a width of 4 micrometres and a high of ~400 nanometres (Figure 50d).



c)



d)

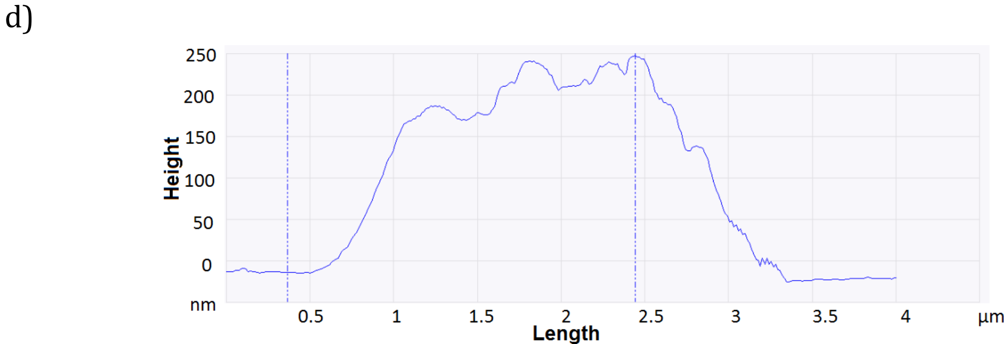


Figure 49: a) AFM image of surface deformations created on TiN contact 15V, 3mA form, resistance read at 0.5V from $9.3\times10^{9}\Omega$ *to* $1.06\times10^{3}\Omega$*; deformation appeared on contact edge . c) profile of superimposed line on previous image. b) AFM image of deformed region created on TiN contact 15V, 3mA form, resistance read at 0.5V from* $5.92\times10^{9}\Omega$ *to* $1.99\times10^{2}\Omega$*; deformation appeared in middle of contact. d) profile of superimposed line on previous image. Images obtained by Mark Buckwell.*

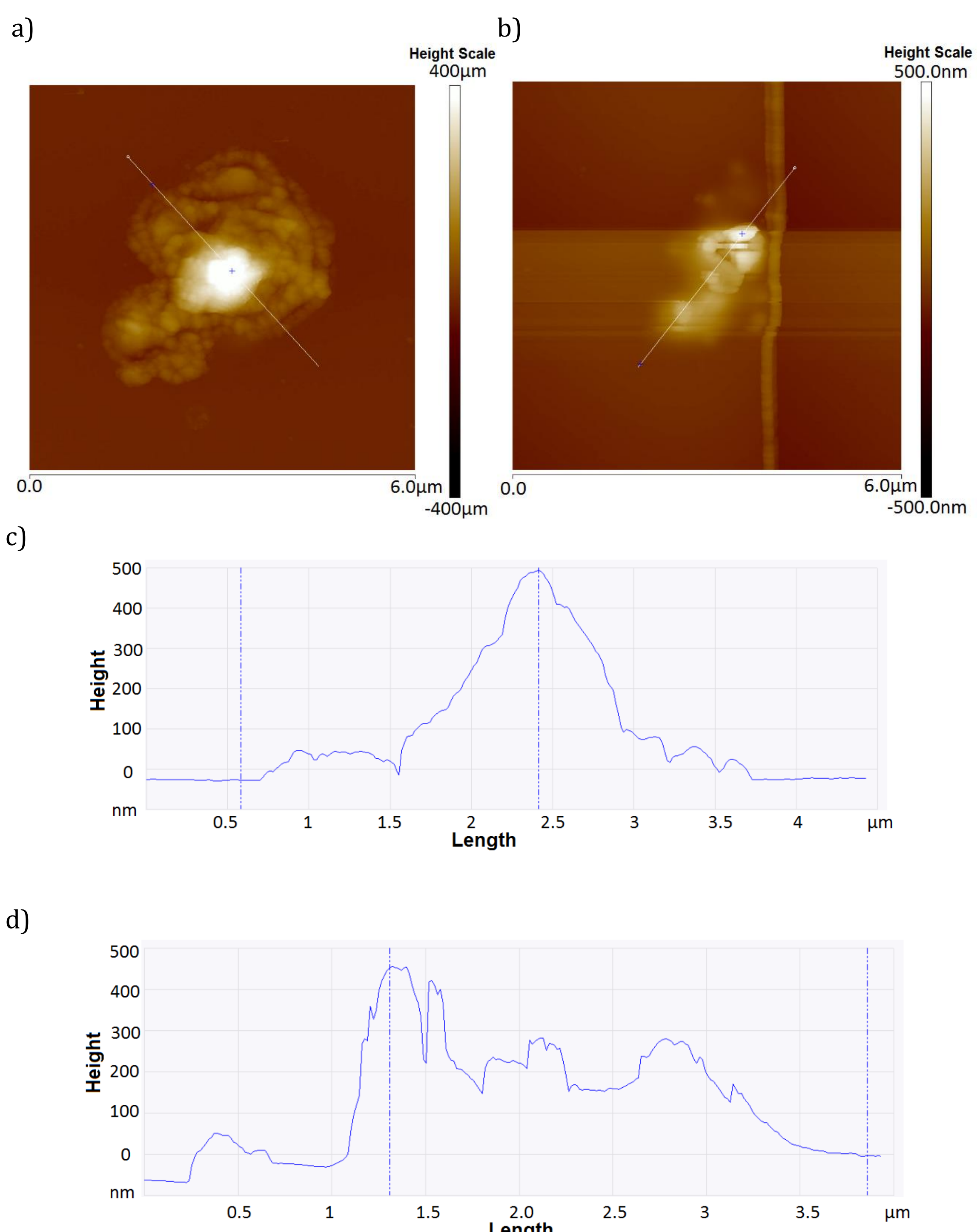


Figure 50: Surface deformations: a) AFM image of deformed region created on TiN contact. Form: 15V, 3mA, reset: 2.7V, 100mA, set: 7V, 3mA, cycled 4 times. Deformation in middle of contact. c) profile of superimposed line on previous image. b) AFM image of deformed region created on TiN contact -15V, 3mA form, resistance read at 0.5V from $6.24 \times 10^{9}\Omega$ to $9.81 \times 10^{4}\Omega$; deformation appeared on contact edge. d) profile of superimposed line on previous image. Images obtained by Mark Buckwell.

Figure 51a shows the AFM image of the deformed region created directly on the SiO_x surface of a $TiN/SiO_x/TiN$ device using a tungsten needle. The device

electroformed using a 20 V pulse with 1 mA current compliance. Resistance went from 4.79×10^{10} Ω to 1.586×10^{2} Ω. It was not possible to reset this device. Figure 51c shows the height profile of the white line found in Figure 51a. Here one observes a width of about 4 micrometres and a height of about 450 nanometres. Figure 51b shows the region SiO_x created by a -20 V (negative) pulse with current compliance of 1 mA. Here the region resistance dropped to $9.66\cdot 10^{3}$ Ω without possible resetting. Figure 51d shows the height profile of the white line found in Figure 51b. One observes 3.7 micrometres of width and a height of about 200 nanometres.

In conclusion, $TiN/SiO_x/TiN$ devices deformed regions attributable to oxygen movement develop through positive and negative biasing and on and off TiN electrode. This shows that the direction of the electric field is not important and, following internal generation of charged ions, one probably has the release of uncharged oxygen species. Similarly, it was possible to learn that UCL $TiN/SiO_x/TiN$ devices are not fully symmetric as devices formed with a negative bias are impossible to reset. This is probably because devices are fabricated from bottom to top and ion implantation must occur. This means that as fabrication occurs each layer will end up containing a certain number of ions implanted out of the deposition of the next layer. For instance, the top part of the bottom electrode will contain some active layer contamination. The top part of the active layer will contain some top electrode contaminants. Devices are therefore not symmetrical.

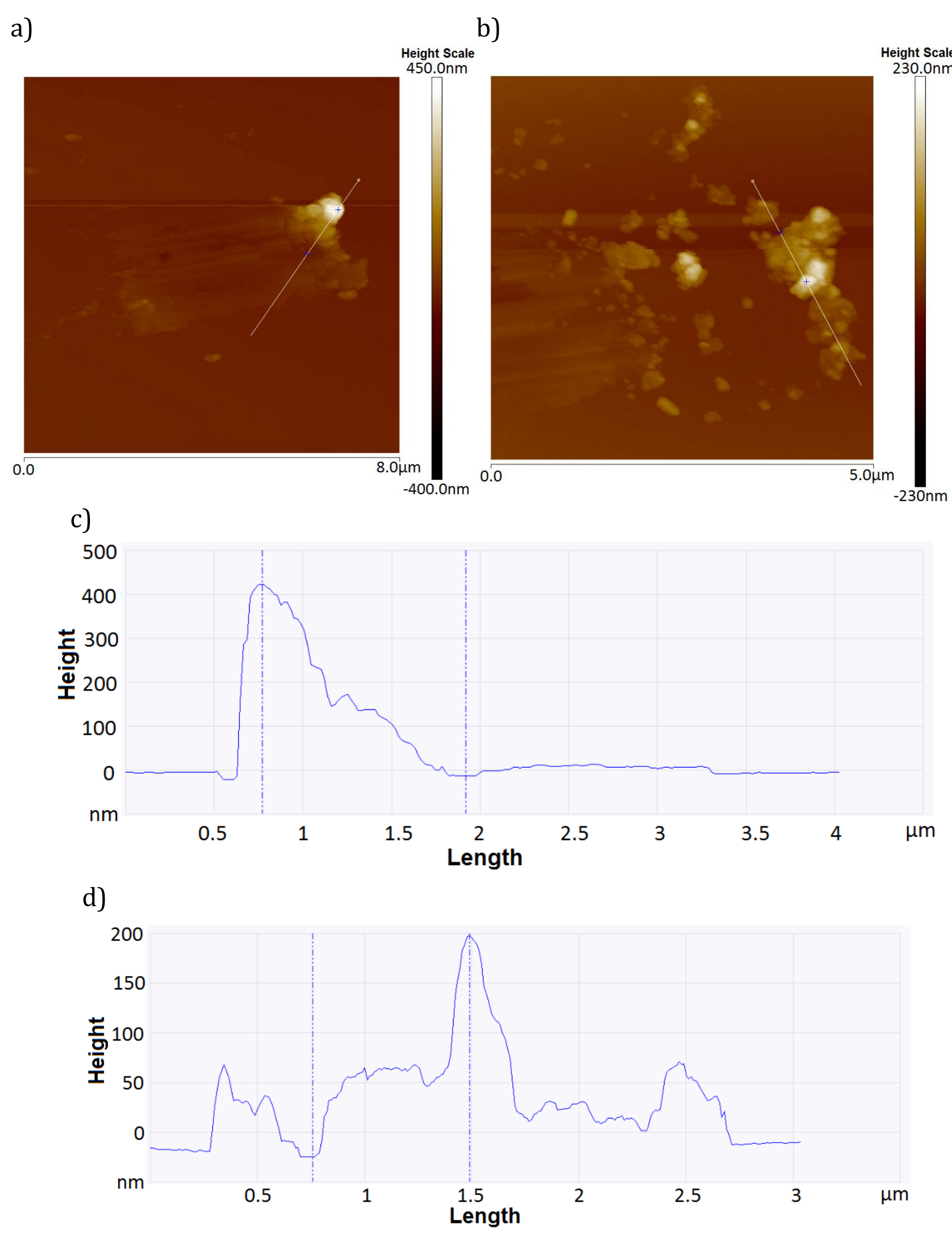


Figure 51: Surface deformations: a) AFM image of deformed region created on SiO_x using a W needle. Form: 20V, 1mA. Resistance dropped from $4.79\times10^{10}\,\Omega$ to $1.586\times10^{2}\,\Omega$. c) profile of superimposed line on previous image. b) AFM image of deformed region created on SiO_x -20V, 1mA form. Resistance dropped to $9.66\times10^{3}\,\Omega$ d) profile of superimposed line on previous image. Images obtained by Mark Buckwell.

5.2 Oxygen Gas Detection

While one of the devices was being electrically swept in high vacuum, a temporary increase in mass-to-charge-ratios 20 and 32 was observed on the Hiden RGA output. This happened while attempting to electroform a pristine device. The mass-to-charge-ratio 32 spike probably corresponds to oxygen gas as no other elements used in the fabrication process, or present in the vacuum background noise, would form molecules of this mass. The mass-to-charge-ratio 20 spike could correspond to an increase in CO gas, given that organic species are present in the chamber. While it might seem appealing to think of these two detections as direct from sample to RGA, the story might be quite more complex. Oxygen gas molecules or superoxide ion emitted into the chamber at molecular regime vacuum would travel in a straight line and quickly hit surfaces such as the one of the chamber and not elastically bounce. This aluminium surface has not been cleaned or baked (a cleaning process involving high temperature) for over 15 years and is likely to be coated in reactive organic compounds. These, as well as aluminium itself, are likely to adsorb the emitted species and therefore slowly desorb them either changed or unchanged as explained in Ref. [120, 121]. It is not unlikely that CO gas is produced by this process. It is therefore most likely that the observed spikes are due to relatively slow desorption rather than direct detection. Nevertheless, it is also appealing to think that the hot ionising RGA filament allows for gaseous organic compound to react with oxygen molecules or superoxide to generate CO gas.

Figure 52 shows the evolution of different species over six subsequent RGA scanning cycles. Each cycle corresponds to an elemental survey lasting about 5.7 seconds. The first shown cycle is “Cycle 1”, which represents background noise

levels. The peak present at around mass-to-charge-ratios 29 (CHO) and 31 (CH_3O) probably represent IPA residue from cleaning. Maximum detection of mass-to-charge-ratio 32 and 28 molecules occurred during “Cycle 2”; the number of counts is here almost three and two times more than in the previous cycle for mass-to-charge-ratios 32 and 28, respectively. “Cycle 3” shows a reduction in mass-to-charge-ratio 28 and 32 counts. By “Cycle 4” counts are back at background levels for the two mass-to-charge-ratios. Thus, a mass 32 gas, probably oxygen, was emitted as a burst during “Cycle 2”, adsorbed by the vacuum chamber walls and desorbed over a few tens of seconds. The desorption stage also involved the release of a mass 28 gas, probably CO that formed from the reaction of oxygen gas and oxygen oxide with the organic contamination present in the chamber. The following cycles, 4, 5 and 6 show that counts are back to background levels.

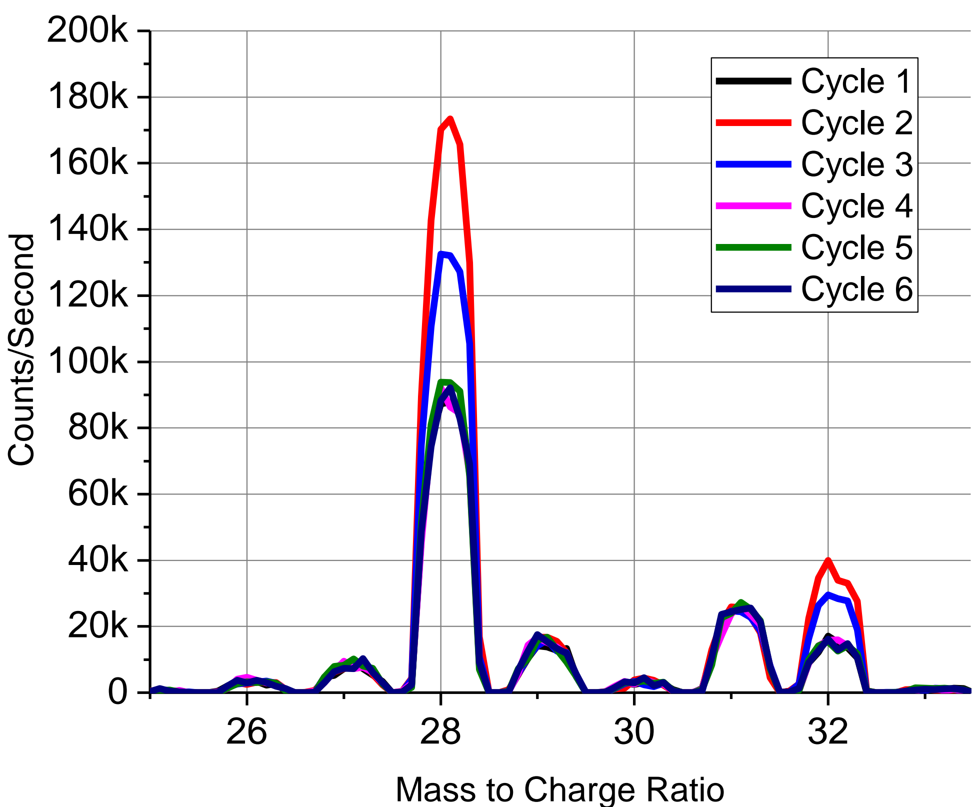


Figure 52: Detection of species between mass 24 and 34 over six sampling cycles. One can observes a transient increase in mass-to-charge-ratios- 28 and 32 indicating an increase in CO and $^{16}O_2$ levels.

Mass-to-charge-ratio 32 detection leads to evidence of oxygen species emission by the sample due to electrical stimulation. Therefore, it was decided to focus on this ratio while increasing the sampling speed of the RGA. The result of this is the curve found in Figure 53, which, once again, suggests that a mass 32 gas emission occurred within the chamber. Also in this case, the number of counts with respect to background noise quickly increased and went back to background over several seconds. Gas desorption from materials is often (but not always) linear when plotted on logarithmic graphs [120]. Therefore, it makes sense to attempt an exponential fit of the type $y = y_0 + Ae^{(-\mathrm{t}-\mathrm{t}_0)/\tau}$, which leads to a time constant $\tau =$ 18.04 seconds. This gives an idea of the relatively quick chamber wall desorption rate. This particular emission occurred at 6.6 V during the forming sweep of a device, which never electrically formed. Tests were only carried out in the positive polarity.

Detecting what appears to oxygen was not an easy task. Although repeatable, this experiment was not very controllable. It was observed that detection only occurs when attempting to electroform or set and not while attempting reset. This reconciles well with the oxygen vacancy theory previously discussed as oxygen ions would migrate toward the anode (or top electrode) only when a filament is being electroformed or set. Similarly, sporadic detection reconciles with the way in which optically-visible deformed regions seem to develop. They do not expand at every switching cycle but rather only abruptly every few cycles. This suggests that oxygen builds up in pockets and, once pressure is sufficient, manages to burst out through the TiN electrode.

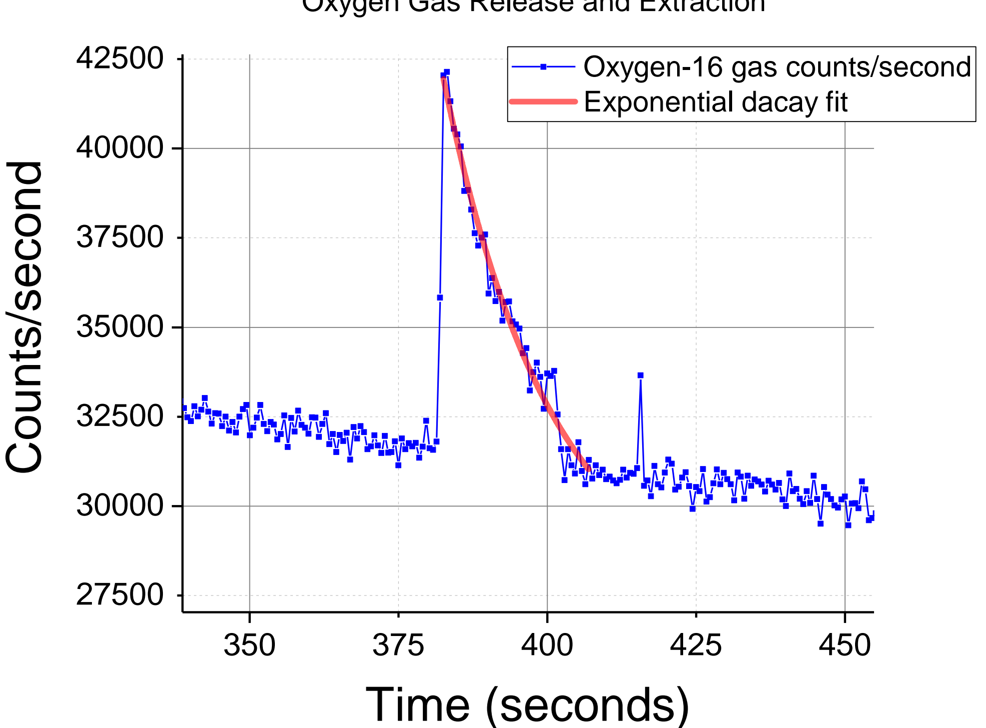


Figure 53: Oxygen gas detection over time. A peak appears at around 385 seconds and disappears into background noise after about 20 seconds. An exponential fit shows a $\tau =$ 18.04 *seconds.*

Lack of literature, other than UCL's [58], involving the detection of gases during the biasing of thin film oxides means that the above results may not be directly compared to similar experiments. Nonetheless, literature supporting the generation of interstitial O^{2-} ions from a theoretical viewpoint does exist [102] as previously discussed in sections 2.4.4 and 5.1. It is stated that amorphous SiO_2 contains wide O-Si-O bonds which can accommodate up to two electrons. This process weakens the Si-O bonds leading to the production of O^{2-} ions.

5.3 Ion Emission Detection

5.3.1 Channeltron Ion Detection without Mass Identification

It was therefore tested whether oxygen is emitted as purely neutral species. The previous results were inconclusive when it comes to this as either charged or

uncharged oxygen may be adsorbed and desorbed by aluminium and/or its layer of organic contaminants.

Using a positively biased channletron detector along with a blanked ion beam, it was possible to convert the detection of negatively charged particles into white streaks on the imaging screen of the FIB microscope. This stems out of the way negative ion/electron images are normally plotted on the screen. The channeltron has a wide aperture, especially compared to the Hiden SIMS/RGA, which facilitates the entry of particles. Streaks often appeared (Figure 54) while trying to form or set a device. The same was not true while reset was attempted. It appears that the form and set processes lead to the release of some negatively charged ionic species.

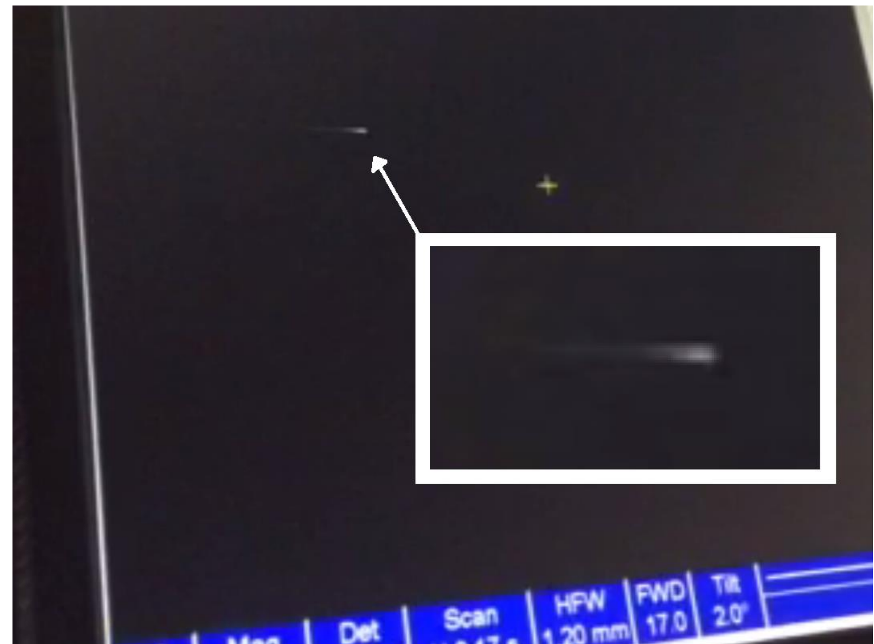


Figure 54: Streak seen on the focused ion beam microscope while forming or setting a device. The channeltron detector is active as if it were imaging but the FIB beam was left blanked to avoid the unwanted images. These streaks represent the short-lived (microsecond-domain) detection of negative particles in the vacuum chamber.

A full screen is drawn every 2.17 seconds and is made of 884 lines. Each line contains 1024 pixel. This means that each pixel represents 2.4 microseconds.

Looking at the image file shown in Figure 54, the width of the image area spans about 1000 pixels. Therefore, each image pixel represents about 2.5 microseconds. Values are approximate as camera titling with respect to the screen was not taken into consideration.

The above can be used to produce a time vs. intensity plot of the streak present in Figure 54. This is shown in Figure 55. Here, the most remarkable feature is a fast decay time. It takes the streak about 25 microseconds to decay back to background level. Using an exponential decay fit, it is possible to see that $\tau = 1.12\times10^{-5}\ seconds$. This time constant is significantly shorter than the one obtained with O_2 and CO detection. This is most probably because emissions are directly detected and no longer rely on chamber wall adsorption and desorption. It must be noted, however, that several delays have not been included in calculations. These are: screen pixel time persistence, camera sensor time persistence, as well as possible delays with the channeltron analog and digital data processing. All these delays may lead to additional pixel persistence, which is not an unwanted feature when the equipment is used for its native purpose of imaging. Therefore, the calculated times above should only be considered as worst cases.

From the above, it is reasonable to conclude that charged species leave samples during electroforming and setting but not resetting. This reconciles with the discussed oxygen vacancy model which relies on an electric field accelerating oxygen ions towards the anode during the former two biasing stages. Similarly, this suggests that oxygen may be released in ionic form.

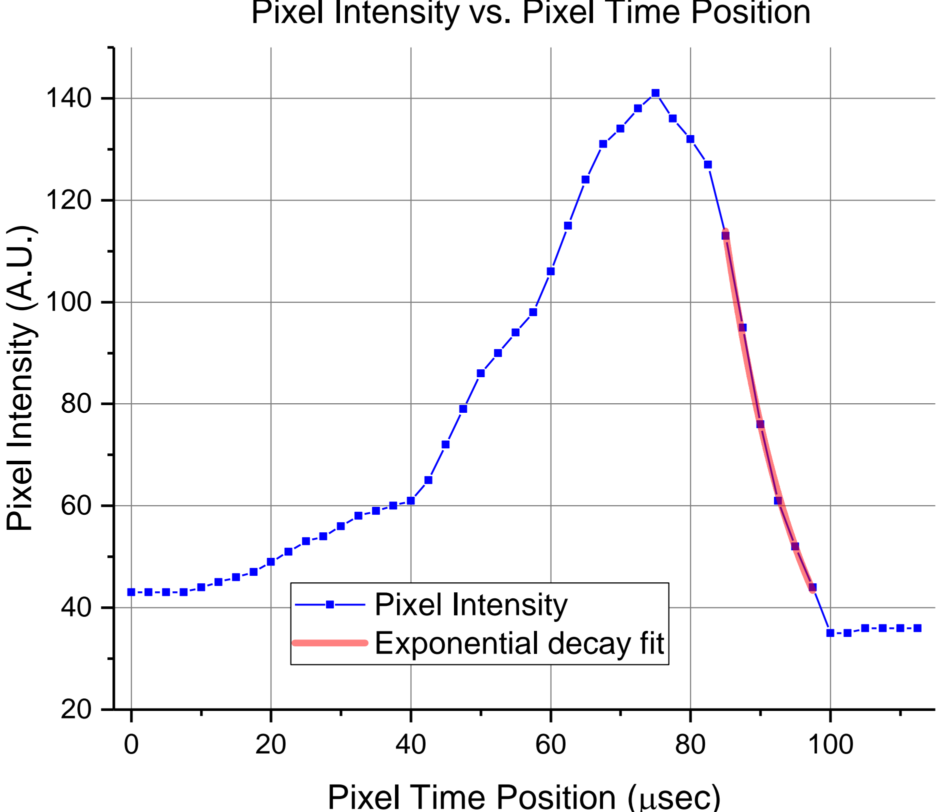


Figure 55: Pixel intensity vs. time of a white streak found in Figure 54. This plot indirectly represents the presence of negatively charged particles in the vacuum chamber during forming and setting of devices.

Once again, a lack of literature, aside from UCL's [58], is present in the field of thin film oxide emission during biasing. It is therefore hard to validate results against those of other groups. It is however reasonable to envision the emission of oxygen ions from devices because of the processes described in Ref. [102] and previously discussed in sections 2.4.4 and 5.2. The fact that pseudo sweeps leave the anode floating between pulses helps explain why oxygen species might not become neutral before leaving the sample. An anode forced at zero volts would increase the likelihood of charged species discharging.

5.3.2 SIMS Ion Detection with Mass Identification

The vacuum chamber also mounted a SIMS detector with an aperture much larger than the one of the Hiden SIMS/RGA. This device is generally used, along with the ion beam, to create SIMS depth profiles of specific negatively or positively

charged ions present in a sample surface. However, blanking the ion beam, allowed the SIMS detector to be used to detect negative ions emitted during device biasing. This approach is similar to the one used in section 5.3.1 but specific mass-to-charge-ratios could now be identified.

Given the previous detection of O_2 gas and unidentified negatively charged particles, it was chosen to attempt detection of O_2^- (superoxide) by monitoring mass-to-charge-ratio 32 while having the SIMS detector under positive bias. The SIMS detector was left sampling the ratio every 1 second for 1 second. At the same time, voltage was ramped up on several different DUTs to attempt electroforming. With some devices, mass-to-charge-ratio 32 counts suddenly increased during this process. In the case shown in Figure 56, this spike was 70 time above the highest background data point collected over 50 minutes. No correlation was found between this spike and electrical formation showing that oxygen movement starts during electroforming before a filament becomes conductive. Background data, as shown in Figure 56, was always limited to 50 minutes at a time due to software limitations. Nevertheless, significant spikes in SIMS counts were only ever detected during or following electrical biasing.

Single spikes detected using a 1 second sampling time, mean that the duration of the detection was one second or less. This means that charged superoxide ions are found in the chamber for at least 10 times less time than uncharged oxygen gas. This is consistent with the idea that highly reactive species are adsorbed by the vacuum chamber and slowly desorbed as uncharged gases but charged emissions are directly detected.

The emission of superoxide from oxide finds support in the literature. For instance Ref. [103] discusses the emission of such species from oxides stimulated either chemically or thermally. Similarly, in Ref. [105], it is explained that inhaled silica can react and emit superoxide, which leads to inflammation. Therefore, even though Ref. [102] (sections 2.4.4, 5.2, and 5.3.1.) talks about the internal generation of O^{2-} species during electroforming, it is possible that these reacted and turned into superoxide species. While no literature exists on the emission of superoxide from an electrically stimulate oxide, these results still do find partial validation in the literature: oxides have been observed to release superoxide when stimulated.

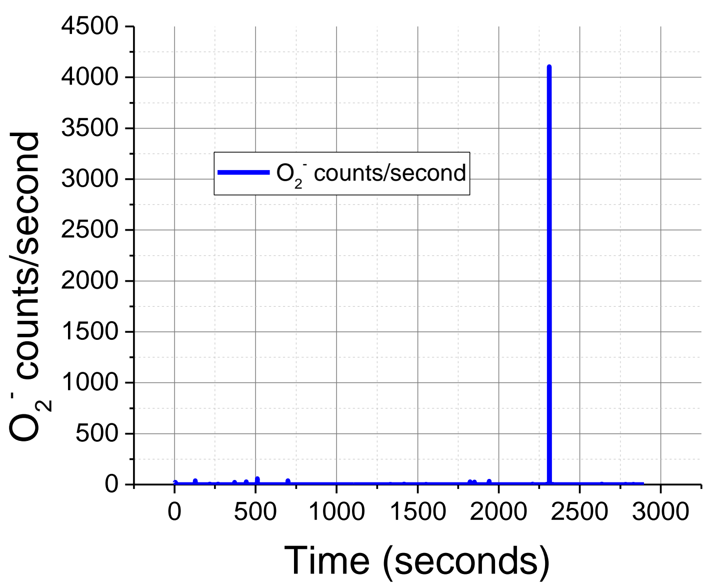


Figure 56: Detection of $^{16}O_2^-$ ions (superoxide) during device formation.

5.4 Conclusions

By combining the three above results stemming from SIMS in RGA mode, channeltron and the SIMS detector, it is reasonable to conclude that oxygen is emitted from devices while attempting to form or set and not while attempting reset. From what is explained in Ref. [102], it is reasonable that electron injection leads to the generation and emission of oxygen species. Similarly, in Ref. [103, 105]

one learns that oxides have been shown to emit superoxide under non-electrical stimulation. Some of these charged and highly reactive species must discharge before leaving the sample (and therefore exit the sample as O_2 gas) and some others must get adsorbed into the vacuum chamber walls. These latter ones will therefore be desorbed as uncharged O_2 gas or as CO gas following reaction with organic contaminants. It is also possible that species leaving the sample as O_2 gas get adsorbed and desorbed by the vacuum chamber being released unchanged as O_2 or as CO gas.

6 $^{18}O_2$ EXCHANGE

Cycling in 400 mbar atmosphere containing 50.5% $^{18}O_2$ allowed the observation of device behaviour with total pressure lower than air but oxygen partial pressure ($p(^{16}O_2)+p(^{18}O_2)+p(^{16+18}O_2)$) higher than in air as discussed in section 4.2. Following device cycling, it was possible to use SIMS and ToF-SIMS depth profiling to observe how atmospheric oxygen interacts with devices. Only those profiles under deformed regions showed embedding of ^{18}O well into the device bulk. The generated 3D tomography provides visualisation of where ^{18}O is found as a function depth. Further measurements using shallow angle SIMS analysis led to slightly different conclusions showing that ^{18}O does not reach into the bottom electrode. These further measurements reduced the probability of implanting species from one layer into the next but did not allow for a tomography to be created.

6.1 $^{18}O_2$-EXCHANGE RESULTS

6.1.1 Initial SIMS Studies

In initial studies, all devices cycled several times in an $^{18}O_2$-rich environment show incorporation of $^{18}O_2$ in the top electrode, active layer, bottom electrode and bulk. This process only appears to occur when analysing through deformed regions. No $^{18}O_2$ is found outside of these regions, showing that atmospheric oxygen only interacts in a localised manner.

The ^{18}O appears to be bonded to the silicon atoms present in the structure. If this were not the case, the non-bonded $^{18}O_2$ molecules would have escaped the sample by outgassing while the high vacuum was pumped. Similarly, the sample was

cleaned using standard acetone, isopropyl alcohol, DI water cycles and this did not remove ^{18}O away from the surface.

Thermally driven oxygen exchange is a process that in all known literature is carried out for tens of minutes, if not hours at temperature in the excess of 500˚C and often up to 1000 ˚C. This usually allows the oxygen to exchange to a depth well beyond the one of thin-films [112-116]. Device filaments are exposed to non-limited current only during reset. This is the occasion when the device might heat-up to temperatures in the order of those necessary for oxygen exchange to occur [65]. The question remains about the length of time the filament remains hot during each reset pulse as these pulses are on only last for around 10 milliseconds as shown in section 3.3.1.2. What must be noted is that following electroforming one no longer has a smooth and heterogenous structure as found in the oxygen exchange literature. Now, one has a structure with cracks and openings; otherwise, oxygen would have not found its way out of the active layer. Because of this, atmospheric oxygen directly reaches many spots within the active layer and top electrode. Local temperature will increase during reset allowing oxygen exchange to occur very quickly on those many surfaces exposed to air. Atmospheric oxygen exchange depth within these surfaces will not be great due to the short heating time but rather through the complexity of the structure. Therefore, despite a short heating time one may well find oxygen to have exchanged in depth and width around the filament location. While filament heating is key to resetting unipolar devices, it could be that some silicon dangling bonds oxidise with atmospheric oxygen even before resetting occurs. However, one would expect this process to become insignificant once heating occurs.

Figure 57 shows depth profiles carried out using the SIMS FIB perpendicularly over the sample. This showed a clear difference in fractional ^{18}O absorption over the deformed and unaffected regions. It is important to rely on the ^{18}O-to-total-oxygen ratio as different sample layers and states following cycling have different oxygen ion yields. Similarly, milling rates can change by the layer or post-cycling state and it is hard to state with certainty which layer is which on the figure in question. Nonetheless, layers may be inferred, especially in the pristine regions, by knowing that the active layer (SiO_x) and bulk (SiO_2) should contain significantly more oxygen than the electrodes (TiN). Caution must be observed with SiO_2 whose insulating nature may deflect the milling and analysis beams and lead to slow milling and low yield, respectively. On deformed regions, it appears that ^{18}O has reached the bulk. This may not be the case (as will be shown in section 6.1.2) as perpendicular ion milling may well implant ^{18}O species deep into the device. On these same regions, it is possible to also observe significant total oxygen depletion in the active layer but also a slower milling rate. This latter observation signals the fact that layers are now harder to mill and probably yield fewer ions per unit of time. This would explain the reduction in oxygen counts. This reconciles with the fact that a lower concentration of oxygen and a higher concentration of silicon have been observed under deformed regions [58] leading to higher density. However, assuming an unchanged yield rate, a reduction in oxygen should not be surprising. This could stem out of the fact that oxygen was released from filaments and only partially replaced slowly leading to well observed device death in a LRS (section 4.1.2.1). Therefore, it is hard to conclude with certainty what might had led to this reduction in oxygen counts.

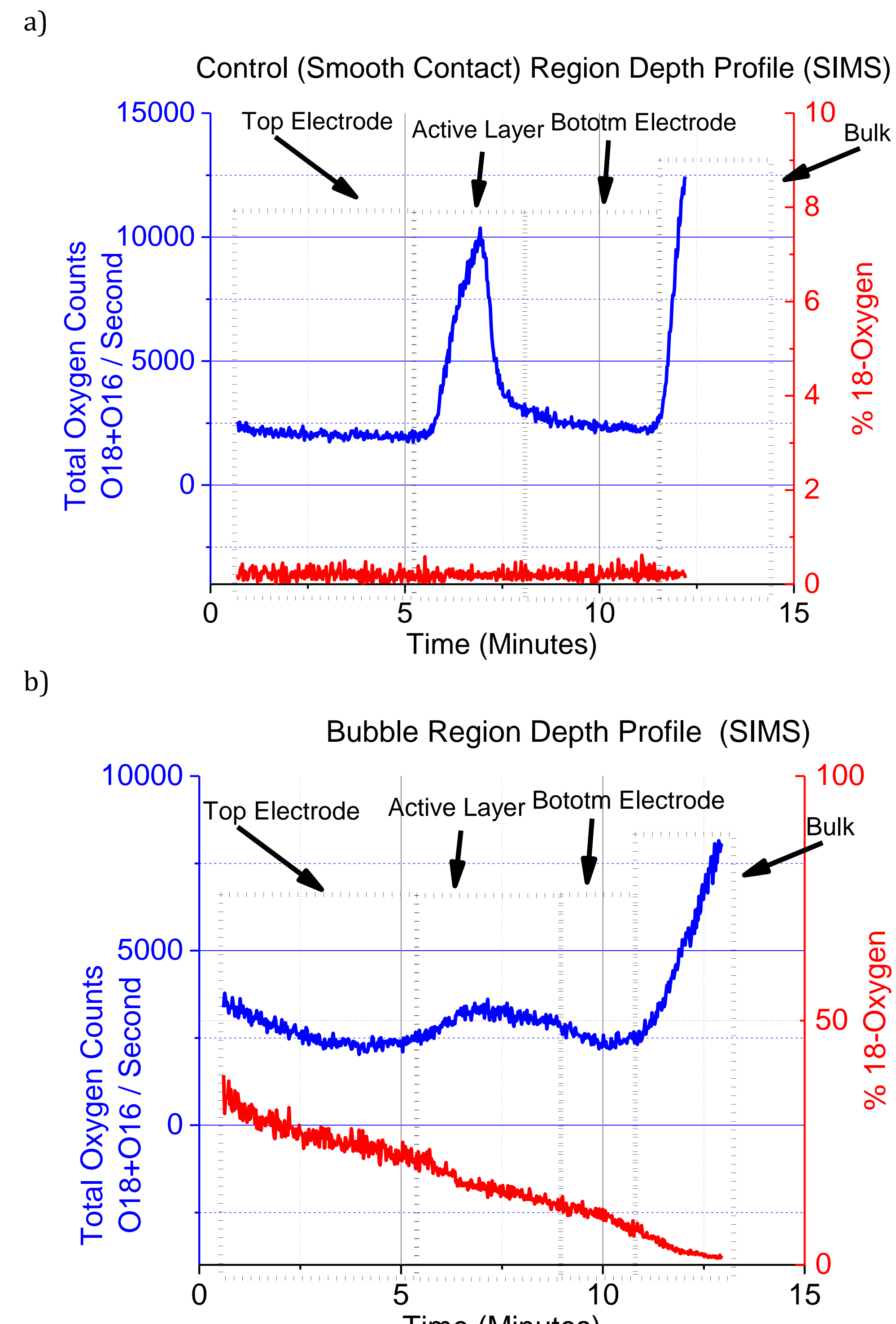
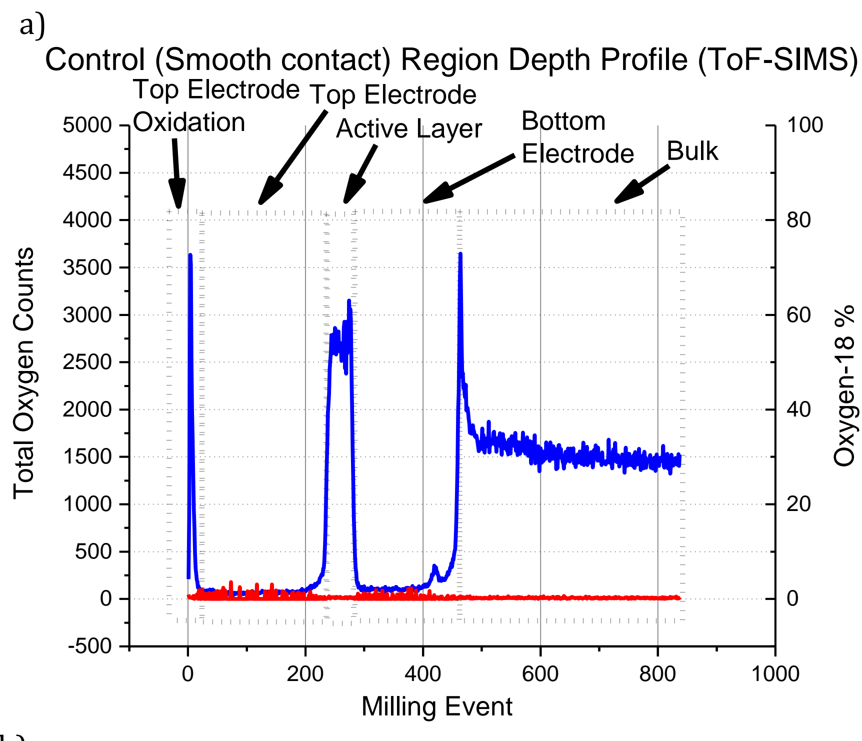


Figure 57: Oxygen depth profiles carried out by FIB-SIMS: a) profile carried out on a pristine region of a TiN anode following cycling in $^{18}O_2$ –rich atmosphere. The ^{18}O ratio in this region is no higher than the natural abundance and the active layer is rich in oxygen. b) similar profile carried out on a deformed region following cycling. ^{18}O is well embedded into the device and, assuming an unchanged yield, a strong depletion of oxygen is observed. In both cases the presumed device layers are highlighted.

Figure 58 shows ToF-SIMS depth profiles. Here, one observes, once again, a clear difference between deformed and unaffected regions. Deformed regions lead to the detection of ^{18}O levels above natural-occurrence down to the bulk. This, once again, appears to be due to ^{18}O implantation by the perpendicular milling beam used in the ToF-SIMS (as discussed in 6.1.2 ^{18}O is not found in the bottom electrode). Milling rates within deformed regions appear, as it was the case with the SIMS, slower as if the different layers have acquired physical strength or become denser following cycling. This makes it hard to directly compare total oxygen counts, which are now inexplicably higher in the cycled active layer than the pristine one. As it was the case with SIMS analysis, within the deformed region, it is possible to have a rough idea of where layers end and start but certainty is not possible as no sharp change in oxygen counts occurs within layers. The drop in yield at around milling event 500 can either be justified by temporary charging or a transient loss in beam current. It is also possible to observe that the ToF-SIMS uses a lower energy beam compared to the FIB SIMS as edges between layers are now sharper. This is because a less energetic beam obtains information by looking at a shallower depth and therefore does not look 'ahead' as much as a higher energy one.

a)



b)

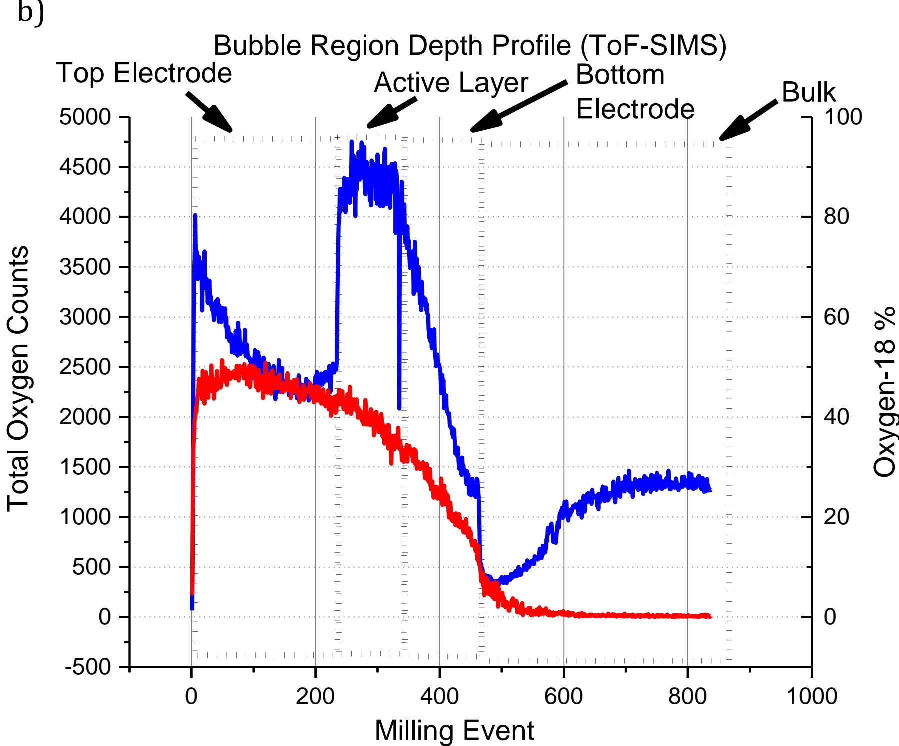


Figure 58: Oxygen depth profiles carried by ToF-SIMS instead of FIB-SIMS. Results are similar but differ in amount of total oxygen detected. A lower beam energy may be observed compared to the previous figure as edges are now sharper.

Overall, from the two different depth profiling instruments one can come up with a few conclusions: total oxygen counts are not to be trusted making it impossible to do direct comparisons of oxygen amounts in pristine and cycled

layers, cycled layers appear slower to mill meaning that layers must have become denser or otherwise become physically stronger, ^{18}O only gets embedded under deformed regions, appears to reach the bulk but beam ions may well be increasing this depth by fostering unwanted implantation. Such Incorporation shows that devices interact with atmospheric oxygen. These results find support in one published paper [75] in which the author observed the presence of ^{18}O in the active layer of devices positively biased in an atmosphere containing $^{18}O_2$. Other than this, the only relevant literature appears to be the one on temperature-driven oxygen exchange in oxides [112-116] which, assuming that high temperatures are reached during reset, supports this work's findings. However, in terms of models, it was only possible to find some that involve lengthy oxide exposure to high temperatures. These are not compatible with this work's dramatically faster reset pulses. Similarly, the literature focuses on diffusion depth on smooth surfaces when here escaping oxygen probably opens a very large number of pathways to the device active layer. This renders the concept of diffusion depth impractical in this work's context as oxygen diffusion is heavily facilitated and thin layers (possibly atomic) of ^{18}O exchange easily given the sponge-like structure observed.

Elemental FIB SIMS analysis could also be carried out, perpendicularly, on a surface to produce a 2D map of fractional ^{18}O with respect to total oxygen. Here, once again, it is best to work with fractions as absolute counts would depend on oxygen ion yield. To come up with fractional maps, two subsequent scans were always used: one for mass-to-charge-ratio 18 and one for mass-to-charge-ratio 16. ^{16}O scans were carried out every two ^{18}O scans. This allowed to generate ratio-maps while collecting more information about ^{18}O than about ^{16}O. Given that the acquisition of each map etches the device surface, each subsequent map

corresponds to an increased depth within the device. While this does not significantly affect fractional results, it allowed to produce a 3D tomography of the ^{18}O fraction.

Tomography analysis was carried out on the deformed region shown as an electron image in Figure 59a. The region appeared during cycling in ^{18}O-rich atmosphere and its size of about 10 x 15 micrometres made it a good candidate. This size is like the one found in the literature for similar devices [60] and is probably related to the amount of emitted oxygen and location of paths of least resistance for its escape. The tomography results are shown in Figure 59b. Here starting from the top left one finds the device surface and moving to the right and down the bulk is reached where one no longer see a high concentration of ^{18}O. Allocating layers to maps was not attempted as this would not be easy. Different regions of the map may undergo milling at a different rate. It is possible that subsequent maps fell victim of ion implantation because of the Ga^+ beam used for milling and analysis. However, what is important is that under the deformed region, ^{18}O concentration does not expand beyond the original shape of the original region. This means that oxygen exchange is constrained to those regions where oxygen previously escaped opening pathways for atmospheric oxygen to enter. The literature [112-116], however, shows that oxygen exchange requires heat to occur in large amount. Here three explanations may be given. The first is that a very large area (at least as large as the deformed region) heats up every reset cycle. The second is that every cycle the filament moves and by doing so heats up new regions where oxygen exchange then occurs. Filament movement, although not found in the literature, also finds support by the fact that deformed regions spread suddenly every few cycles (section 5.1). This means that new paths of oxygen least resistance must be found or else new

deformed regions would not be needed hence justifying a moving oxygen source/filament. The third explanation would be that once oxygen pathways have been opened, atmospheric oxygen quite simply reacts, irrespectively of temperature, with some highly reactive silicon dangling bonds leading to low grade ^{18}O incorporation. This last explanation, however, does not justify the high ^{18}O concentration found under deformed regions.

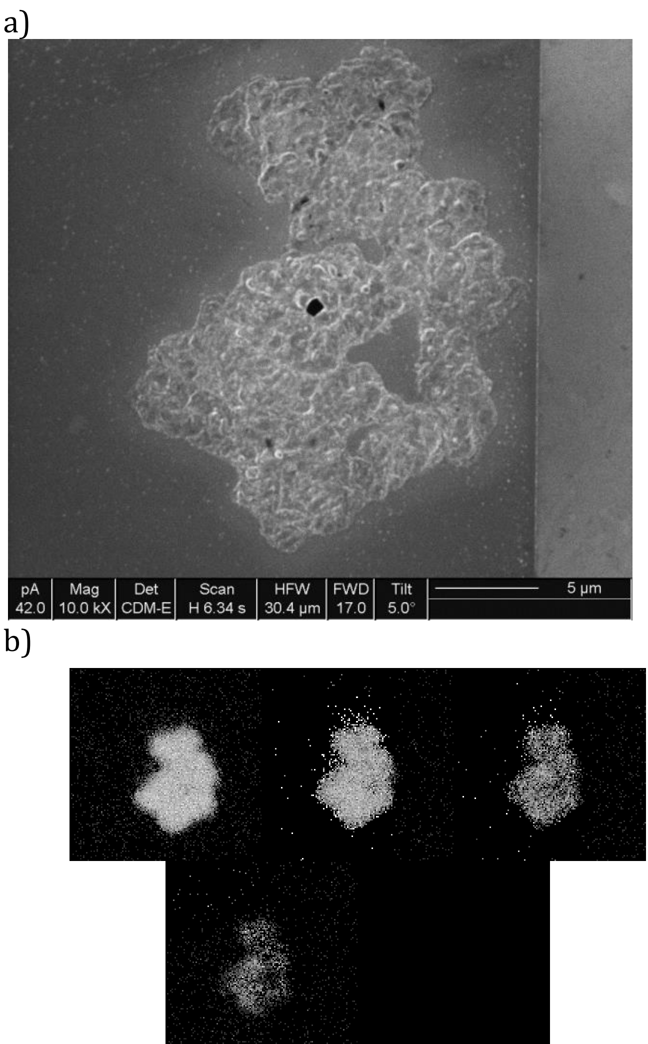


Figure 59: Images of surface deformations: a) FIB electron image showing surface deformations created from cycling a device 5 times in $^{18}O_2$-rich atmosphere. b) ^{18}O percentage maps of part of the same deformed region at a certain depth; the contour follows the topographical contour of the deformed region visible on the surface. As depth increases (bulk by 4th image), the amount of absorbed $^{18}O_2$ decreases.

At this point it has been shown that devices acquire environmental oxygen during cycling. This, along with previous evidence of superoxide ion (and possibly oxygen gas) emission, shows the devices as being extremely reliant on oxygen and interacting with atmosphere to release and obtain it. This dependency could probably be reduced by using layers of materials that while conductive act as oxygen sinks and sources such as graphene [56].

6.1.2 Further SIMS Studies

Further SIMS studies were carried out by Dr Richard Chater using samples previously cycled in this study in the same $^{18}O_2$-rich atmosphere already discussed. SIMS analysis found in section 6.1.1 relied only on the use of an electron beam normal to the sample surface. It was assumed that this might have led to the implantation of ions into subsequent layers and therefore distorted results by showing ^{18}O deeper within the sample than it was.

Dr Chater used the same FIB-SIMS previously discussed. The ion beam was first used to dig ramps at a 4 ° angle on and off deformed regions. This exposed the different device layer virtually magnifying them by around $\frac{\sin(86^{\circ})}{\sin(4^{\circ})} = 14.3$ times. This process also led to limited ion implantation from previous layers. This is because implantation now only occured within the extremities of large regions. Following this step, the ion beam was used for analysis on the now exposed (and magnified) layers. In particular, the two TiN electrodes were looked at leading to confirmation that ^{18}O is present in the top electrode but showing that ^{18}O is not present beyond natural abundance in the bottom one. Results may be found in Figure 60 where ^{18}O fraction may be seen for the top and bottom electrodes of 4° angle ramps carried out on deformed a) and pristine b) regions. These figures are to be used to look at the

^{18}O isotopic fraction first obtained on the now exposed top electrode and then, following beam relocation, on the bottom electrode. One can see that within the top electrode there is between 20% and 10% ^{18}O with respect to total oxygen on the deformed region. An insignificant amount (natural aboundance) of ^{18}O is present in the bottom electrode within the same region. A similar amount of ^{18}O is also present in both electrodes within the tested smooth region. These results shows that some ion implantation occurred during the depth profiles and tomography carried out in section 6.1.1.

It would be of great interest to also obtain results of ^{18}O fractional concentration in the active (SiO_x) layer using this same method along with even shallower angles. However, the high abundance of ^{18}O previously found in this layer cannot purely be due to ion implantation. It would be highly improbable for the shallow ramp approach not to find any ^{18}O in this layer.

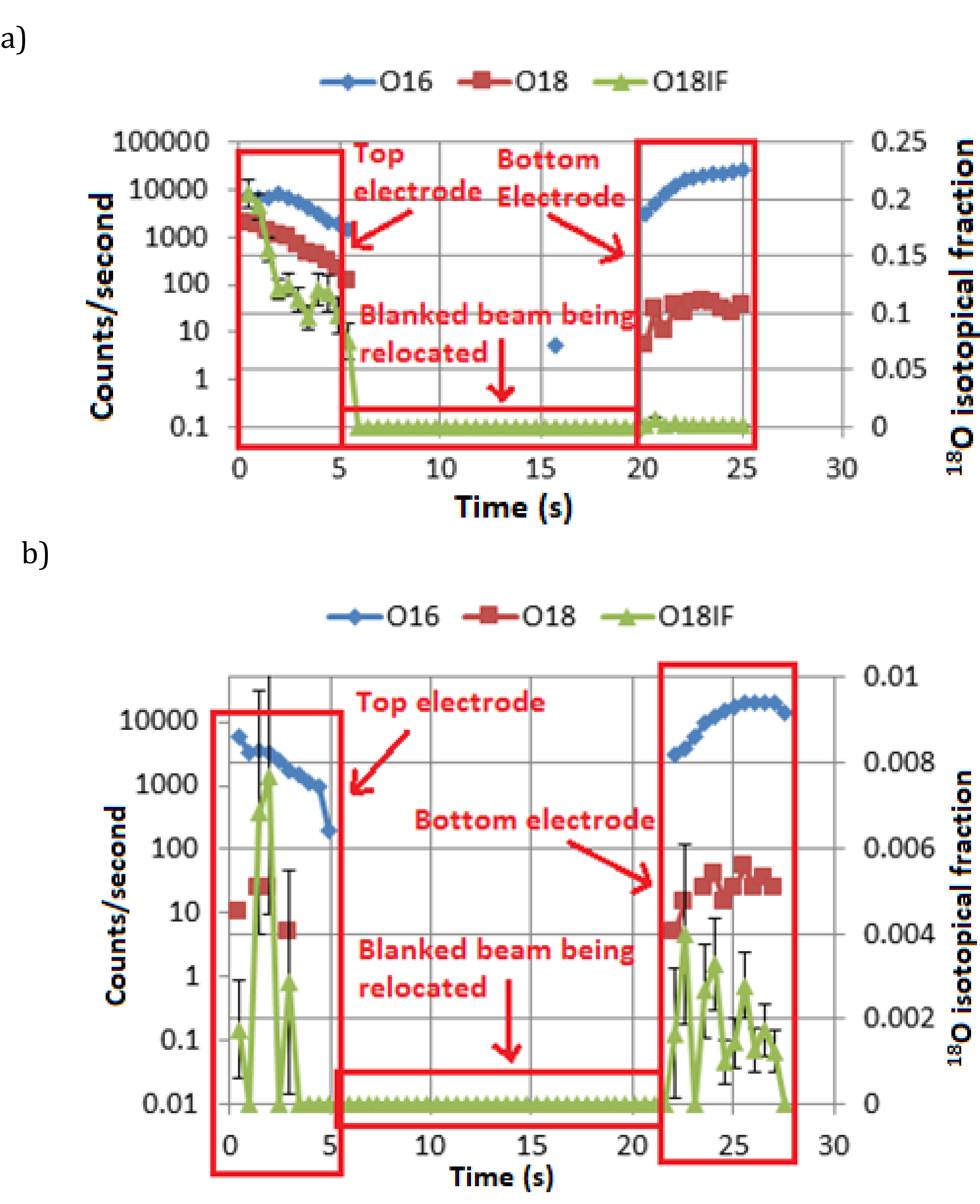


Figure 60: Oxygen profiles of top and bottom electrodes ramps: a) ^{18}O fraction of top and bottom electrode extracted through 4° ramp on deformed region. b) ^{18}O fraction of top and bottom electrode extracted through 4° ramp on pristine region. Both figures adapted from figures by Dr Richard Chater.

The finding that results found in section 6.1.1 might have been affected by ion implantation does not change the important results. It can still be stated that ^{18}O embedding only occurs under deformed regions and therefore that atmospheric penetration pathways are presents within such regions. This finding strengthens the

one from section 4.2.2.4 further confirming that devices electroformed in air need atmospheric oxygen to reset. Also, the finding from section 4.2.3.3 is strengthened as it has now been shown that atmospheric oxygen reaches the inside of devices; this helps explain why atmospheres very rich in oxygen may hinder set performance.

7 Summary, Conclusions and Future Work

Using already existing UCL's unipolar TiN/SiO_x/TiN silicon rich silica devices it was possible to study some of their electrical, physical and chemical properties to find out how their endurance could be improved. To do so, several analysis techniques were used: IV curve characterisation, pulsed characterisation (with and without compliance), in situ RGA analysis, SIMS surface analysis, ToFSIM surface analysis, as well as operation in modified atmospheres.

Chapter 4 was used to present the electrical findings. Firstly, it is shown that when it comes to data storage, devices perform similarly to the experimental ones found in the literature with low switching voltages, and good retention and reliability. This is established using IV curves and pulses with current compliance. Comparing this family of experimental devices to state of art ones does not make sense as the objective was not to improve performance but to learn more about device behaviour. Novel results are then obtained by studying how the same devices respond to quick electrical stimuli in the tens of nanoseconds without current compliance. Devices toggled between states with identical pulses. Expanding this programming approach also led to neuromorphic behaviour involving potentiation, depression and short term memory emulation. These results open doors to simplified programming circuits in data storage applications and the possibility of using TiN/SiO_x/TiN devices in circuits that attempt to emulate some of the biological brain behaviour. Programming through fast pulses, did not lead to endurance hindering electrode deformations that are generally attributed to oxygen movement in this family of devices; this may be due to tighter control on timing. Fast pulse

behaviour, however, showed high variability which must be overcome by reducing device process variation in future generations.

Still within chapter 4, novel results are presented involving electrical device behaviour in modified atmosphere. Here it is shown that device behaviour is affected by its external environment. It appears easier to melt down devices in high vacuum than it is in air, a higher than air oxygen partial pressure makes setting hard, and resetting is favoured by the presence of oxygen in the atmosphere. These results show that a high level of interaction occurs between device and atmosphere.

Chapter 5 presents novel results that expand on those found in the second half of chapter 4. Here, by switching devices in situ in a high vacuum, the emission of oxygen species during electroforming is observed. Superoxide ions and well as uncharged molecular oxygen are detected. This strengthens the previous results which showed device-atmosphere interaction.

Chapter 6 expands on the previous atmospheric interaction results by showing that environmental oxygen is incorporated within devices during cycling. While atmospheric oxygen is only found under those contact regions that show physical deformation, it was impossible to determine whether incorporated oxygen is used by the switching mechanism.

Overall, this work on unipolar $TiN/SiO_x/TiN$ devices shows results that open doors to simplified programmers, neuromorphic applications and solutions to endurance issues. Using fast pulses leads to a lack of detrimental electrode deformations. This can be directly translated into increased endurance. However, having seen the high level of atmosphere-device interaction, it cannot be ruled out that some electrode deformations may still occur under stressful operating

conditions. Therefore, trying to place a barrier between devices and atmosphere may be useful. Such barrier may be a conductive layer of a material (such as graphite) which can store oxygen without undergoing physical of chemical changes. Some of these findings may be applicable to bipolar devices as these devices are also used in neuromorphic applications. However, bipolar devices do not seem to show the same endurance problems as unipolar devices.

A. APPENDIX

A.1 LIST OF PUBLICATIONS ARISING FROM THIS WORK

A.1.1 Journal Publications

1) “Nanosecond Analog Programming of Substoichiometric Silicon Oxide Resistive RAM,” L Montesi, M Buckwell, K Zarudnyi, L Garnett, S Hudziak, A Mehonic, A Kenyon, IEEE Transactions on Nanotechnology (Volume:15 , Issue: 3, Pages 428 - 434), May, 2016.

2) "Nanoscale Transformations in Metastable, Amorphous, Silicon-Rich Silica,"A. Mehonic, M. Buckwell, L. Montesi, M. S. Munde, D. Gao, S. Hudziak, R. J. Chater, S. Fearn, D. McPhail, M. Bosman, A. L. Shluger, A. J. Kenyon, Advance Materials (Volume 28, Issue 34, Pages 7486 – 7493), September, 2016.

3) “Conductance tomography of conductive filaments in intrinsic silicon-rich silica RRAM,” M. Buckwell, L. Montesi, S. Hudziak, A. Mehonic, A. J. Kenyon, Nanoscale (Volume 7, Issue 43, Pages 18030 - 18035), October, 2015.

4) “Structural changes and conductance thresholds in metal-free intrinsic SiO_x resistive random access memory,” A. Mehonic, M. Buckwell, L. Montesi, L. Garnett, S. Hudziak, S. Fearn, R. Chater, D. McPhail and A. J. Kenyon, Journal of Applied Physics (Volume 117, Issue 12, Pages 124505-1 – 124505-8), March, 2015

5) “Microscopic and spectroscopic analysis of the nature of conductivity changes during resistive switching in silicon-rich silicon oxide,” M. Buckwell, L. Montesi, A. Mehonic, O. Reza, L. Garnett, M. Munde, S. Hudziak and A. J. Kenyon, Physica Status Solidi (C) Current Topics in Solid State Physics (Volume 12, Issue 1–2, Pages 211–217), January, 2015.

A.1.2 Conference Presentations Given

1) “Silicon Oxide ReRAM,” A. Mehonic, L. Montesi, M. Munde, A. J. Kenyon, Materials Reasearch Society, Spring 2014, San Francisco, USA, BB13.04.

2) “Dynamic Mass-Spectrometry Characterization of Oxygen Emission during Operation of SiOx ReRAM Devices”, L. Montesi, M. Munde, M. Buckwell, L. Garnett, A. Mehonic, R. Chater, S. Fearn, S. Hudziak, D. McPhail, A. Kenyon. Materials Reasearch Society, Fall 2014, Boston, USA, M3.03.

3) “Nano-Second Scale Neuromorphic Behaviour in SiOx Thin-Film ReRAM Devices”, L. Montesi, M. Buckwell, L. Garnett, A. Mehonic, A. Hudziak, A.J. Kenyon, European Materials Reasearch Society, Spring 2015, Lille, France.

4) “Nano-Second Pulse Programming of Resistive RAM Devices and Its Benefits”, L. Montesi, M. Buckwell, A. Mehonic, S. Hudziak, A.J. Kenyon, Electrochemical Society, Spring 2015, Chicago, USA.

5) “Structural investigation of resistance switching in silicon-rich silica films”, M. Buckwell, L. Montesi, A. Mehonic, A. J. Kenyon, S. Fearn, R. J. Chater, D. McPhail, IEEE Nano 2015, Rome, Italy.

6) “Using SiOx Thin-Film ReRAM Devices to Emulate Neuronal Behaviour”, L. Montesi, M. Buckwell, A. Mehonic, S. Hudziak, A. J. Kenyon, Materials Reasearch Society, Fall 2015, Boston, USA, KK10.17.

7) “The Crucial Role of Oxygen in Oxide-Based ReRAM: Oxygen Emission and Device Behaviour in a Labelled Oxygen Environment”, L. Montesi, M. Buckwell, A. Mehonic, S. Hudziak, R. Chater, D. McPhail, A. J. Kenyon, Materials Reasearch Society, Fall 2015, Boston, USA, KK10.18.

8) “Exploring ReRAM through isotopically labelled oxygen atmosphere: investigating filamentation and switching mechanics”, L. Montesi, M. Buckwell, C. van den Bosch, R. Chater, S. Fearn, A. Mehonic, A. Aguadero, A. J. Kenyon, European Materials Reasearch Society, Spring 2015, Lille,France, O.9.33.

9) “The Interaction of SiO x-Based Resistive RAM Devices with Oxygen: Observations on Release and Absorption”, L. Montesi, M. Buckwell, C. A. M. van den Bosch, R. J. Chater, S. Fearn, A. Mehonic, A. Aguadero, and A. J. Kenyon, Electrochemical Society, PRiME 2016, Honolulu, USA.

10) “Nanosecond Programming of SiO x Based Resistive RAM: Simplified Digital Programmers and Analog Neuromorphic Type Behavior”, L. Montesi, M. Buckwell, K. Zarudnyi, L. Garnett, S. Hudziak, A. Mehonic, and A. J. Kenyon, Electrochemical Society, PRiME 2016, Honolulu, USA.

11) “Oxygen “Breathing" in Intrinsic SiOx-Based ReRAM Devices”, L. Montesi, M. Buckwell, C. A. M. Bosch, R. J. Chater, S. Fearn, A. Aguadero, A. Mehonic, A. Kenyon, Materials Reseach Soceity, Fall 2016, Boston, USA, EM10.3.03.

A.2 Description of Software Scripts and Drivers

A.2.1 Keithley 4200-SCS SMU Matlab Scripts

Script 1: Allows pulse width and current compliance to be defined. Current is sampled during measurement and resistance is calculated, displayed and stored to file as comma delimited.

Script 2: A reading voltage is first specified. Programming pulses are then specified (these last about 14 milliseconds unless compliance kicks in) as before but they are now always followed by a reading pulse during which current sampling occurs. Similar user interaction as Script 1.

Script 3: Reading voltage, delay between pulses, pulse duration above minimum, electroform voltage (and compliance) set voltage (and compliance) and reset voltage are specified at the beginning of the script code. Upon execution, the script asks whether the device needs electroforming. If so, an electroforming pulse is sent. Following this, the device will be reset and set. The resistance of the two states will be shown to the user to determine whether the device is working. If the user decides so, the device will be switched between states as many times as the user specifies. Data are stored in a comma delimited file

and plotted, allowing for easy assessment of endurance.

Script 4: Allows the device to be placed in a certain state and then have its resistance sampled at specific time intervals using a specified reading voltage. This allows for reliability over time to be tested. Data are displayed and saved to file as comma delimited.

A.2.2 Keithley 4200-SCS PMU C Drivers

Driver 1: Allows a single pulse to be applied. This pulse can be as short as 20 nanoseconds with 20 nanosecond rise and fall times (Section 3.3 shows output curves). This is mainly intended for device programming.

Driver 2: Allows a single pulse to be applied. Current is sampled between 75 and 90% of the pulse width (this does not include rise and fall times). The shortest pulse that may be applied is 20 nanoseconds long. The average of both sampled voltage and current are returned. The sampled voltage is important due to impedance mismatch issues which will be discussed later.

Driver 3: A pulse sequence can be passed to this driver. This is specified by start/stop voltages as well as segment duration. Here each segment can be as short as 20 nanoseconds (if sampling occurs this minimum length increases to 35 nanoseconds, but 160 nanoseconds is recommended). Data points must be continuous. Sampling occurs between 75% and 90% of a pulse only when requested. Data are passed in the form of an array and returned as an array. Once again,

both voltage and current are sampled.

A.2.3 Keithley 4200-SCS PMU Matlab Scripts to Control C Drivers

Script 1: Interactively allows the user to apply a pulse of a specific voltage and duration. Voltage and current are sampled and resistance is displayed for a specified reading voltage. The device is left floating between pulses.

Script 2: Carries out sweeps where pulse duration is increased to easily determine when electroforming of a device occurs. The device is left floating between pulses.

Script 3: Reads a comma delimited file containing thousands of data points and applies them to the device under test (DUT). Collected data is therefore saved in a new comma delimited file, which includes the input. Sampled measured resistance is plotted on a graph, which also shows the input voltage specified (not measured) and timing. The voltage specified in the input waveform is always applied to the device; during waveform execution the device cannot be left floating.

A.2.4 Keithly 6430 Matlab Scripts

Script 1: Start and stop voltages are defined along with the required step. A reading voltage is also specified. During the sweep, each programming pulse is followed by a smaller reading pulse. Resistance measured during the reading pulse is what is displayed and recorded. The instrument output is brought back to floating between pulses.

Script 2: A regular sweep is carried out. Here sampling of current occurs during the programming pulse which makes up the sweep (this takes about 1.24 seconds depending on the stability of the resistance state). However, the output is brought back to floating between pulses.

A.3 REFERENCES

[1] R. Waser *et al.*, "Introduction to Nanoionic Elements for Information Technology," in *Resistive Switching*: Wiley-VCH Verlag GmbH & Co. KGaA, 2016, pp. 1-30.

[2] A. Mehonic, "Resistive Switching in Silicon-rich Silicon Oxide," PhD, University College London, 2014.

[3] A. Chen, "Emerging research device roadmap and perspectives," in *2014 IEEE International Conference on IC Design & Technology*, 2014, pp. 1-4.

[4] F. Pan, S. Gao, C. Chen, C. Song, and F. Zeng, "Recent progress in resistive random access memories: Materials, switching mechanisms, and performance," *Materials Science and Engineering: R: Reports,* vol. 83, no. 0, pp. 1-59, 9// 2014.

[5] F. Pan, C. Chen, Z.-s. Wang, Y.-c. Yang, J. Yang, and F. Zeng, "Nonvolatile resistive switching memories-characteristics, mechanisms and challenges," *Progress in Natural Science: Materials International,* vol. 20, no. 0, pp. 1-15, 11// 2010.

[6] "The international Technology Roadmap for Semiconductors 2015 Edition," vol. Beyond CMOS*,*

[7] L. M. Grupp *et al.*, "Characterizing flash memory: Anomalies, observations, and applications," in *2009 42nd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 2009, pp. 24-33.

[8] S. Lee, "Scaling Challenges in NAND Flash Device toward 10nm Technology," in *2012 4th IEEE International Memory Workshop*, 2012, pp. 1-4.

[9] G. W. Burr, B. N. Kurdi, J. C. Scott, C. H. Lam, K. Gopalakrishnan, and R. S. Shenoy, "Overview of candidate device technologies for storage-class memory," *IBM Journal of Research and Development,* vol. 52, no. 4.5, pp. 449-464, 2008.

[10] J. A. Carballo, W. T. J. Chan, P. A. Gargini, A. B. Kahng, and S. Nath, "ITRS 2.0: Toward a re-framing of the Semiconductor Technology Roadmap," in *2014 IEEE 32nd International Conference on Computer Design (ICCD)*, 2014, pp. 139-146.

[11] W. T. J. Chan, A. B. Kahng, S. Nath, and I. Yamamoto, "The ITRS MPU and SOC system drivers: Calibration and implications for design-based equivalent scaling in the roadmap," in *2014 IEEE 32nd International Conference on Computer Design (ICCD)*, 2014, pp. 153-160.

[12] P. Gargini, "The International Technology Roadmap for Semiconductors (ITRS): "Past, present and future"," in *GaAs IC Symposium, 2000. 22nd Annual*, 2000, pp. 3-5.

[13] A. B. Kahng, "The ITRS design technology and system drivers roadmap: Process and status," in *Design Automation Conference (DAC), 2013 50th ACM/EDAC/IEEE*, 2013, pp. 1-6.

[14] A. B. Kahng, "Roads not taken," *IEEE Design & Test of Computers,* vol. 28, no. 2, pp. 74-75, 2011.

[15] A. Sheikholeslami and P. G. Gulak, "A survey of circuit innovations in ferroelectric random-access memories," *Proceedings of the IEEE,* vol. 88, no. 5, pp. 667-689, 2000.

[16] Y. Kato, T. Yamada, and Y. Shimada, "0.18-μm nondestructive readout FeRAM using charge compensation technique," *IEEE Transactions on Electron Devices,* vol. 52, no. 12, pp. 2616-2621, 2005.

[17] S. Parkin, J. Xin, C. Kaiser, A. Panchula, K. Roche, and M. Samant, "Magnetically engineered spintronic sensors and memory," *Proceedings of the IEEE,* vol. 91, no. 5, pp. 661-680, 2003.

[18] M. Julliere, "Tunneling between ferromagnetic films," *Physics Letters A,* vol. 54, no. 3, pp. 225-226, 1975/09/08 1975.

[19] W. J. Gallagher and S. S. P. Parkin, "Development of the magnetic tunnel junction MRAM at IBM: From first junctions to a 16-Mb MRAM demonstrator chip," *IBM Journal of Research and Development,* vol. 50, no. 1, pp. 5-23, 2006.

[20] L. Montesi, Z. Zilic, T. Hanyu, and D. Suzuki, "Building Blocks to Use in Innovative Non-volatile FPGA Architecture Based on MTJs," in *VLSI (ISVLSI), 2012 IEEE Computer Society Annual Symposium on*, 2012, pp. 302-307.

[21] R. Waser, *Nanoelectronics and Information Technology*, 3rd ed. Wiely, 2012.

[22] L. Montesi, *Use of Magnetic Tunnel Junction Devices to Implement a Non-Volatile FPGA Based on a Pre-Existing Architecture*. McGill University Libraries, 2011.

[23] S. Raoux *et al.*, "Phase-change random access memory: A scalable technology," *IBM Journal of Research and Development,* vol. 52, no. 4.5, pp. 465-479, 2008.

[24] G. Sun, X. Dong, Y. Xie, J. Li, and Y. Chen, "A novel architecture of the 3D stacked MRAM L2 cache for CMPs," in *2009 IEEE 15th International Symposium on High Performance Computer Architecture*, 2009, pp. 239-249.

[25] K. Kyung Min, J. Doo Seok, and H. Cheol Seong, "Nanofilamentary resistive switching in binary oxide system; a review on the present status and outlook," *Nanotechnology,* vol. 22, no. 25, p. 254002, 2011.

[26] D. R. Lamb and P. C. Rundle, "A non-filamentary switching action in thermally grown silicon dioxide films," *British Journal of Applied Physics,* vol. 18, no. 1, p. 29, 1967.

[27] C. M. Osburn and D. W. Ormond, "Dielectric Breakdown in Silicon Dioxide Films on Silicon: I . Measurement and Interpretation," *Journal of The Electrochemical Society,* vol. 119, no. 5, pp. 591-597, May 1, 1972 1972.

[28] A. M. S. G. Dearnaley, D.V. Morgan, "Electrical Phenomena in Amorphous Oxide Films," *Rep. Prog. Phys.,* vol. 33, pp. 1129-1191, 1970.

[29] T. W. Hickmott, "Low-Frequency Negative Resistance in Thin Anodic Oxide Films," *Journal of Applied Physics,* vol. 33, no. 9, pp. 2669-2682, 1962.

[30] W. Zhang, Z. Chai, J. Ma, J. F. Zhang, and Z. Ji, "Analysis of RTN signals in resistive-switching RAM device and its correlation with device operations," in *2016 13th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, 2016, pp. 174-177.

[31] A. Bricalli, E. Ambrosi, M. Laudato, M. Maestro, R. Rodriguez, and D. Ielmini, "SiOx-based resistive switching memory (RRAM) for crossbar storage/select elements with high on/off ratio," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 4.3.1-4.3.4.

[32] R. Waser and M. Aono, "Nanoionics-based resistive switching memories," *Nat Mater,* 10.1038/nmat2023 vol. 6, no. 11, pp. 833-840, 11//print 2007.

[33] A. Hanada, K. Kinoshita, K. Matsubara, T. Fukuhara, and S. Kishida, "Developmental mechanism for the resistance change effect in perovskite oxide-based resistive random access memory consisting of Bi2Sr2CaCu2O8+δ bulk single crystal," *Journal of Applied Physics,* vol. 110, no. 8, p. 084506, 2011.

[34] A. Sawa, "Resistive switching in transition metal oxides," *Materials Today,* vol. 11, no. 6, pp. 28-36, 6// 2008.

[35] Q. Lu and B. Yildiz, "Voltage-Controlled Topotactic Phase Transition in Thin-Film SrCoOx Monitored by In Situ X-ray Diffraction," *Nano Letters,* vol. 16, no. 2, pp. 1186-1193, 2016/02/10 2016.

[36] A. Mehonic *et al.*, "Quantum Conductance in Silicon Oxide Resistive Memory Devices," *Sci. Rep.,* Article vol. 3, 09/19/online 2013.

[37] E. Miranda, A. Mehonic, J. Suñé, and A. J. Kenyon, "Multi-channel conduction in redox-based resistive switch modelled using quantum point contact theory," *Applied Physics Letters,* vol. 103, no. 22, pp. -, 2013.

[38] P. Jubong *et al.*, "New Set/Reset Scheme for Excellent Uniformity in Bipolar Resistive Memory," *Electron Device Letters, IEEE,* vol. 32, no. 3, pp. 228-230, 2011.

[39] L. Montesi *et al.*, "Nanosecond Analog Programming of Substoichiometric Silicon Oxide Resistive RAM," *IEEE Transactions on Nanotechnology,* vol. PP, no. 99, pp. 1-1, 2016.

[40] G. Wang, A. C. Lauchner, J. Lin, D. Natelson, K. V. Palem, and J. M. Tour, "High-Performance and Low-Power Rewritable SiOx 1 kbit One Diode–One Resistor Crossbar Memory Array," *Advanced Materials,* vol. 25, no. 34, pp. 4789-4793, 2013.

[41] C. Schindler, G. Staikov, and R. Waser, "Electrode kinetics of Cu–SiO2-based resistive switching cells: Overcoming the voltage-time dilemma of electrochemical metallization memories," *Applied Physics Letters,* vol. 94, no. 7, pp. 072109-072109-3, 2009.

[42] C. Kuan-Chang *et al.*, "Characteristics and Mechanisms of Silicon-Oxide-Based Resistance Random Access Memory," *Electron Device Letters, IEEE,* vol. 34, no. 3, pp. 399-401, 2013.

[43] L. Ji *et al.*, "Integrated One Diode–One Resistor Architecture in Nanopillar SiOx Resistive Switching Memory by Nanosphere Lithography," *Nano Letters,* vol. 14, no. 2, pp. 813-818, 2014/02/12 2013.

[44] S. Jungho *et al.*, "Effect of Program/Erase Speed on Switching Uniformity in Filament-Type RRAM," *Electron Device Letters, IEEE,* vol. 32, no. 7, pp. 958-960, 2011.

[45] B. J. Choi *et al.*, "Electrical Performance and Scalability of Pt Dispersed SiO2 Nanometallic Resistance Switch," *Nano Letters,* vol. 13, no. 7, pp. 3213-3217, 2013/07/10 2013.

[46] B. S. Dmitri and K. L. Konstantin, "CMOL FPGA: a reconfigurable architecture for hybrid digital circuits with two-terminal nanodevices," *Nanotechnology,* vol. 16, no. 6, p. 888, 2005.

[47] W. Zhao, E. Belhaire, V. Javerliac, C. Chappert, and B. Dieny, "Evaluation of a Non-Volatile FPGA based on MRAM technology," in *2006 IEEE International Conference on IC Design and Technology*, 2006, pp. 1-4.

[48] G. Indiveri, B. Linares-Barranco, R. Legenstein, G. Deligeorgis, and T. Prodromakis, "Integration of nanoscale memristor synapses in neuromorphic computing architectures," *Nanotechnology,* vol. 24, no. 38, p. 384010, 2013.

[49] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale Memristor Device as Synapse in Neuromorphic Systems," *Nano Letters,* vol. 10, no. 4, pp. 1297-1301, 2010/04/14 2010.

[50] C. Ting, Y. Yuchao, and L. Wei, "Building Neuromorphic Circuits with Memristive Devices," *Circuits and Systems Magazine, IEEE,* vol. 13, no. 2, pp. 56-73, 2013.

[51] R. Yang *et al.*, "On-Demand Nanodevice with Electrical and Neuromorphic Multifunction Realized by Local Ion Migration," *ACS Nano,* vol. 6, no. 11, pp. 9515-9521, 2012/11/27 2012.

[52] S. Yu, B. Gao, Z. Fang, H. Yu, J. Kang, and H. S. P. Wong, "A Low Energy Oxide-Based Electronic Synaptic Device for Neuromorphic Visual Systems with Tolerance to Device Variation," *Advanced Materials,* vol. 25, no. 12, pp. 1774-1779, 2013.

[53] M. Buckwell, L. Montesi, S. Hudziak, A. Mehonic, and A. Kenyon, "Conductance tomography of conductive filaments in intrinsic silicon-rich silica RRAM," *Nanoscale,* 10.1039/C5NR04982B 2015.

[54] S. U. Sharath *et al.*, "Towards forming-free resistive switching in oxygen engineered HfO2−x," *Applied Physics Letters,* vol. 104, no. 6, p. 063502, 2014.

[55] A. Mehonic *et al.*, "Resistive switching in silicon suboxide films," *Journal of Applied Physics,* vol. 111, no. 7, pp. -, 2012.

[56] M. Buckwell *et al.*, "Microscopic and spectroscopic analysis of the nature of conductivity changes during resistive switching in silicon-rich silicon oxide," *physica status solidi (c),* vol. 12, no. 1-2, pp. 211-217, 2015.

[57] A. Mehonic *et al.*, "Structural changes and conductance thresholds in metal-free intrinsic SiOx resistive random access memory," *Journal of Applied Physics,* vol. 117, no. 12, p. 124505, 2015.

[58] A. Mehonic *et al.*, "Nanoscale Transformations in Metastable, Amorphous, Silicon-Rich Silica," *Advanced Materials,* pp. n/a-n/a, 2016.

[59] A. Mehonic *et al.*, "Electrically tailored resistance switching in silicon oxide," *Nanotechnology,* vol. 23, no. 45, p. 455201, 2012.

[60] Y. Wang, X. Qian, K. Chen, Z. Fang, W. Li, and J. Xu, "Resistive switching mechanism in silicon highly rich SiOx ($x < 0.75$) films based on silicon dangling bonds percolation model," *Applied Physics Letters,* vol. 102, no. 4, pp. -, 2013.

[61] U. Celano *et al.*, "Three-Dimensional Observation of the Conductive Filament in Nanoscaled Resistive Memory Devices," *Nano Letters,* vol. 14, no. 5, pp. 2401-2406, 2014/05/14 2014.

[62] U. Celano, Y. Yin Chen, D. J. Wouters, G. Groeseneken, M. Jurczak, and W. Vandervorst, "Filament observation in metal-oxide resistive switching devices," *Applied Physics Letters,* vol. 102, no. 12, p. 121602, 2013.

[63] H. Akinaga and H. Shima, "Resistive Random Access Memory (ReRAM) Based on Metal Oxides," *Proceedings of the IEEE,* vol. 98, no. 12, pp. 2237-2251, 2010.

[64] J. Yao, Z. Sun, L. Zhong, D. Natelson, and J. M. Tour, "Resistive Switches and Memories from Silicon Oxide," *Nano Letters,* vol. 10, no. 10, pp. 4105-4110, 2010/10/13 2010.

[65] D. Ielmini, F. Nardi, and C. Cagli, "Universal Reset Characteristics of Unipolar and Bipolar Metal-Oxide RRAM," *IEEE Transactions on Electron Devices,* vol. 58, no. 10, pp. 3246-3253, 2011.

[66] S. Long *et al.*, "Quantum-size effects in hafnium-oxide resistive switching," *Applied Physics Letters,* vol. 102, no. 18, p. 183505, 2013.

[67] T. Tohru, H. Tsuyoshi, T. Kazuya, and A. Masakazu, "Conductance quantization and synaptic behavior in a Ta 2 O 5 -based atomic switch," *Nanotechnology,* vol. 23, no. 43, p. 435705, 2012.

[68] S. Tappertzhofen, I. Valov, and R. Waser, "Quantum conductance and switching kinetics of AgI-based microcrossbar cells," *Nanotechnology,* vol. 23, no. 14, p. 145703, 2012.

[69] B. Butcher *et al.*, "Modeling the effects of different forming conditions on RRAM conductive filament stability," in *2013 5th IEEE International Memory Workshop*, 2013, pp. 52-55.

[70] A. Schonhals, R. Waser, S. Menzel, and V. Rana, "3-bit read scheme for single layer Ta2O5 ReRAM," in *Non-Volatile Memory Technology Symposium (NVMTS), 2014 14th Annual*, 2014, pp. 1-4.

[71] L. Dongsoo *et al.*, "Resistance switching of the nonstoichiometric zirconium oxide for nonvolatile memory applications," *IEEE Electron Device Letters,* vol. 26, no. 10, pp. 719-721, 2005.

[72] R. Jiang, X. Du, Z. Han, and W. Sun, "Investigation of chemical distribution in the oxide bulk layer in Ti/HfO2/Pt memory devices using x-ray photoelectron spectroscopy," *Applied Physics Letters,* vol. 106, no. 17, p. 173509, 2015.

[73] S. Menzel, "Simulation and Modeling of the Switching Dynamics in Resistive Switching Devices," *ECS Transactions,* vol. 69, no. 3, pp. 19-32, September 9, 2015 2015.

[74] D. S. Jeong, H. Schroeder, U. Breuer, and R. Waser, "Characteristic electroforming behavior in Pt/TiO2/Pt resistive switching cells depending on atmosphere," *Journal of Applied Physics,* vol. 104, no. 12, pp. -, 2008.

[75] C. Yoshida, K. Kinoshita, T. Yamasaki, and Y. Sugiyama, "Direct observation of oxygen movement during resistance switching in NiO/Pt film," *Applied Physics Letters,* vol. 93, no. 4, pp. -, 2008.

[76] H. Jiun-Jia, C.-W. Kuo, W.-C. Chang, and T.-H. Hou, "Transition of stable rectification to resistive-switching in Ti / TiO 2 / Pt oxide diode," *Applied Physics Letters,* vol. 96, no. 26, pp. 262901-262901-3, 2010.

[77] G. Bersuker *et al.*, "Metal oxide resistive memory switching mechanism based on conductive filament properties," *Journal of Applied Physics,* vol. 110, no. 12, pp. -, 2011.

[78] H. Shima, N. Zhong, and H. Akinaga, "Switchable rectifier built with Pt/TiOx/Pt trilayer," *Applied Physics Letters,* vol. 94, no. 8, pp. -, 2009.

[79] S. Long *et al.*, "Voltage and Power-Controlled Regimes in the Progressive Unipolar RESET Transition of HfO2-Based RRAM," *Sci. Rep.,* Article vol. 3, 10/14/online 2013.

[80] N. Takeki, K. Koji, M. Shunsaku, Y. Ryutaro, M. Takumi, and W. Zhiqiang, "Conductive Filament Expansion in TaO x Bipolar Resistive Random Access Memory during Pulse Cycling," *Japanese Journal of Applied Physics,* vol. 52, no. 11R, p. 114201, 2013.

[81] P. Bousoulas, I. Michelakaki, and D. Tsoukalas, "Influence of oxygen content of room temperature TiO2−x deposited films for enhanced resistive switching memory performance," *Journal of Applied Physics,* vol. 115, no. 3, pp. -, 2014.

[82] F. A. S. Al-Ramadhan and C. A. Hogarth, "Observation and compositional studies of the metallic conducting filaments in the low-resistance state (ON-state) of SiO/V2O5 thin films used as memory elements," (in English), *Journal of Materials Science,* vol. 19, no. 6, pp. 1939-1946, 1984/06/01 1984.

[83] H. Boris *et al.*, "Interface engineered HfO 2 -based 3D vertical ReRAM," *Journal of Physics D: Applied Physics,* vol. 49, no. 21, p. 215102, 2016.

[84] E. Filatova *et al.*, "NEXAFS study of electronic and atomic structure of active layer in Al/indium tin oxide/TiO2 stack during resistive switching," *Science and Technology of Advanced Materials,* vol. 17, no. 1, pp. 274-284, 2016/01/01 2016.

[85] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*. Wiley, 2006.

[86] F.-C. Chiu, "A Review on Conduction Mechanisms in Dielectric Films," *Advances in Materials Science and Engineering,* vol. 2014, p. 18, 2014, Art. no. 578168.

[87] K. K. Ng and S. M. Sze, *Physics of Semiconductor Devices 3rd Edition*. 2006.

[88] G. A. and S. S., *Extraction of trap-assisted tunneling parameters by graphical method in thin n-Si/SiO2 structures* (no. 3). Bucuresti: INOE 2000, 2005, p. 5.

[89] A. Latreche, "An Accurate Method for Extracting the Three Fowler-Nordheim Tunnelling Parameters Using IV Characteristic," *Microelectronics and Solid State Electronics,* vol. 2, no. 4, pp. 59-64, 2013.

[90] D. J. DiMaria *et al.*, "Charge transport and trapping phenomena in off-stoichiometric silicon dioxide films," *Journal of Applied Physics,* vol. 54, no. 10, pp. 5801-5827, 1983.

[91] A. Fischer, M. S. Mousa, and R. G. Forbes, "Influence of barrier form on Fowler–Nordheim plot analysis," *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena,* vol. 31, no. 3, p. 032201, 2013.

[92] X. Lian *et al.*, "Quantum point contact model of filamentary conduction in resistive switching memories," in *2012 13th International Conference on Ultimate Integration on Silicon (ULIS)*, 2012, pp. 101-104.

[93] Y. Wang *et al.*, "Memory switching properties of e-beam evaporated SiOx on N++ Si substrate," *Applied Physics Letters,* vol. 100, no. 8, p. 083502, 2012.

[94] Y.-F. Chang *et al.*, "Study of polarity effect in SiOx-based resistive switching memory," *Applied Physics Letters,* vol. 101, no. 5, p. 052111, 2012.

[95] Y.-F. Chang *et al.*, "Demonstration of Synaptic Behaviors and Resistive Switching Characterizations by Proton Exchange Reactions in Silicon Oxide," *Scientific Reports,* vol. 6, p. 21268, 02/16 2016.

[96] Z. Wang *et al.*, "Electrochemical metallization switching with a platinum group metal in different oxides," *Nanoscale,* 10.1039/C6NR01085G 2016.

[97] S. M. Bishop, H. Bakhru, J. O. Capulong, and N. C. Cady, "Influence of the SET current on the resistive switching properties of tantalum oxide created by oxygen implantation," *Applied Physics Letters,* vol. 100, no. 14, pp. -, 2012.

[98] Keithley, Model 4200-SCS Semiconductor Characterization System Reference Manual, Keithley, ed., 2013. [Online]. Available.

[99] C. Hong-Yu, Y. Shimeng, G. Bin, H. Peng, K. Jinfeng, and H. S. P. Wong, "HfOx based vertical resistive random access memory for cost-effective 3D cross-point architecture without cell selector," in *Electron Devices Meeting (IEDM), 2012 IEEE International*, 2012, pp. 20.7.1-20.7.4.

[100] S. Mahadevan, S. M. Hardas, and G. Suryan, "Electrical breakdown in semiconductors," *physica status solidi (a),* vol. 8, no. 2, pp. 335-374, 1971.

[101] G. J. Leggett, "Scanning Probe Microscopy," in *Surface Analysis – The Principal Techniques*: John Wiley & Sons, Ltd, 2009, pp. 479-562.

[102] Z. G. David, E.-S. Al-Moatasem, and L. S. Alexander, "A mechanism for Frenkel defect creation in amorphous SiO 2 facilitated by electron injection," *Nanotechnology,* vol. 27, no. 50, p. 505207, 2016.

[103] M. Anpo, M. Che, B. Fubini, E. Garrone, E. Giamello, and M. C. Paganini, "Generation of superoxide ions at oxide surfaces," *Topics in Catalysis,* journal article vol. 8, no. 3, pp. 189-198, 1999.

[104] M. A. Schoonen, C. A. Cohn, E. Roemer, R. Laffers, S. R. Simon, and T. O Riordan, "Mineral-induced formation of reactive oxygen species," *Reviews in mineralogy and geochemistry,* vol. 64, p. 179, 2006.

[105] B. Fubini and A. Hubbard, "Reactive oxygen species (ROS) and reactive nitrogen species (RNS) generation by silica in inflammation and fibrosis," *Free Radical Biology and Medicine,* vol. 34, no. 12, pp. 1507-1516, 2003.

[106] T. Chang, S.-H. Jo, and W. Lu, "Short-Term Memory to Long-Term Memory Transition in a Nanoscale Memristor," *ACS Nano,* vol. 5, no. 9, pp. 7669-7676, 2011/09/27 2011.

[107] M. Prezioso, F. Merrikh-Bayat, B. D. Hoskins, G. C. Adam, K. K. Likharev, and D. B. Strukov, "Training and operation of an integrated neuromorphic network based on metal-oxide memristors," *Nature,* Letter vol. 521, no. 7550, pp. 61-64, 05/07/print 2015.

[108] Z. Wang, S. Ambrogio, S. Balatti, and D. Ielmini, "A 2-transistor/1-resistor artificial synapse capable of communication and stochastic learning in neuromorphic systems," (in English), *Frontiers in Neuroscience,* Original Research vol. 8, no. 438, 2015-January-15 2015.

[109] A. Pascual-Leone *et al.*, "Characterizing Brain Cortical Plasticity and Network Dynamics Across the Age-Span in Health and Disease with TMS-EEG and TMS-fMRI," *Brain Topography,* vol. 24, no. 3, p. 302, 2011// 2011.

[110] T. A. Keller and M. A. Just, "Structural and functional neuroplasticity in human learning of spatial routes," *NeuroImage,* vol. 125, pp. 256-266, 1/15/ 2016.

[111] R. J. Chater, "Oxygen self-diffusion and surface exchange measurements of the polycrystalline oxide (La0.8Sr0.2)(Cr0.2Fe0.8)O3-d using Secondary Ion Mass Spectrometry," PhD, Department of Materials, Imperial College London, 2014.

[112] R. D. Bayliss, S. N. Cook, S. Kotsantonis, R. J. Chater, and J. A. Kilner, "Oxygen Ion Diffusion and Surface Exchange Properties of the α- and δ-phases of Bi2O3," *Advanced Energy Materials,* vol. 4, no. 10, pp. n/a-n/a, 2014, Art. no. 1301575.

[113] M. Kubicek, Z. Cai, W. Ma, B. Yildiz, H. Hutter, and J. Fleig, "Tensile Lattice Strain Accelerates Oxygen Surface Exchange and Diffusion in La1−xSrxCoO3−δ Thin Films," *ACS Nano,* vol. 7, no. 4, pp. 3276-3286, 2013/04/23 2013.

[114] J. A. Kilner, S. J. Skinner, and H. H. Brongersma, "The isotope exchange depth profiling (IEDP) technique using SIMS and LEIS," *Journal of Solid State Electrochemistry,* journal article vol. 15, no. 5, pp. 861-876, 2011.

[115] R. J. Chater, S. Carter, J. A. Kilner, and B. C. H. Steele, "Development of a novel SIMS technique for oxygen self-diffusion and surface exchange coefficient measurements in oxides of high diffusivity," *Solid State Ionics,* vol. 53, pp. 859-867, 1992/07/01 1992.

[116] M. G. M. Uematsu, K. M. Itoh, "Oxygen Self-Diffusion in Silicon Dioxide: Effect of the Si/SiO2 Interface," *Defect and Diffusion Forum,* vol. Vols. 258-260, pp. pp. 554-561, 2006.

[117] Tektronix, "MDO4000 Series Mixed Domain Oscilloscopes User Manual," ed.

[118] Keithley, "Model 6430 Sub-Femtoamp Remote SourceMeter Instruction Manual," ed.

[119] C. U. B. T. Dingle, D.B. Cook, D.D. Do, R.L. Gerlach, J-Z Li, J.Rasmussen, "The Integration of a High Performance Quadupole SIMS Facility with a Ga^+ LMIS based FIB Instrument.," in *Secondary Ion Mass Spectrometry*, University of Muenster, Menster, Germany, 1995.

[120] R. J. Elsey. Outgassing of vacuum materials-l. *The Theory and Practice of Vacuum Science and Technology in Schools and Colleges.*

[121] R. J. Elsey. Outgassing of vacuum materials-II. *The Theory and Practice of Vacuum Science and Technology in Schools and Colleges.*

[122] A. Brudnik, M. Bućko, M. Radecka, A. Trenczek-Zając, and K. Zakrzewska, "Microstructure and optical properties of photoactive TiO2:N thin films," *Vacuum,* vol. 82, no. 10, pp. 936-941, 6/3/ 2008.

[123] D. McPhail and M. Dowsett, "Dynamic SIMS," in *Surface Analysis – The Principal Techniques*: John Wiley & Sons, Ltd, 2009, pp. 207-268.

[124] J. C. Vickerman, "Molecular Surface Mass Spectrometry by SIMS," in *Surface Analysis – The Principal Techniques*: John Wiley & Sons, Ltd, 2009, pp. 113-205.

[125] J. A. Kilner, B. C. H. Steele, and L. Ilkov, "Oxygen self-diffusion studies using negative-ion secondary ion mass spectrometry (SIMS)," *Solid State Ionics,* vol. 12, pp. 89-97, 1984/03/01 1984.

[126] IonTOF, "ToF.SIMS5," ed. https://www.iontof.com/download/IONTOF_TOF-SIMS_5_Brochure.pdf.