Formation and Evolution of Conductive Filament in TaO_x Resistive Random-Access Memory

Submitted in partial fulfillment of the requirements for

The degree of

Doctor of Philosophy

in

Materials Science and Engineering

Yuanzhi Ma

B.S., Materials Physics, Jilin University

M.S., Materials Science, Carnegie Mellon University

Carnegie Mellon University

Pittsburgh, PA

December 2019

Acknowledgements

First of all, I would like to thank my advisor Prof. Marek Skowronski, who taught me how to think independently and provided me guidance when I was lost. He helped me realize the necessity of taking the ownership of my projects, which greatly improved my research planning ability, awareness to utilize all resources, and to always consider the “unknown unknown”. Additionally, I would like to thank Prof. James Bain, who is one of my committee members, for providing me with a lot of insight throughout my graduate school. I would also like to thank the rest of my committee members Prof. Paul Salvador and Prof. Daniele Ielmini, who have helped me see my projects from a more comprehensive point of view.

During my time at Carnegie Mellon University, I have received a tremendous amount of help and advice from the staff members and my colleagues, including but not limited to: Noel Thomas Nuhfer, Betsy Clark, Adam Wise, Dan Flaherty, Norman Gottron, Mason Risley, Marygrace Antkowski, Suzanne Smith, Darshil Gala, Phoebe Yeoh, Jonathan Goodwill, Dasheng Li, Jonghan Kwon, Abhishek Sharma, Qiyun Xu, Yiqi Yu, Bingyuan Zhao, Jingjia Meng, Peilin Tian, Liting Shen, Yuezhang Zou, and Luke Lyle. I have also collaborated with the staff members from Oak Ridge National Laboratory and National Institute of Standards and Technology, who have greatly contributed to my research projects: David Cullen, Jonathan Poplawsky, Brian Sneed, Karren More, and Andrew Herzing.

Moreover, my family and friends have been nothing but supportive throughout graduate school. I would like to specially thank my parents Yongwu Ma and Shuqing Ma, who have always trusted me and supported me no matter what obstacle I was faced with. I would also like to thank all my friends for the help I received during hard times and the joy from good times, including but not limited to: Youjin Nam, Runke Liao, Guannan Tang, Hokon Kim, Yi Shi, Alice Perrin, Fanchen Kong, Yang Yu, Yu Zhao, Bing Zhou, Enbo Zhang, Yang Liu, Cheng-Ming Chow, Tong Mo, Zhao Lu, Xiaoting Zhong, Yining He, Han Wang, Nan Gao, Yuwei Qin, Tong Mo, Dai Tang, Yu Lin, Jun Wang, Humphrey Hu, Tim Hsu, Elizabeth Mao, Tyler Martin, Vineetha

Bheemarasetty, Rubayyat Mahbub, Saransh, Maxwell Li, Ke-Wei Jin, Yao-Hsuan Tseng, Li-Yang Chang, Jaejun Lee, and Christian Kurniawan.

Finally, this work was made possible by several funding sources. This work was in part supported by FAME, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA, by NSF Grant DMR 1409068, NSF Grant DMR-1905648, and Data Storage Systems Center at Carnegie Mellon University. This work made use of deposition facilities provided by ONR DURIP equipment Grant #N000141310874, the Materials Characterization Facility at Carnegie Mellon University supported by grant MCF-677785, the microscopy facilities at Oak Ridge National Laboratory's Center for Nanophase Materials Sciences, which is a U.S. Department of Energy, Office of Science User Facility, and the instrumentation provided by the U.S. DOE Office of Nuclear Energy, Fuel Cycle R&D Program, and the Nuclear Science User Facilities.

Abstract

The increasing demand on memory from the next-generation technologies facilitated the pathfinding and development of emerging memories, among which Resistive Random-Access Memory (RRAM) is one of the most competitive options and least understood. Many attempts have been made to understand the resistive switching phenomena in metal-oxide RRAM, most of which suffered artifacts introduced from the testing or characterization processes. In my PhD work, I selected TiN/TaO_2/TiN-based RRAM to understand the mechanism of electroformation, resistive switching, filament evolution, and endurance failure of such memory cells using electrical characterization and electron microscopy techniques.

The findings of my thesis work indicate that the behavior of TaO_2-based RRAM is a mixture of both Valence Change Model (VCM) and Thermochemical Model (TCM). During electroformation, the Ta moves laterally towards the hot spot and down the direction of the electric field, whereas the O moves up the electric field. The motion of both Ta and O results in the $TaO_{0.4}$ filament core and Ta_2O_5 gap after forming, corresponding to the high resistance state (HRS). The localized heating during forming also induces temperature activated ionic interdiffusion of O and Ti across the interfaces. The resistive switching is induced by the electric field applied across the device leading to Ta-rich sub-filaments forming and breaking the connection between the filament core and the electrode. Repeated resistive switching provides a virtual annealing in the conductive filament, causing the material in proximity of the filament to phase separate into metallic Ta and Ta_2O_5. The phase separation continues if provided with longer electrical stressing and will result in SET failure due to the Ta particles isolated by the Ta_2O_5 matrix. The intrinsic cycle-to-cycle variability of metal-oxide RRAM causing stochastic sub-filament overgrowth can lead the device to RESET failure. Endurance can be potentially improved by reducing interdiffusion and controlling the temperature in the memory cell during programming operations.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Emerging Non-Volatile Memory (NVM)

Information storage and transfer are essential to the technology in our everyday life, and will be more and more relied on as technology advances. In the past decades, it was mainly needed for the memory to gain higher capacity, reaction speed and reduced cost in conventional Von-Neumann computer architecture (levels of cache by static random-access memory (SRAM), random-access memory (RAM) by dynamic random-access memory (DRAM), and hard disk drive (HDD) etc.). Recently, with the emerging technologies such as Internet of Things (IoT), bio-inspired computing, and in-memory computing, it became urgent to explore beyond the existing memory solutions due to the difficulty of efficiently further increasing their performance as well as the data transfer speed between the levels of the memory hierarchy [1].

In recent years, the implementation of 3D NAND products as solid-state drive has made significant progress and has very promising outlook in the market share for information storage versus its opponent HDD due to its much shorter read and write times and potential of further increasing its capacity by increasing number of layers in the vertical dimension. As the 3D NAND related technology becomes more mature, drawbacks of such application have started to appear as well. As the number of layers increases (up to 96 layers as of 2018), the requirement for high aspect ratio via directly results in larger feature size in the fabrication, and the area reserved for wordline staircase also becomes larger, which further decreases the cell efficiency (the percentage of the area taken by memory cells on the whole die), etc. [2]. Therefore, as the advancement of 3D NAND persists for the next decade, other potential approaches should also be considered in order to meet the demands in the memory industry for the coming decades. The recent commercialization of 3D Crosspoint (speculated to have adopted phase-change memory (PCM)) [3]–[5] and the demonstration of spin torque transfer (STT) random-access memory (RAM) [6] show the candidates for the alternative solutions for information storage.

On the other hand, the emerging application of neuromorphic computing and in-memory computing could also use newer structures dedicated for parallel computing [1], [7]. Yu *et al*. [8] demonstrated binary neural network (BNN) on 16 Mb resistive random-access memory (RRAM) macro-chip for the classification of MNIST handwritten dataset, which achieved almost the same accuracy as the ideal software implementation. In a more recent study, Burr *et al*. [9] implemented PCM based hardware arrays generalized for a number of common machine-learning tasks and achieved accuracies equivalent to those of software-based training but at two orders of magnitude lower energy consumption than those of today's graphical processing units (GPU).

Emerging non-volatile memory (NVM) not only has potential of replacing existing memories and being implemented in brain-inspired computing as mentioned above, but also has been considered promising as storage class memory to reduce the complexity of the currently available memory-storage hierarchy, and low-power computing solutions for applications that are more limited on power supply such as IoT [10].

Clearly, the research on NVM has been developing rapidly and shown potential of being implemented in many next-generation technologies. While the operation mechanism of the prototypical devices such as PCM and STT-RAM has been more sufficiently studied and understood, same thing cannot be said for resistive random access-memory (RRAM), albeit its high scalability and stackability, retention, endurance, and compatibility with the complementary metal-oxide semiconductor (CMOS) technology [11]–[15]. More importantly, the resistance of RRAM can be adjusted continuously and used for storing more than binary information, which could be potentially advantageous over other emerging NVM for analog computing [1], [16], [17]. Therefore, it is important to understand the operation mechanism of metal-oxide based RRAM.

1.2 Electrical Characteristics of Metal Oxide Based RRAM

Schematic diagram of a RRAM cell is shown in Figure 1.1. The device has a metal-insulator-metal (MIM) sandwich structure, where the two metal layers serve as electrodes, and the metal-oxide insulator serves as the functional layer.

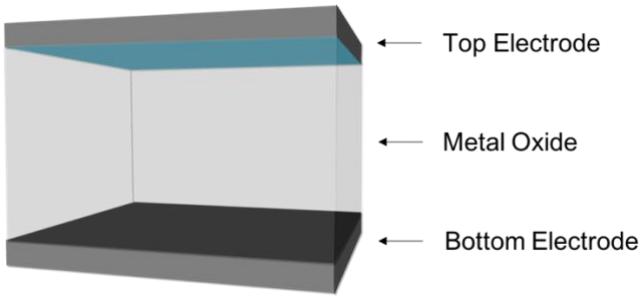

Figure 1.1 A schematic illustration of a resistive switching memory cell.

The as-fabricated devices are usually very resistive (> 1 MΩ), and require a one-time conditioning procedure, referred to as electroforming or forming, which brings the resistance of the device down to 10-100 kΩ before it can start resistive switching. Goodwill *et al*. [18] argued that as-deposited TaO_x based metal-insulator-metal structures exhibit a positive feedback loop between the Joule heating and the electrical conductivity and undergo a thermal runaway during the voltage sweep used for electroformation. During such runaway, the device resistance rapidly drops, which appears as a "snap" in the device *I-V* (dashed part of blue curve, Figure 1.2 (a)). This process is known as "threshold switching", which is volatile and does not induce permanent changes in the memory cell. The threshold switching can induce instantaneous localized high temperature in the device which adds instability to the system therefore should be avoided if possible. By selecting the value of load resistance (R_{LOAD}) so that the load line has only one intersection with the intrinsic device *I-V*, the threshold switching event can be eliminated, the evolution of current can be gradual, and the entire S-type negative differential resistance (S-NDR) characteristics (black curve, Figure 1.2 (a)) can be accessed experimentally. The RRAM *I-V* characteristics shown in Figure 1.2 (b) were collected with R_{LOAD} = 58 kΩ to reveal the full S-NDR by avoiding the threshold switching, and also to reduce the current overshoot due to discharge of parasitic capacitances during the process [19]–[21].

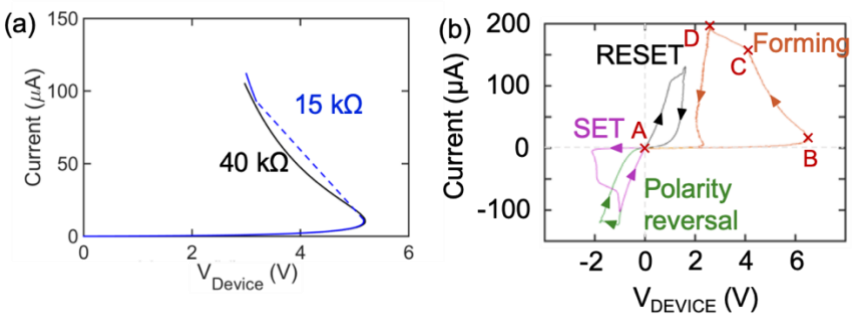

Figure 1.2 (a) The S-type Negative differential resistance (S-NDR) (black), and threshold switching (blue) characteristics of RRAM cells. (b) *I-V* characteristics of an RRAM cell, including forming (orange), SET (purple), RESET (black), and polarity reversal (green). The four crosses marked with "A", "B", "C", and "D" denote the turning points in the forming *I-V*. A load resistor of 58 kΩ was used during the *I-V* sweep.

Figure 1.2 (b) (orange trace) shows a typical forming *I-V* of an RRAM memory cell. As can be seen, the as-fabricated structures exhibit an S-type negative differential resistance (S-NDR)) extending between points A and C. In this range, there are no permanent changes of the device structure or electrical properties and the *I-V* can be retraced back and forth many times. The forming occurs when the current goes beyond point C when the device resistance decreases with time once the process has started. Upon reaching point D, the time evolution stops and the D→A part of the characteristics could be retraced repeatedly. However, the A→B→C route cannot be accessed anymore because at this stage, the device resistance is permanently reduced.

A typical set of resistive switching characteristics following formation are shown in Figure 1.2 (b), where the device switches from high resistance state (HRS) to low resistance state (LRS) under applied negative bias (purple trace marked SET), and switches from HRS back to LRS when under positive bias (RESET, black trace). When the SET and RESET operations occur at two opposite voltage polarities, the device exhibits bipolar switching characteristics, and when both SET and RESET occur in the same polarity, the switching is unipolar. Clearly, the switching *I-V* from Figure 1.2 (b) is bipolar. In bipolar resistive switching devices, if a memory cell at LRS continues to be biased at a higher voltage, the resistance can switch back to HRS, which is known as the polarity reversal, or complementary switching (Figure 2 (b), green) [22], [23]. After polarity reversal, the bias polarities for subsequent resistive switching will be opposite to the original switching bias polarities.

1.3 Currently Existing Models for Resistive Switching

1) Conductive Bridge Memory Model

The first model involves memory cells using electrochemically active metal electrodes (typically Ag or Cu) sandwiching the solid electrolyte (metal oxide or chalcogenide) [24], [25]. During operation, the metal ions diffuse into the electrolyte under the effect of electric field applied across the device, forming a “conductive bridge” connecting the two electrodes. The bridge consists of metallic Cu or Ag small diameter filament. Reversing the bias polarity leads to removal of metal ions from the electrolyte and appearance of a gap in th filament. This corresponds to switching between high and low resistance states (HRS and LRS), as shown in Figure 1.3. This model is called conductive bridge memory (CBM) [26]. Since the metal ions are positively charged, they move in opposite directions when the device is applied with opposite biases, resulting in the connection and break of the filament between the electrodes upon the application of alternate voltage polarities, which makes CBM a bipolar resistive switching memory.

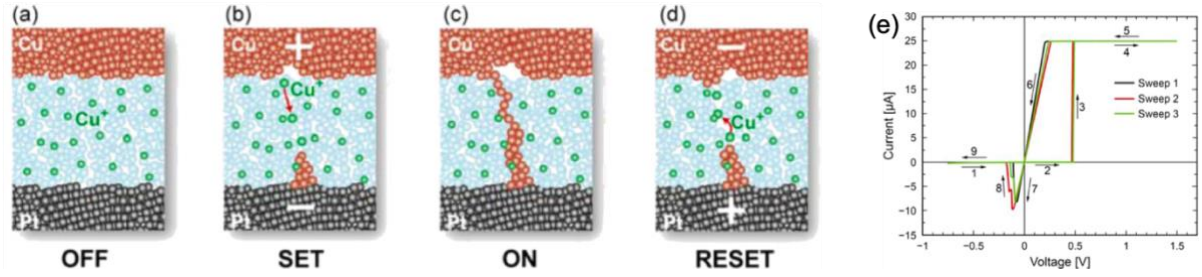

Figure 1.3 Schematic presentation of the operation principles of cation based ReRAM memory cells with a Cu active electrode and Pt counter electrode. (a) depicts the initial or OFF state. (b) shows the dissolution of metallic Cu to Cu+ ions and the formation of the metallic filament in the *‘*program*’* or *‘*set*’* operation. (c) is the short-circuited ON state. (d) shows the dissolution of the filament in the *‘*erase*’* or *‘*reset*’* step under reverse bias and the return to the OFF state. (e) *I-V* plot for a Cu-doped SiO_2 device with a current compliance of 25 μA. *(a-d) are reprinted from* [27]*, and (e) is reprinted from* [28].

2) Valence Change Memory Model

The second model is known as “valence change memory (VCM)” [26], where the functional layer in the device is typically transition-metal oxide such as TaO_x, HfO_x, or TiO_x [14], [29], [30]. Jeong *et al.* [31] postulated that initially high device resistance is lowered by the oxygen leaving the functional oxide and

entering the anode due to the potential difference created by applied bias. Oxygen vacancies ($V_O^{\bullet\bullet}$) generated in the process act as donors in the functional layer and make it conductive. It is not clear why this process occurs only locally, but it is generally agreed that the result is a small diameter filament that connects the two electrodes. When forming the filament, O ions within the filament was argued to leave the oxide and diffuse into the electrode [32]–[34] or form oxygen bubbles at the interface [35], [36]. Figure 1.4 shows a schematic illustration of the VCM switching mechanism, where the O^{2-}/ $V_O^{\bullet\bullet}$ drift in the direction of the electric field in the functional layer. The filament can either be continuous corresponding to the LRS of the device or have a gap responsible for the HRS. Since in the VCM class of the devices the electric field effect dominates the ion motion, the switching is inherently bipolar with SET and RESET occurring in opposite polarities. Recently, Wedig *et al*. have suggested that in addition to O, metal ions are also mobile in materials systems that are generally accepted as operating by VCM mechanism (TaO_x, HfO_x, and TiO_x) [37].

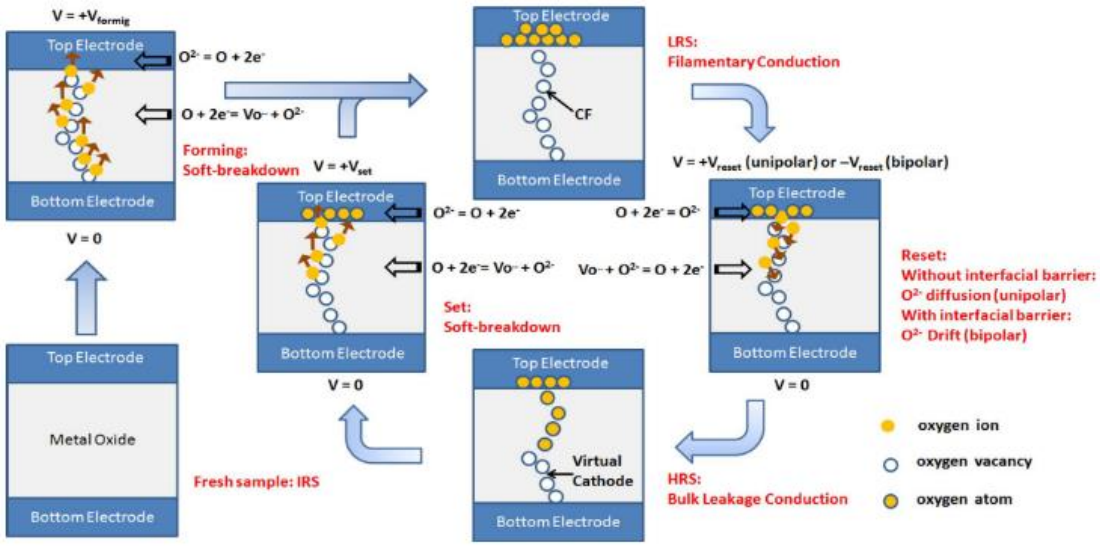

Figure 1.4 Schematic illustration of the switching process in the simple binary metal–oxide RRAM. Reprinted from [11].

3) Thermochemical Model

Unipolar switching, observed for example in TiO_2 [38] and NiO [39], [40], was explained by thermo-chemical model (TCM) that emphasized thermal effects [41], [42]. Electrical conductivity in most oxides is thermally

activated and as-fabricated metal/oxide/metal structures exhibit S-type Negative Differential Resistance (S-NDR) *I-V* characteristics [43], [44]. Such devices are unstable when biased into the NDR region exhibiting threshold switching events and bifurcation of current density into high and low current density domains. This leads to appearance of a small diameter hot spot in the device in which the temperature can easily exceed 1000 K [44]–[46]. Since at high temperatures, lower oxidation states of metal cations become energetically more stable as evidenced by Ellingham diagram [47], the appearance of the hot spot should lead to compositional change with the high temperature region becoming metal-rich and oxygen-poor. Such composition change can occur by diffusion of metal cations in the direction of temperature gradient or diffusion of O against it. Which one of these two mechanisms dominates remains a matter of dispute [48]–[51], but both have been reported. This effect describes the demixing of an alloy by an imposed temperature gradient is known as the Soret effect [52]. Since the temperature gradient associated with current constriction is mostly lateral, the motion of ions is expected to be mostly horizontal rather than vertical (along the direction of the electric field), which explains the unipolar *I-V* behavior (Figure 1.5 (c)).

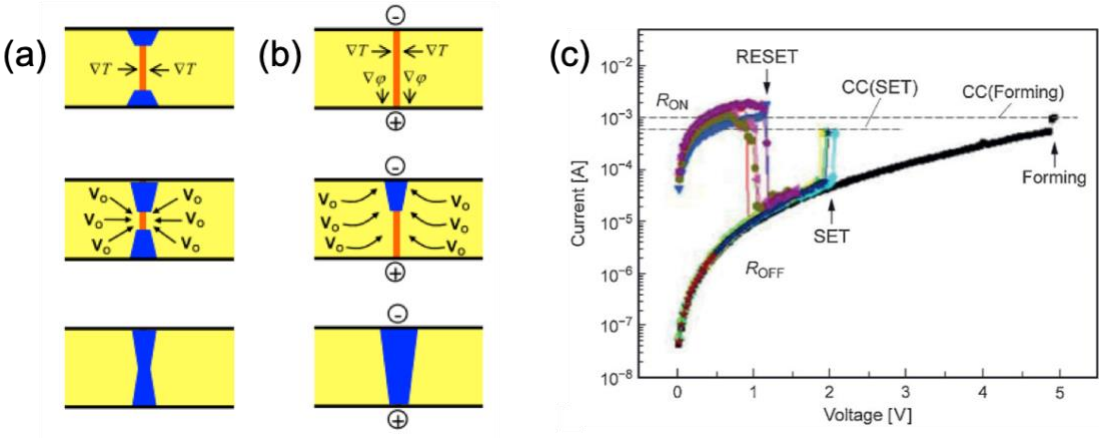

Figure 1.5 Unipolar SET switching schematic illustrations for the side view of the device stack with: (a) neutral vacancies (symmetric channel formation) and (b) positively charged vacancies (asymmetric channel influenced by drift). The process is initiated in both cases by a purely electrical (threshold) transition that creates a radial temperature gradient in the film due to Joule heating. The Soret force attracts vacancies toward the electrical filament, and if the vacancies are positively charged, they will be attracted toward the negative electrode. (c) Unipolar current-voltage characteristics of a Pt/NiO/Pt stack with a NiO film thickness of 50 nm and 100 µm x 100 µm Pt top electrodes. *(a) and (b) are reprinted from* [38]. *(c) is reprinted from reference* [40].

1.4 A Brief Introduction to Soret Effect (Thermophoresis)

In a multi-component material system, the material can build up a time-dependent concentration gradient in response to the imposed temperature gradient until stead state is reached (Figure 1.6) [52]. This effect is called Soret effect, also known as Ludwig-Soret effect, thermophoresis, thermal diffusion, or thermomigration.

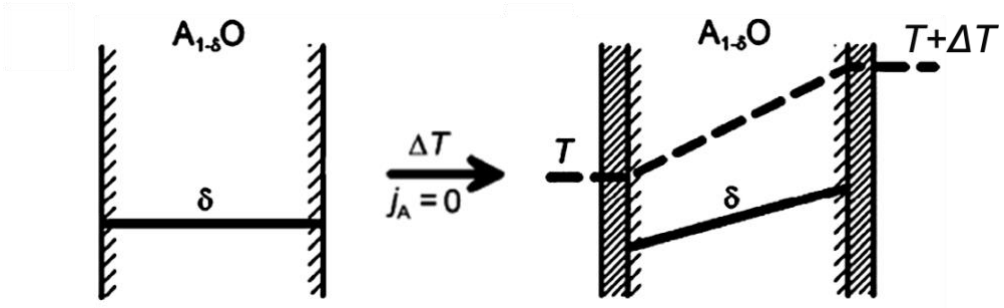

Figure 1.6 Illustration of Soret effect where a binary oxide forms a concentration gradient in response to the temperature gradient. δ stands for the deviation from stoichiometry. *Reprinted from* [53].

In experiments with the material exposed to a temperature gradient, Fick's law of diffusion needs to be modified to describe the flux of a mobile species k [52]:

$$J_k = -D_k \nabla c - D_k' c_k \nabla T \qquad \text{Equation 1.1}$$

where J_k is the number of moles of k crossing the yz plane per unit time in the direction of the positive x axis, D_k is isothermal diffusion coefficient of k, c_k is the molar concentration of k, D_k' is the thermal diffusion coefficient of k, and T is the absolute temperature. For steady state ($J_k = 0$), Equation 1.1 becomes [52]:

$$\frac{\partial \ln (c_k)}{\partial T} = -\frac{D_k'}{D_k} = -s_k \qquad \text{Equation 1.2}$$

where s_k is known as the Soret coefficient, which describes the extent of diffusion of species k in response to the temperature gradient.

The mechanism of Soret effect i.e. diffusion under temperature gradient is generally perceived as phonon assisted defect migration [54]. Consider a material system where phonon migration accounts for the major heat transport: as phonons move along the temperature gradient, the interaction between phonons and the mobile point defects such as vacancies or interstitials can pose a driving force on the defects in a certain direction. The direction in which the defects move as a result of the phonon-defect interaction is, however, not yet well documented and not straightforward to predict. It has been argued that for a species that travels

by vacancy migration, the direction of thermal diffusion is dictated by the relative magnitude of the vacancy formation energy h_f and the vacancy migration energy barrier h_v [55], [56]. If $h_v > h_f$, the net flux of vacancies will be flowing towards the higher temperature, meaning a concentration gradient will be built opposite to the temperature gradient; whereas for $h_v < h_f$, the concentration gradient will form in the same direction of the temperature gradient; and when $h_v = h_f$, no concentration gradient appears in response to the imposed temperature gradient. However, this argument itself has been argued to have underestimated the h_v and hence have overestimated the Soret effect [57]. The direction of Soret effect in materials systems with components diffusing interstitially have been argued to be even less intuitive and harder to predict [55].

Typically, the Soret flux J_k of component k in a system with diffusion based on the motion of vacancies of the same species can be described [52]:

$$J_k = -\left(\frac{c_k D_k (q_k^* - h_v)}{k_B T^2}\right) \cdot \nabla T \qquad \text{Equation 1.3}$$

where k_B is Boltzmann's constant, and q_k^* is known as the heat of transport whose magnitude and sign are unknown. For component k in a system with the Soret flux due to the migration of interstitials of the same species, it can be described as [52]:

$$J_k = -\left(\frac{c_k D_k (q_k^* + h_i)}{k_B T^2}\right) \cdot \nabla T \qquad \text{Equation 1.4}$$

where the h_i is the migration energy barrier of an interstitial.

To date, only $Cu_{2-\delta}O$ (cuprite) and $UO_{2+\delta}$ (uranium dioxide) have been investigated for Soret effect in binary oxide compounds [53], [58]. For the cuprite with depleted Cu, Timm *et al*. reported Soret effect by the diffusion of Cu vacancies and a concentration gradient of Cu opposite to the temperature gradient [53]. The Soret effect in $UO_{2+\delta}$ was reported to be based on O vacancies, however unfortunately the heat of transport of O vacancy migration was shown to be strongly dependent on composition therefore hard to predict the direction of the thermal diffusion [58].

1.5 Understanding RRAM: Previous Attempts and Challenges

Many reports have addressed the fabrication and electrical testing of resistive switching devices. Comparatively fewer studies used techniques such as transmission electron microscopy (TEM) [37], [59]–[63] or X-ray spectro-microscopy [48], [59], [64], [65] in order to image the filament to differentiate between existing models. Imaging studies face several challenges. One key challenge concerning all of the electron microscopy-based techniques is capturing the filament within the thin lamellae prepared by focused ion beam (FIB) milling or other sample preparation techniques. The thickness of the TEM specimen is typically 50 nm and this is the precision with which one has to locate the filament (possibly with a diameter smaller than 10 nm) for the cross-sectional sample. A common way to identify the filament location is to use the morphology changes of the top electrode caused by formation/switching as a fiduciary mark. Such features, including craters or electrode delamination, have been reported by many groups [66]–[68]. However, these features are not intrinsic to the filament formation or the switching process. They are the result of excessive power dissipation likely associated with capacitive discharge during forming [19], [20], [69], [70]. Such filaments can exhibit numerous artifacts not associated with filaments formed under appropriate “gentle” conditions.

Some groups have used plan view geometry for imaging and fabricated their devices on electron and X-ray transparent membranes [48], [59]. This approach allows for inspection of the entire active area of the device, but at the very significant cost of reducing the rate of heat removal. Estimated thermal resistances of devices on Si_3N_4 membranes versus those on oxidized Si wafers differ by an order of magnitude [71]. Since the dissipated power in experiments performed on membranes was less than tenfold than that used for SiO_2/Si structures [64], [72]–[74], it is likely that the devices on membranes experienced much higher temperature excursions that could lead to high temperature phenomena not intrinsic to the formation/switching processes. The same challenge of adequate heat-sinking and possible excessive temperature excursions appears in all *in situ* TEM experiments, in which bias is applied to the device located within a thin TEM lamellae [30], [75], [76].

It has been generally accepted that the electrodes such as Ag and Cu exhibit high diffusivities in the functional material forming the conductive filament and thus responsible for the switching events.

However, this cannot explain the switching mechanism in the devices without highly diffusive electrode materials. To date, few studies have reported the artifact-free visualization of the conductive filament without jeopardizing the normal working environment of a device, which leads to interpretations of non-intrinsic effects of the operations in the devices. Moreover, the understanding of the resistive switching in oxide-based devices is supported by even less experimental evidence that is free of experimental artifacts. Most of the work investigating the working mechanism of RRAM cells by microscopy methods applied in-situ technique, where the devices are being programmed as they are observed in the microscopes [30], [75], [76]. These works reported to some extent do reveal the physics of the resistive switching behavior, but also has significant limitation because the original thermal environment of the devices during programming was sacrificed for the visibility in the microscopes during the in-situ experiments. In the in-situ experiments, the devices were almost completely thermally isolated, whereas those in real working condition are well thermally grounded, which can introduce unwanted artifacts especially if the resistive switching behavior is affected by temperature. The author of this thesis found only one report focused on microscopy analysis on the evolution of the filament during endurance cycling, and device failure at the end of endurance due to the in-diffusion of the Cu from the metal contacts [77]. Using microscopy techniques to understand the intrinsic physics behind the formation, switching, evolution, and the failure mechanism of the metal-oxide resistive switching device is of great importance in the on-going process of implementing this application into the next-generation technologies.

1.6 Materials Selection

In this work, substoichiometric amorphous TaO_x was selected as the model material system for its desired behavior in resistance, electrical characteristics, and high visibility in TEM. The composition of TaO_x can be controlled by adjusting the O_2 flow during the reactive sputtering and therefore control its electrical conductivity in order to obtain the devices with optimal electrical characteristics. In addition, MIM structures using TaO_x as the insulating layer have shown both threshold switching and memory switching [18], [78], which helps this work understand both processes with one batch of devices. Finally, the atomic number of

Ta is greatly larger than that of O, which helps lower the difficulty interpreting the TEM data, specifically in the imaging mode where the contrast is highly dependent on the atomic number of the sample (Z-contrast), such as High-Angle Annular Dark-Field (HAADF) in Scanning Transmission Electron Microscopy (STEM). Other candidates with similar electrical behavior have also been considered, such as HfO_x, VO_x, and NbO_x, etc., but they all suffered from the lack of one or more advantages with TaO_x mentioned above, therefore were not chosen as the material system for this work [79]–[83].

TiN was selected as the material for electrodes because it is less reactive and easy to etch for the fabrication of the device. It is also less likely to diffuse into the functional layer compared with copper or silver, eliminating the effect of CBM from this study. Moreover, the huge difference between the atomic numbers of Ta and Ti makes it possible to interpret the Z-contrast using HAADF images – the contrast within the structure is mostly caused by Ta rather than any other element present in the entire structure (O, Ti, N, Si).

1.7 Thesis Outline

The work reported in this thesis will focus on understanding the electroformation, resistive switching, device evolution and failure mechanism by relating the electrical characteristics of RRAM cells with the structure and composition of the conductive filament via TEM/STEM (S/TEM) analysis.

The research will first set out to inspect the conductive filament post-forming in plan-view by S/TEM and identify the lateral driving force for forming (Chapter 2). The S/TEM investigation continues in Chapter 3 in cross-sectional geometry to study the vertical structure of the conductive filament, the vertical driving forces during forming, and also the mechanism of resistive switching. Chapter 4 is focused on electrical characterization to study the dynamics of electroformation and extraction of activation energy of this process. In Chapter 5, by inspecting the distribution of elements in the device post forming, we reveal the thermally activated ionic interdiffusion and re-evaluated the switching model of TaO_x-based RRAM. Finally, TEM analysis of devices after repeated switching leads to a better understanding of the evolution of the conductive filament.

Chapter 2 Formation of the Conducting Filament in TaOx Resistive Switching Devices by Thermal-Gradient-Induced Cation Accumulation

2.1 Introduction

This chapter presents the TEM analysis of physical changes in the functional layer of TaOx-based devices induced by the electroformation process, which were formed with dissipated power below 50 μW and arguably free of artifacts from excessive heating. The author argues that the first step in the electro-formation is a thermal runaway leading to spontaneous current constriction and appearance of a small hot spot. This process is referred to as threshold switching and is volatile i.e. does not induce any physical changes in the device structure. Only after a long pulse, the diffusion will lead to the permanent changes. The analytical TEM experiments have been performed in the plan-view geometry in order to monitor the lateral redistribution of ions. I have detected local accumulation of metal with the metal-poor ring around it, providing a strong evidence of metal-ion transport via the Soret effect. This observation does not exclude possibility of the vertical motion of either oxygen or tantalum as the geometry employed did not allow for observation of motion along the direction of the electric field.

2.2 Experimental

1) Device fabrication

The devices used in studies described in this chapter have a simple crossbar structure (Figure 2.1 (a)) and were fabricated on Si substrates covered by 1 μm thick thermal SiO_2 layer. The bottom electrode (400 nm in width) was defined using electron-beam lithography, followed by sputter deposition of 5 nm Ta and 20 nm TiN and the lift-off process. The 120 nm of functional TaO_x was deposited by reactive sputtering using Ta target and a mixture of oxygen (4 sccm) and argon (56 sccm) gas at chamber pressure of 3 mTorr. The

2 nm Ta/20 nm TiN top electrode layer (370 nm in width) was defined by electron-beam lithography and lift-off. The SEM top view of a device was shown in Figure 2.1 (b).

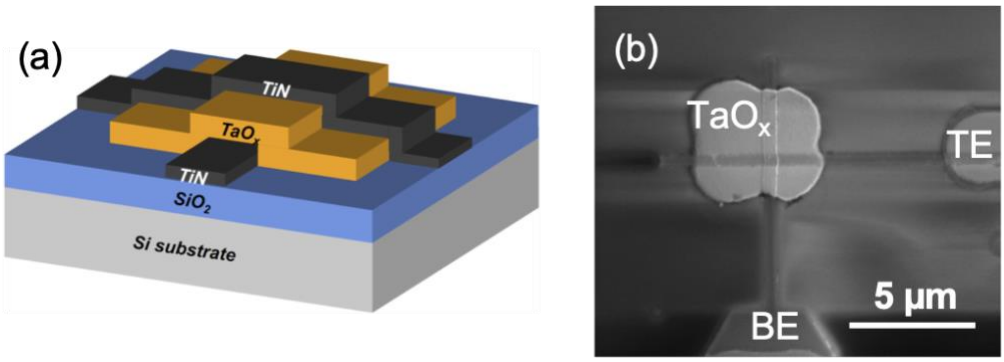

Figure 2.1 (a) An illustration of the cross-bar structured device using TiN as the electrodes and TaOx as the functional layer. (b) An SEM top view of the device. BE: bottom electrode; TE: top electrode.

2) Electrical testing

The devices were tested using a quasi-DC voltage source (Agilent 4155C Semiconductor Parameter Analyzer) with a load resistor of 1 MΩ connected in series with the device. No external current compliance was used. The devices were first electroformed with some of them switched to either low resistance or high resistance states. All samples show the same microstructure in the TEM/STEM images (in a plan view geometry). Examples of electroformation and resistive switching *I-V* characteristics are shown in Figure 2.2 (a) and (b), respectively. Forming currents for the device used in Figure 2.4, Figure 2.5 (a) and Figure 2.5 (b) were 20 μA, 21 μA and 15 μA, respectively.

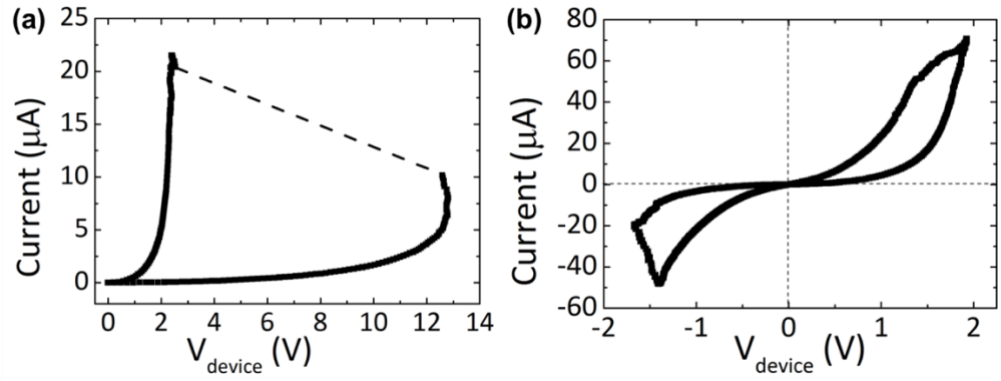

Figure 2.2 Typical *I-V* characteristics of (a) electroforming and (b) resistive switching.

3) Microstructural Analysis

The devices were lifted-out and thinned down by focused ion beam (FIB) and characterized by TEM/STEM. Large slabs (80 × 60 × 10 μm^3) of material containing devices of interest were first lifted-out in plan-view with Xe^+ plasma focused ion beam FIB (FEI Helios) and thinned down to electron transparency using Ga+ FIB (FEI NOVA Nanolab 600). In order to leave the free-standing TaO_x functional layer for TEM analysis, the material beneath the functional layer (Si substrate, 1 µm SiO_2 insulation layer, 5 nm Ta adhesion layer, and 20 nm TiN bottom electrode) was first removed, then the top electrode on the other side of the lamella was removed, both using Ga^+ FIB at 30 keV, 1 nA-50 pA, 89˚-90˚ incident angle. Afterwards, Ga^+ FIB at 5 keV, 70 pA, 85˚ incident angle was used to polish the side walls of the FIB lamella on both sides to remove the surface damage. It has been reported that such polishing condition limits the sidewall surface damage thickness to 2.5 nm, and the ion beam heating on the sample is less than 2 ˚C [84]–[86]. The final thickness of the region of interest in the FIB lamella was measured to be ~50 nm by SEM and log-ratio measurement in energy-filtered TEM. Given the above, it is safe to assume the TEM analysis on the as-prepared FIB lamella reveals mostly the intrinsic properties of the device rather than the artifacts introduced by FIB.

The HAADF images in Figure 2.4 were taken using FEI Titan 80-300 operating at 300 keV; the HAADF and XEDS analysis in Figure 2.5(a) and (b) were performed using FEI Talos F200X at 200 keV. The HAADF images were collected with an angle higher than 50 mrad to ensure exclusive Z-contrast in the images. The EELS O-K map in Figure 2.6(c) was extracted from EELS spectrum image at 532-572 eV. The EELS spectrum image was taken using Hitachi HF3300 S/TEM operating at 300 keV.

4) Electrothermal simulation

The heat and charge flow in devices were simulated using COMSOL Multiphysics finite element software package. Figure 2.3 shows the structure used in the simulation. The nominally 400×400 nm^2 square crossbar devices were modeled as a disc consisting of bottom electrode and the oxide with the diameter of 3 μm on 10 μm diameter disc of SiO_2/Si substrate. The active area of the device was represented by 400 nm diameter top electrode (Figure 2.3). The 5 µm thick and 10 µm diameter silicon disc is sufficient to mimic

the heat sinking of the real 500 μm thick Si substrate. In the sandwich structure, the bottom electrode is 5 nm Ta/20 nm TiN, and top electrode is 2 nm Ta/20 nm TiN.

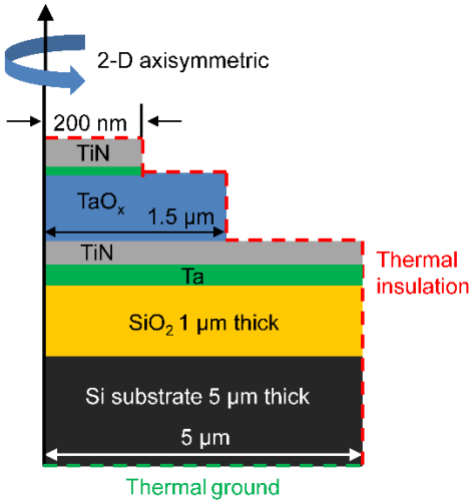

Figure 2.3 Schematic diagram of the device structure used for electro-thermal simulation.

The temperature distribution and electrical conductivity are solutions of coupled equations describing heat and charge flow. The temperature distribution T(x,y,z,t) in the device is the solution of the heat flow equation:

$$\rho C_p \frac{\partial T(x,y,z,t)}{\partial t} - \kappa \Delta T(x,y,z,t) = \frac{J^2(x,y,z,t)}{\sigma(T)} \qquad \text{Equation 2.1}$$

where ρ is the material density, C_p is the heat capacity, κ is heat conductivity, J is current density, σ is the electrical conductivity and Δ is the Laplace operator. All the material parameters used in the simulation are listed in Table 2.1.

The time-dependent electrical current flow J_e is described by combining the conservation of currents equation and Gauss's Law:

$$\nabla \left(\boldsymbol{J_e}(x,y,z,t) + \frac{\partial (\varepsilon_0 \varepsilon_r \boldsymbol{E}(x,y,z,t))}{\partial t} \right) = 0 \qquad \text{Equation 2.2}$$

Where ε_0 is the permittivity of vacuum, ε_r is the relative permittivity, and **E** is the electric field in the materials. The σ in the equation is a characteristic of the functional material and was modelled by three-dimensional Poole-Frenkel formula [18]:

$$\sigma_{PF}(E,T) = \frac{\sigma_0(T)}{E} \left(\frac{k_B T}{\beta}\right)^2 \left\{ 1 + \left(\frac{\beta\sqrt{E}}{k_B T} - 1 \right) \exp\left(\frac{\beta\sqrt{E}}{k_B T} \right) \right\} + \frac{\sigma_0(T)}{2} \qquad \text{Equation 2.3}$$

where

$$\sigma_0(T) = q\mu N_c(\frac{N_d}{N_t})^2 exp\left(-\frac{E_d+E_t}{2k_BT}\right) \qquad \beta = \left(\frac{q^3}{\pi\varepsilon_0\varepsilon_i}\right)^{\frac{1}{2}}$$ Equation 2.4

k_B is the Boltzmann constant, q is the electron charge, ε_0 is the permittivity of free space, ε_i is the relative dielectric constant of TaO_x, E is the electric field (device voltage/oxide thickness), T is the temperature, N_c is the effective density of states in the conduction band, N_d, E_d, N_t and E_t correspond to the densities and ionization energies of donors and traps, respectively, $\sigma_0(T)$ is the low field conductivity, and μ is the electron mobility. The values of activation energy ($\frac{E_d+E_t}{2}$) and pre-exponential factor ($q\mu N_c(\frac{N_d}{N_t})^2$) used in this work were 0.35 eV and 1397 S/m by fitting the forming *I-V* with the experimental (Figure 2.7 (a)).

Material	Density ($kg \cdot m^{-3}$)	Thermal conductivity ($W \cdot m^{-1} \cdot K^{-1}$)	Electrical conductivity ($S \cdot m^{-1}$)	Heat capacity ($J \cdot kg^{-1} \cdot K^{-1}$)	Relative permittivity
Si	2329	130	N/A	700	N/A
SiO_2	2200	1.4	N/A	730	N/A
TiN	5210	5	4E5	545	4
TaO_x	8200	0.08	Described above	174	22

Table 2.1 Material parameters used in the simulation.

During the simulation, the temperature of the bottom surface of the Si was fixed at 300 K while the top and side surfaces of the structure were thermally insulated. Electrically, the bottom surface of the bottom electrode was at electrical ground, and the top surface of the top electrode was connected to an outside circuit with 1 MΩ load resistor and the voltage source. The amplitude of the applied source voltage was the same as during electrical testing, which was 22 V with a ramp rate of 4 V/s. It should be noted here that in Figure 2.7 (a) and 2.2 current is plotted as a function of device voltage (across the device) while the voltage mentioned here is the source voltage (across the device and the load resistor). A thermal boundary conductance of 100 MW/m^2K was included at the metal/oxide interfaces. The TaO_x sandwiched in between is 120 nm thick.

2.3 TEM/STEM Analysis of Post-Forming Devices

The investigated devices have relatively thick functional layers, which typically require high forming and switching voltages.[29], [87] This effect was alleviated by lowering the resistivity of the material through the reduction of the oxygen partial pressure during deposition of the functional layer. The forming voltage was

around 10 V for all devices examined in this report and all exhibited stable memory switching characteristics. Devices were formed and/or switched using a circuit with the 1 MΩ load resistor connected in series with the device and the voltage source (all *I-V*'s are plotted as a function of voltage across the device, not source voltage). This limited the current during rapid changes of device resistance to 20 μA - 60 μA [88].

The methodology used in this report relies on TEM imaging and spectroscopy as have several other studies [61], [89], [90]. However, two aspects of our experimental approach made it possible to reveal unreported structural characteristics of the filaments. First, the imaging was done in plan-view (*i.e.* with the electron beam parallel to both the direction of current flow and the filament length) on devices with relatively thick TaO_x layers (120 nm). This geometry maximizes the interaction length between the beam and the filament and increases signal to noise ratio for all imaging modalities compared with a cross-sectional view. Contrast was further enhanced by removing the top and bottom electrodes (TiN) and the SiO_2/Si substrate, leaving only the functional layer in the beam path. We have been able to examine the entire region of the functional oxide in 300 nm × 400 nm crossbar devices, which is clearly not possible in cross-sectional specimens. This geometry allows only for the elemental mapping of the lateral distributions. Accordingly, we comment only on the lateral motion mechanisms leaving the motion along the direction of the field for Chapter XXX. Secondly, leaving only the functional oxide makes it easy to characterize the material by the high angle annular dark field (HAADF) imaging mode. HAADF images are obtained by collecting electrons scattered at high angles by Rutherford scattering, which increases rapidly with the atomic number [91]. Since the atomic number of Ta is 73 while that of other elements present in the device is much lower (N:14; O:16; Ti:22), the intensity of the collected HAADF images was high and was mostly determined by the local tantalum content.

A low magnification, plan-view HAADF image of a device after formation and 10 switching cycles (forming current of 22 μA) is shown in Figure 2.4(a). All electrical testing was performed on-chip to assure good heat sinking of the device and preventing excessive temperature excursions expected in the thin STEM samples. The device was in the low resistance state of the memory switch. However, we have also tested devices just after forming, as well as after both forming and memory switching left at the high resistance state. We

have not noticed any significant differences in the atomic distribution in the filament. Apparently, the changes detected in plan-view geometry have been created during electro-formation and not affected by subsequent memory switching.

The overlaid dashed lines in Figure 2.4 (a) indicate the positions of the edges of electrodes in the crossbar structure. The contrast within the active device area is uniform with the exception of a bright, circular region close to the top edge. Closer inspection also reveals a larger, low intensity ring around the high intensity core (Figure 2.4 (b)). A single feature of this type was observed in every one of the seven electroformed and/or switched devices examined while it was not present in control as-fabricated sample. Accordingly, we attribute this contrast to the filament formation. None of the imaged devices displayed any additional contrast that could be interpreted as due to multiple and/or partial filaments. This is in agreement with results of multiple groups indicating formation of only one filament [76], [92]–[95].

Figure 2.4 (b) shows a high magnification image of the same filament. The diameter of the bright core of the filament is ≈ 65 nm while that of the dark ring is ≈ 110 nm. In addition to these features, both the core and ring show finer scale intensity changes. In the core, the lateral size of these features is ≈ (5 – 8) nm with size decreasing towards the periphery. The sample thickness is uniform within the active device region due to the sample preparation method with both sides polished by Ga beam. The contrast in the vicinity of the filament is, therefore, due to increased high angle scattering. The intensity changes imply that the core of the filament has increased Ta content while the ring is depleted of Ta compared to the initial film composition. The contrast in Figure 2.4 (b) was quantified by rotationally averaging the HAADF intensity (normalized by the number of pixels) around the center of the filament region (black line in Figure 2.4 (c)). Close to the center of the filament, the distribution shows large variations due to the fine scale contrast and small integration volume. The smoothed average is shown as the red dashed line with the black dashed line denoting the average signal intensity away from the filament. The HAADF intensity in the filament core is approximately $20 \pm 5\%$ higher than that in the unaffected oxide material far from the core and ring regions. The intensity in the dark ring is approximately 2.5 ± 0.7 % lower than the matrix. The local variations of signal intensity in the filament core are higher and range from 85 % to 140 % of the background intensity.

Somewhat similar observations. albeit with lower spatial resolution, have been reported by other groups. Wei *et al*. has observed a dark feature in cross-sectional bright field image of TaO_x switching device [89]. Since bright field and HAADF modes typically show reversed contrast compared to bright field, this should also be interpreted as the local increase of Ta content. Similarly, Strachan *et al*. detected a local increase of the Ta signal in the plan view map of x-ray fluorescence [96]. Neither group commented on the origin of the contrast.

It is important to note here that the techniques such as bright and dark field TEM or X-ray fluorescence monitor total number of atoms interacting with the beam rather than the ratio of Ta/O as is the case for spectroscopy methods. Local increase of HAADF (as well as XEDS and EELS) intensity, indicates larger number of Ta ions in the TEM sample. Moreover, what the STEM experiments detect is the total Ta content in the beam path. If the sample thickness is uniform, the intensity translates into Ta concentration or density. Nonuniformities of sample thickness would produce apparent changes of both. The author would also like to assert that, since that the sample thickness was uniform and the device surface remained flat as verified by Atomic Force Microscopy, the total number of Ta ions increased at the location of the filament. This can only be explained if Ta ions are mobile during electroforming/switching and accumulate within the filament. This possibility was not considered in most of the existing models of the resistive switching [72], [97]–[102] even though the formation energies of metal interstitials [103] and their mobility are well documented [103]–[105] and are comparable to those of oxygen. More recent data do indeed indicate that Ta is mobile and electrochemically active in resistive switching devices [37], [106]. In addition to the Ta mobility, the HAADF images indicate that the Ta ions move laterally from the area of the ring to the filament core. Since this flux of Ta is perpendicular to the electric field and occurs in the material with the uniform initial composition, it cannot be driven by the electric field or the composition gradient. Outside of the filament, one can discern a pattern of dark wavy lines. The origin of these features was argued to be the self-shadowing effect of small surface corrugations. The protrusions on the layer surface block the arriving ions during sputtering and result in lower density of deposit in the valleys. This effect was analyzed in a recent publication [107].

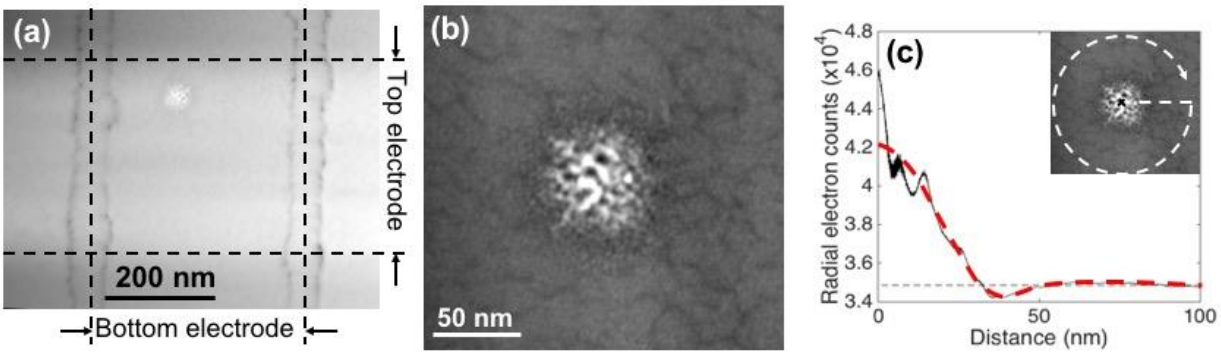

Figure 2.4 (a) Low magnification plan-view high angle annular dark field (HAADF) image of the active area of the TaO_x device. (b) High magnification HAADF image of the filament region. (c) Rotationally averaged HAADF line profile normalized by the number of pixels. The red dashed line indicates a smoothed profile.

The elemental distribution of Ta and O in the formed and switched devices was imaged using XEDS and EELS. The image in Figure 2.5 (a) is the HAADF plan view image of the filament area in a sample that was only electro-formed and serves as a reference (There were no significant differences in any of the maps or spectra obtained from devices that were only formed and those that were formed and switched. Typical forming and switching *I-V* curves are shown in Figure 2.2(a)). The two examples of XEDS spectra are collected on bright spot in the core of the filament and the background away from it (shown in Figure. 2.5 (b) as red and blue traces). The spectra show only peaks associated with Ta and O emission which is expected as the electrodes have been removed in the sample preparation process. The intensity of Ta M-line collected on a brightest spot within the filament is about 1.8 times as intense as the corresponding line collected away from the filament. The intensity of the oxygen K-line from the same spot is significantly lower than that in the surrounding area. This result suggests the increase of Ta content in agreement with the HAADF images and concomitant oxygen depletion. It is, however, difficult to quantify the concentration of oxygen as the O K-line is strongly absorbed by tantalum. Since the density of Ta is higher within the filament, the apparent decrease of O X-ray emission could be due to increased reabsorption. As a consequence, we have refrained from interpreting the XEDS oxygen intensity distribution maps. Table 1 lists the photon counts obtained on different locations within the filament including bright and dark spots within a core, Ta-depleted ring, and background. The values have been averaged over many locations and are in general agreement with the line profile in Figure 2.5 (c).

Location	$I_{Ta\text{-}M\alpha}$ (counts)	$I_{O\text{-}K}$ (counts)
Core (bright spots)	2636	453
Core (dark spots)	1886	488
Ring	1922	510
Background	1988	506

Table 2.2 Intensities of Ta-M and O-K lines at different locations in the filament.

The Ta XEDS M-line does not suffer from the absorption effect and was used to map the Ta distribution shown in Figure 2.5 (c). The map shows one-to-one correlation with HAADF intensity, which confirms attribution of the HAADF contrast to the Ta density changes. The filament region exhibits higher Ta concentration with the local increase by $\approx 20 \pm 5$ % compared with the background (Figure 2.5 (c)). This value corresponds to the concentration change averaged through the sample thickness and is likely much smaller than the local changes associated with the fine structure. A darker ring around the bright core corresponds to $\approx 3.5 \pm 0.7$ % decrease of X-ray intensity. This value has a relatively large error due to changing thickness of the sample close to the ring area. Even though the ring area is bigger than the central core, the integrated HAADF and XEDS signals give positive values corresponding to the net increase of Ta in the oxide. We believe that this effect is due to possible error in background subtraction and the fact that the thickness of the TEM sample was approximately 50 nm and was less than half of the entire functional layer thickness. Potentially, a vertical motion of ions could produce such an increase in the total Ta content in the specimen. Another possible contribution could come from the thin Ta layer that was part of the top electrode. Given a small thickness of this layer (2 nm) compared to the functional layer thickness (120 nm) it appears unlikely that the Ta layer could affect Ta content in the filament more than 5%.

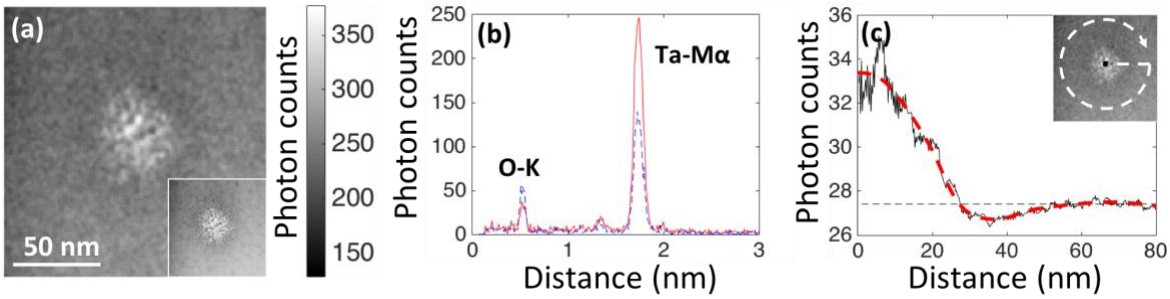

Figure 2.5 (a) XEDS map of Ta M line intensity (insert shows a reference HAADF image of the same area) (b) Examples of XEDS spectra obtained on bright HAADF spot in the core of the filament (red continuous line) and that of background (dashed blue line). (c) Radially averaged line profile extracted from (a). The

red dashed line indicates the smoothed-out profile, the black dashed line is the background intensity, and the inset shows the schematics for averaging.

The oxygen distribution was assessed by EELS on a different sample than the two shown in Figures 2.4 and 2.5. The device was electroformed and switched 10 times. Figures 2.6 (a) and (b) show the reference HAADF map of the filament and the two EELS O-K edge spectra collected within (red) and away (blue) from the filament, respectively. There is a clear difference in the intensities of the oxygen signal suggesting lower oxygen content within the filament. The oxygen distribution map created using these data is shown in Figure 2.6 (c). However, the raw electron count in EELS needs to be corrected for the changes of the beam intensity. In EELS, electrons are collected only within the small aperture around the optical axis of the instrument. In the regions with higher Ta concentration, part of the beam is scattered at high angles missing the detector and effectively lowering the beam intensity. The change of beam intensity was determined by collecting the zero-energy-loss peak intensity map at the same time as the EELS oxygen map. The results are shown in Figure 2.6 (c). The zero-loss-peak map shows a clear decrease of intensity within the filament, similar to that seen in bright field images (not shown) and inverse of that in HAADF. The corrected map of oxygen distribution (Figure 2.6 (d)) is uniform with the precision of experimental error with no indication of depletion of oxygen reported by others. This apparent contradiction warrants a more detailed discussion. The oxygen EELS maps reported so far have not been corrected even though the authors reported the change of the beam intensity in the bright field image [61], [89]. It is possible, that the reported oxygen depletion was an experimental artifact. The changes of oxygen distribution reported by Kumar *et al.* [48], [108] have occurred only after prolonged switching with high power dissipation. These authors have not reported any changes of oxygen distribution after forming performed in conditions similar to ours. In this respect, the two results are in agreement. A more significant potential issue is that the local increase of one element could lead to large increase of strain energy unless compensated by the decrease of the other element density. In other words, there has to be a local depletion of oxygen to allow for extra Ta ions in the sample with uniform thickness. For the sake of discussion, let us assume that there is no strain caused by the Ta accumulation and the starting composition of the oxide is TaO_2. In such material, Ta occupies less than 8 % of the volume due to large difference in ionic radii difference between Ta^{4+} and O^{2-} [109]. Increase of Ta ion density by 20 % in a unit volume, can be compensated by about 2 % decrease

of oxygen. Since the experimental error of oxygen content in our EELS maps is 5 %, we should not be able to detect the oxygen density change induced by the Ta increase. It should be noted here that changing composition of the oxide does not have to cause appearance of uncompensated charges. It is safe to assume that oxygen is always in the O_{2-} state. The charge of the tantalum ions, however, is likely changing with composition. In Ta_2O_5, it is 5+ while in TaO_2 only 4+ always resulting in electrically neutral oxide. The above does not imply that there is no change at all of the oxygen content within the filament compared to that of the initial film composition. Oxygen density could be somewhat smaller that of the matrix, but the relative change of oxygen atomic density is certainly much smaller than that of tantalum.

The corrections of XEDS and EELS maps to account for re-absorption of X-rays and high angle scattering by heavy metal ions are clearly significant at least in our experimental conditions. The amount of correction can change depending on the thickness of the functional layer and the elemental composition, but these effects need to be carefully assessed before reaching conclusions on the relative changes of composition.

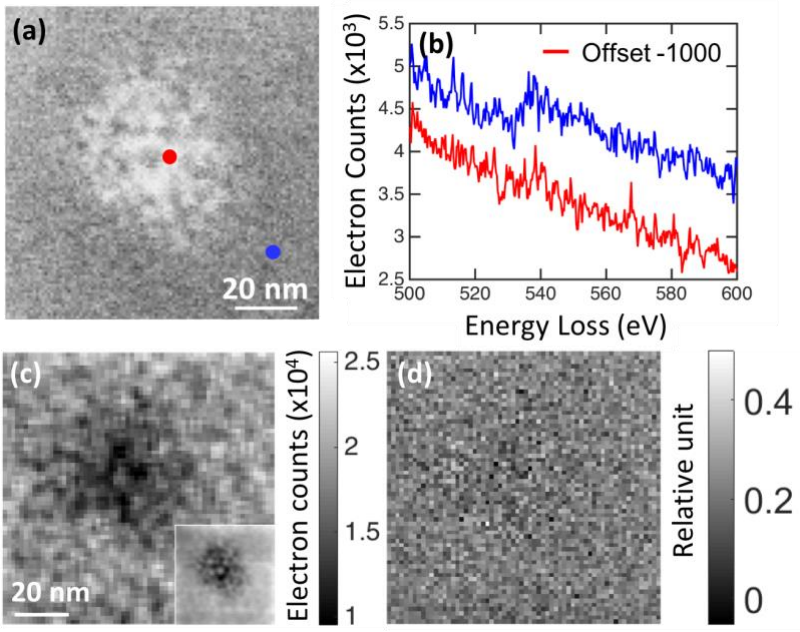

Figure 2.6 (a) High angle annular dark field image of the filament with marked locations where the EELS spectra were collected (b) EELS spectra obtained at red and blue dots in (a), (c) Map of O-K edge intensity. The inset shows the corresponding zero-loss image. (d) Map of O-K edge in (c) normalized by the zero-loss map.

The experimental results above can be summarized as follows. Devices that have been formed/switched at low dissipated power (~50 μW), do not exhibit any electrode morphology changes due to excessive local heating. They do show one well-defined filament with increased Ta content in the center of the filament and a metal-depleted ring around it. The lateral relative oxygen density changes were too small to be detected by EELS. Also, the inside of the filament exhibits fine scale contrast also associated with local changes of Ta content.

2.4 Soret Effect During the Forming Event

The interpretation of electro-formation and observed local changes of oxide composition relies on a two-step process similar to that proposed by Ielmini *et al* [42], [74]. These authors proposed that, in the first step, a small hot spot is created in the device by any number of mechanisms. This is followed by Soret effect which drives demixing of TaOx and produces lasting changes of composition.

We will start the discussion with the step one: formation of the small hot spot in what is initially a uniform device. Typically, as-fabricated metal/oxide/metal devices exhibit current that increases super-linearly with applied voltage. The $\partial I/\partial V$ is increasing with bias and at threshold voltage becomes infinite. In the circuit that has a large load resistor, the device can enter into part of characteristics where the slope becomes negative (Negative Differential Resistance (NDR) range, see detailed discussion of *I-V* characteristics of unformed TaO_x devices in Sharma *et al.* [44] and Goodwill *et al*. [18]). In addition to TaO_x, the NDR characteristics have been observed in TiO_x [44], VO_x [110] and NbO_x [111], [112]. The prevailing model of such behavior relies on fast increase of conductivity with device temperature [18], [82], [111]. Within NDR region, the current flow is no longer uniform across the device. Instead, the current spontaneously collapses to a narrow constriction and creates a hot spot due to Joule heating [113]. After the bias is removed, the constriction disappears leaving no physical changes, the device retains its original characteristics. The size of current constriction has been documented experimentally by Scanning Joule Expansion Microscopy in case of TaO_x [46], which agrees with the simulation of the electrothermal model that has also managed to reproduce the experimental *I-V* characteristics and dynamics of switching [18], [114]. The model predicts

the size of the constriction in the 0.1-1.0 μm range, depending on the composition of the functional layer, and the temperatures exceeding 1000K [18]. The model is used here to estimate the current density distribution and the temperature in structures used in the TEM experiments described above.

The forming experiments have been conducted using a circuit consisting of a voltage source, a switching device and a series resistor with R_{LOAD} = 1 MΩ with the experimental *I-V* shown as black line in Figure 2.7 (a). As described above, the *I-V* characteristics exhibit NDR region in which the current spontaneously constricts [18]. If the voltage sweep is stopped soon after the knee (the point at which $\partial I/\partial V$ becomes infinite), the device can return to its initial state if the bias is removed. With further increase of source voltage (the voltage across the device used in Figure 2.7 (a) is dropping), current constricts further and, at some point, the device threshold switches and transitions along the load line (black dashed line) to a new state with even smaller and hotter constriction [18]. Almost immediately afterwards, the device electro-forms permanently changing its characteristics: the return trace of the characteristics (upper branch of the *I-V*) differs from the initial one (Sharma *et al*. Fig. 4 [44]).

The experimental threshold switching *I-V* in Figure 2.7 (a) was simulated using a finite element model, which solved two coupled differential equations describing charge and heat flow. The simulated *I-V* is generated from the same model as the one shown in Figure 1.2 (a), the difference being the different device lateral size, as well as the activation energy and the pre-exponential factor of the electrical conductivity of the TaO_x layer. The results are shown as blue squares in the *I-V* plot. The model reproduced the main features of the experimental *I-V* with correct value of the knee voltage and the switching event without any adjustable parameters such as the size of constriction or conductivity. In parallel, this approach produced the current density (Figure 2.7 (b)) and temperature distribution (Figure 2.7 (c)) in the device. As is apparent from Figure 2.7 (b), the limit imposed by the load resistor on the current in the forming circuit led to the current constriction with the current forming a cone with the tip near the bottom electrode and the wider base at the top. The shape of the constriction is caused by the larger than typical thickness of the functional oxide. In thinner oxides, the current density distribution has more cylindrical shape with the maximum temperature in the middle of the oxide. For the device structure investigated here, the maximum of

temperature is shifted towards the bottom electrode. It is evident that the steepest temperature gradient is oriented laterally with a smaller but still significant component pointing vertically toward the bottom electrode. The upper range of simulated temperatures (> 1000 K) has significant uncertainty as the values of electrical and thermal conductivities of materials used in the device structure are not well known in this temperature range. The transition from the OFF to ON states of the threshold switch takes less than 10 ns, a time that does not allow for appreciable diffusion distance [114]. One can, therefore, approximate the entire forming process by two distinct steps. First, the device quickly forms the hot spot during threshold switching event, which is followed by slower diffusion of ions in the imposed temperature distribution.

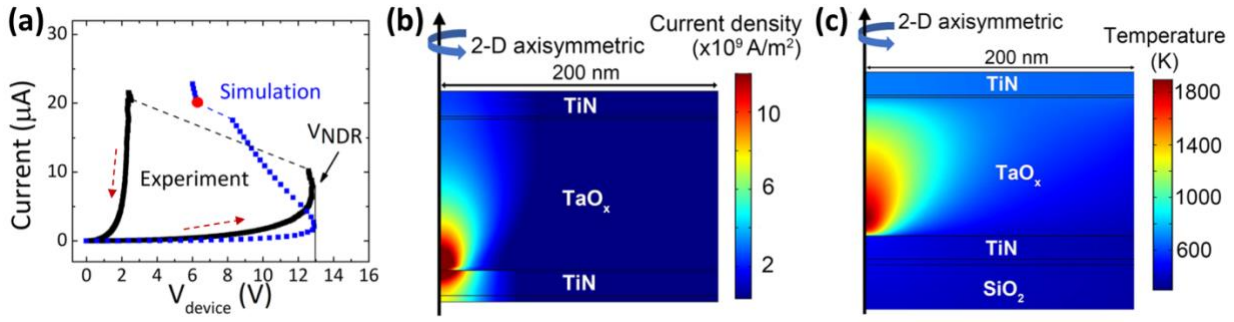

Figure 2.7 (a) Experimental (black squares) and simulated (blue squares) forming *I-V* of the device. Simulated current density (b) and temperature distribution (c) right after the threshold switch (red point in (a)).

The segregation of oxygen and tantalum in the TaO_x functional layer somewhat resembles elemental segregation in phase change threshold and memory devices [115]. The driving forces that have been proposed in this case include carrier wind, the electric field, the temperature gradient (Soret effect), the mechanical stress induced by the local lattice expansion, and the instability of the oxide [115], [116]. The last of these forces originates from the fact that there are only two stable phases in the Ta-O binary system: Ta and $TaO_{2.5}$ [117]. The starting composition in our devices [O]/[Ta] is between 2 and 2.5 based on XEDS and atom probe tomography measurements (data not included). Under prolonged heating, the functional layer should separate into these two crystalline phases. In switching experiments, the high temperature was not maintained long enough to cause complete segregation and crystallization, but one could expect partial segregation. Among the five segregation phenomena listed above, the first two could only induce vertical motion of ions, which could not be detected in plan-view geometry. The same applies to the compressive stress induced by thermal expansion: lateral motion of ions could not relieve the accumulated

strain energy. In any event, the change of lattice parameter between room temperature and 1000 K is only 0.4 % [118] and could only induce a comparable amount of atom density change. This leaves only the last two to account for the observed concentration profiles. Once the segregation produces concentration differences, another driving force, the concentration gradient, would act as the opposing agent trying to homogenize the oxide. The balance between these forces would lead to steady state distribution.

The segregation driven by the temperature gradient has to produce the distribution with the symmetry reflecting that of the driving force. Among two segregation length scales, large corresponding to accumulation in the center and depletion within the ring and fine giving blotched contrast throughout, it is the first that has the required symmetry. The large-scale Ta density profile in Figure 2.5 (c) could be interpreted as follows. The lateral flux of Ta or O ions due to Soret effect can be described by the modified Fick's law shown in Equation 1.1 and 1.2. The sign of S_{Ta} is positive and the flux is directed toward the center of the constriction. Far from the constriction, the diffusion coefficient D_{Ta} is very low and the flux is close to zero. As one moves toward the hot spot, D is increasing monotonically with temperature. For a given volume between distance r from the center and r+Δr, the flux from the outside of the ring is lower than the flux out towards the center. This creates a locally Ta depleted region at the periphery of the hot spot. At certain distance from the center, the flux reaches maximum and gradually decreases to zero at the center. Using temperature distribution shown in Figure 2.7 (b) and the values of Ta diffusion coefficient, one can estimate the size of the permanent filament. We have assumed that one can detect the changes in the composition of the film if the mean square displacement due to diffusion during 10 ms at temperature is equal to 1 nm. This gives the value of Ta diffusion coefficient of 5×10^{-17} m_2/s. Unfortunately the values of Ta diffusion coefficient in amorphous TaO_2 is not known, and an estimate was used by the value determined by Gries *et al*. for diffusion of Nb in crystalline Ta_2O_5 [119]. This gave the temperature at the outer boundary of the filament of ~1000 K. The map of the radial temperature distribution in the constriction in the plane parallel to and 80 nm above the bottom electrode interface is shown in Figure 2.8 (a). This particular distance was selected as the sample used in STEM experiments was likely located in the upper part of the device. The diameter of the circle corresponding to 1000 K is approximately 80 nm and this would be the

estimated diameter of the outer Ta-depleted ring. The diameter observed in the HAADF image was 110 nm and well within a factor of two of the estimate.

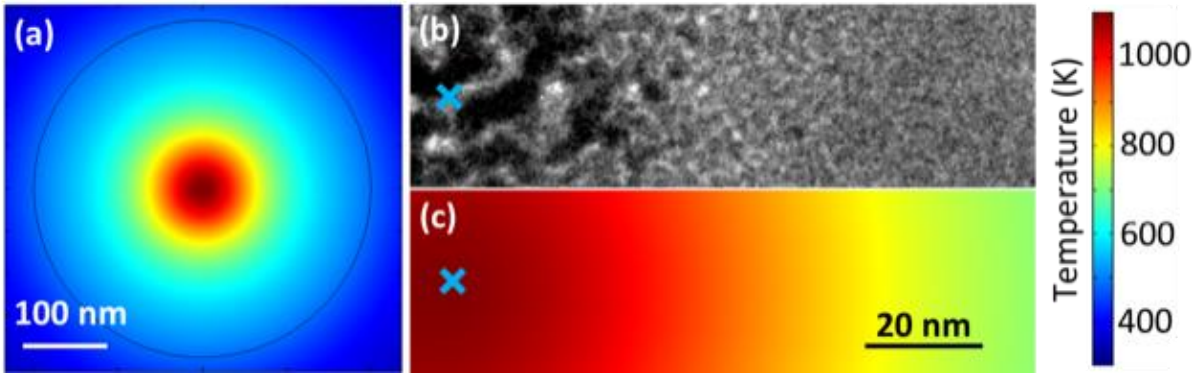

Figure 2.8 (a) Modeled plan-view temperature distribution during threshold switching in a plane 80 nm above the interface between TaO_x functional layer and bottom electrode. The black circle is the boundary of the active device region. (b) Cropped bright-field image of the filament and surrounding region (blue cross marks the center of the). (c) Corresponding temperature distribution in the same area as (b).

The direction of Ta motion reported here is opposite to that of oxygen suggested by others. It is of interest to explore the origin of such difference. The Soret fluxes of atomic transport due to the motion of vacancies and interstitials were described in Equation 1.3 and 1.4 from Chapter 1, respectively. The Soret effect was also discussed in detail by Strukov *et al.* in RRAM cells [38], however the heat of transport (q_v*) was neglected in their derivation reached the conclusion that the Soret coefficient of O should be negative, which would result in O diffusing away from the hot center of the current constriction. Similarly, the transport of Ta due to the motion of interstitials in response to the temperature gradient would be opposite to that of O if one neglects the q_i*, i.e. the temperature gradient would drive tantalum ions toward the high temperatures in agreement with the experiment. It should be noted, however, that the absolute value of the heat of transport can be higher than the migration enthalpy of Ta interstitials and can be of either sign. At this point in time, there is no available model of the heat of transport and we cannot predict the sign of the Soret coefficient. Its values are typically determined from experiment and have been found be either positive or negative. For example, the magnitude of the Soret coefficients for Ge and Sb in $Ge_2Sb_2Te_5$ have opposite sign [115], [116].

The remaining issue is a possible effect of phase separation and the origin of fine contrast in the elemental maps presented above. It is easy to notice that the scale of fine contrast is changing with the temperature at any given location: the size of the bright and dark areas in the center of the filament is about 5 nm and decreases to 1 nm at the distance of 40 nm from the center (Figure 2.8 (b)). The contrast is equiaxial (at least in the plan view presented here): the intensity changes along the radial direction as well as perpendicular to it. The azimuthal variation cannot be caused by the temperature gradient as it has only radial component. The fine scale segregation is, in our opinion, due to decomposition of TaO_x toward the two stable phases in the phase diagram [117]. The change of size is associated with Ostwald coarsening with the size of the feature increasing with local temperature (the time at temperature was approximately the same).

2.5 Discussion

Figure 2.9 presents a sequence of events proposed here as the mechanism for the electro-formation of conducting filaments in metal oxides. The initial state corresponds to a uniform distribution of Me_{x+} ions and a uniform current flow at low voltages (a). At the threshold voltage, the device threshold switches to the volatile ON state characterized by a small diameter hot constriction with large lateral temperature gradient (b) and a smaller vertical component marked by the red arrows. The gradient produces the flux of Ta ions by the Soret effect (c) and results in accumulation of excess metal in the hot center of the filament (d). It should be noted here that the proposed sequence does not address bipolar switching.

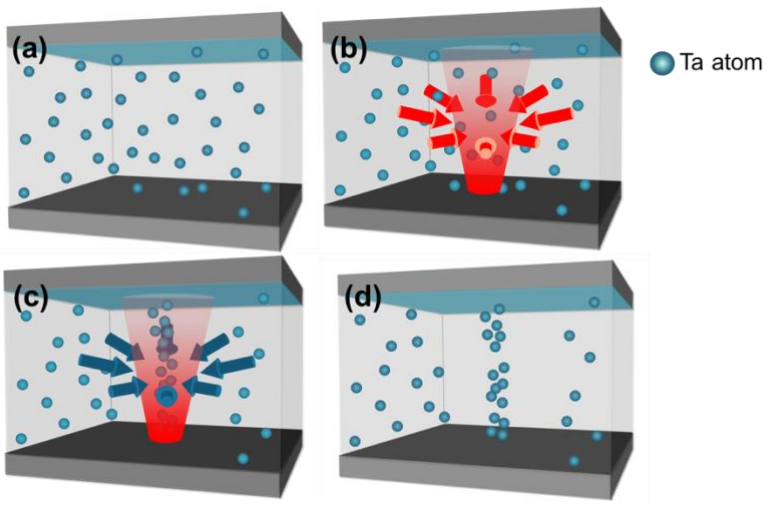

Figure 2.9 Proposed model for the formation of the conducting filament in TaO_x based RRAM. (a) As deposited uniform distribution of Ta ions in the functional layer. (b) Threshold switching-induced volatile current constriction and the temperature gradient distribution. (c) Ta atoms migrate in the functional oxide. (d) Permanent Ta-rich filament is created as result of the forming event.

As proposed, the process includes two positive feedback loops that could result in runaway events. The first one is the thermal feedback resulting in the threshold switching event. This process can be illustrated by considering the device response to applied voltage pulse. Immediately after the beginning of the pulse, the current reaches value corresponding to the conductivity at the stage temperature. The dissipated Joule heat causes a temperature rise which in oxides results in current increase. If the dependence of electrical conductivity on temperature is steep enough, this positive feedback loop will produce rapid increase of temperature and current. The process will stop whenever the external load or some extraneous process within the device (for example melting and delamination of the electrodes) will limit current increase in the circuit. The size and maximum temperature of the hot spot created by this event is determined mostly by the circuit (load) external to the device rather than by intrinsic device or material characteristics. Since the load is rarely mentioned in reports on resistive switching (current compliance set on the source meter cannot replace the load located next to the device [88]), the temperatures reached during electro-formation are not known with any accuracy resulting in a large range of effects observed. In the experiments described above, the load was carefully selected, and its effects included in the model of temperature distribution.

The second runaway is of chemical nature and is caused by positive feedback between oxide composition

and current/conductivity. As the temperature gradient drives the Ta ions toward the center of the filament, the changing stoichiometry causes an increase of conductivity and total current. This, in turn, increases the temperature and the temperature gradient i.e. the driving force for elemental segregation. In many devices structures/electrical bias scenarios, this can lead to a second runaway event, this time the runaway of composition. The consequences include rapid rather than gradual change of permanent filament resistance [120], [121]. It also is expected that the diameter of the cylinder with increased Ta content, would decrease as the result. At the end of the process, one could form a very small metallic Ta filament. Once such a filament forms, the conductivity would become constant with current increasing only with the diameter of the filament. The compositional runaway should stop. It is likely that the process would come to steady state before formation of the metallic Ta phase if the dependence of conductivity on composition levels off at high oxygen deficiencies. The limited available data on conductivity do not allow for making a definitive assessment of the final filament composition. The highest increase of local Ta atomic percent obtained from the XEDS is 80% (assuming the initial composition to be TaO_2). This fraction is almost certainly an underestimate as the composition values are averaged over the thickness of the sample. It could well be that some of the bright spots in the Ta intensity maps correspond to Ta phase. The filament thus created resembles very closely the filaments reported in the conducting bridge memory devices in which the conduction occurs between metallic inclusions created by coalescence of cations [122]–[124].

Independently of these arguments and similarly to the thermal runaway, the chemical runaway could also be forced to stop by the current limit imposed by the circuit. This would make the size and composition of the filament very strongly dependent on experimental conditions of forming and switching. Accordingly, these characteristics are not universal and particular values reported here should only be applied to other devices/processes with caution. The value of the results rests in identification of processes and their relative magnitudes.

The significant role played by the Soret effect in electroformation process explains the relative stability of the filament over as many as 10_{12} switching cycles reported in literature.[125] It is clear that the non-volatile filament gets hot during switching as the power dissipated over small volume of the filament must lead to

large temperature excursions [72], [126], [127]. This has to lead to high diffusion coefficients and atomic fluxes down the concentration gradient. As the consequence, concentration-gradient-induced diffusion would result in dissolution of the filament after a few switching events. Clearly, this does not happen. The Soret flux of Ta toward the core compensates the outward diffusion flux providing for the filament neither shrinking or dissolving. The same argument applies to oxygen and oxygen vacancies. It is also important to note that Soret effect should play an important role in switching in addition to that played in electro-formation. The temperature gradients close to interfaces with metallic electrodes are expected to exert a force on the ions driving both oxygen and metal ions in the vertical direction as well as the lateral one. This force should be quantified and included together with electrostatic and carrier wind components in a comprehensive model of resistive switching. Moreover, the fact that Ta segregation induced by the temperature gradient is much more pronounced than that of oxygen could indicate that the same applies to motion induced by the electric field i.e. that the switching is due to mostly to the motion of metal ions rather than that of oxygen. Such a hypothesis would require a careful assessment of elemental drift and discussion in cross section geometry.

In this chapter, through STEM analysis, the lateral motion of metal ions forming the Ta-rich filament was discovered, and the short-scale phase segregation in the filament was also revealed. The electrothermal simulation by finite element modeling confirmed the lateral motion of Ta was driven by the lateral temperature gradient due to the localized high temperature corresponding to the current constriction during NDR. The model described in this chapter parallels the model proposed for the TCM switching, where the temperature gradient plays the major role responsible for the resistive switching.

Chapter 3 Stable Metallic Enrichment in Conductive Filaments in TaO_x-Based Resistive Switches Arising from Competing Diffusive Fluxes

3.1 Introduction

The results from Chapter 2 suggest that the memory cell undergoes a positive feed-back loop between Joule heating and electrical conductivity, which induces a thermal runaway during the voltage sweep for forming and causes the formation of the conductive filament. From the *I*-*V* shown in Figure 2.2 (a), one can see the formation of the filament (the "snap" in the *I*-*V*) happened very soon after the *I*-*V* entered NDR. Due to the relatively high resistance and large thickness of the as-fabricated functional oxide of the devices used in Chapter 2, the forming required a large V_{Device}. During forming, the voltage across the device decreased over a large range of voltage, which could cause significant energy discharge from the parasitic capacitances (such as device self-capacitance [20] and parasitic capacitance from the circuit [128]) onto the device. The significant parasitic capacitance discharge could explain why the forming happened almost as soon as theses device entered NDR. Moreover, the discharge could also increase the temperature in the memory cell and introduce potential uncontrolled artifacts during forming. Therefore, a new device structure was designed with a more conductive functional layer and on-chip resistor to reduce the forming voltage and hence the energy stored in the parasitic capacitance, as well as reducing the magnitude of current spike due to capacitive discharge.

Studying the filament in plan-view provided quantitative information of the filament composition and identified lateral driving force of the filament formation. However, plan view geometry did not allow to assess the motion of ions along the field direction. In particular, it was not possible to evaluate the filament changes between HRS and LRS. This information could be obtained only in the cross-sectional TEM analysis.

Using the devices with new structure and composition, in this chapter, a combination of high-angle annular dark-field scanning electron microscopy (HAADF) and electron energy loss spectroscopy (EELS) are performed in cross-sectional and plan-view sample geometry to provide quantitative information on tantalum and oxygen distributions including filament diameter, composition, and gap width. The maps are interpreted in terms of electro-formation and switching mechanisms. Additionally, finite element modeling is used for the electrothermal simulation to extract the temperature in the device during programming, by fitting the experimental *I-V* with the simulation.

3.2 Experimental

1) Device fabrication

The devices were fabricated with an inverse-via structure (Figure 3.1). A blanket 40 nm thick TiN layer was first deposited on the thermal oxide/Si substrate followed by photolithography and dry etching to define the bottom electrode and the on-chip load resistor (R_{Load}). The patterned TiN was covered by reactively sputtered 50 nm thick TaO_x and 10 nm SiO_x blanket layers. The active area of the device was defined by e-beam lithography and dry etching of a 150 x 150 nm^2 via in the SiO_x. The stack was finished by sputtering another 40 nm thick TiN top electrode layer and a lift-off process. The TiN layers were sputtered at 150 W, 3 mTorr ambient pressure with 60 sccm Ar flow; the TaO_x layer was sputtered at 50 W, 3 mTorr ambient pressure with 58.5 sccm Ar and 1.5 sccm O_2 flow.

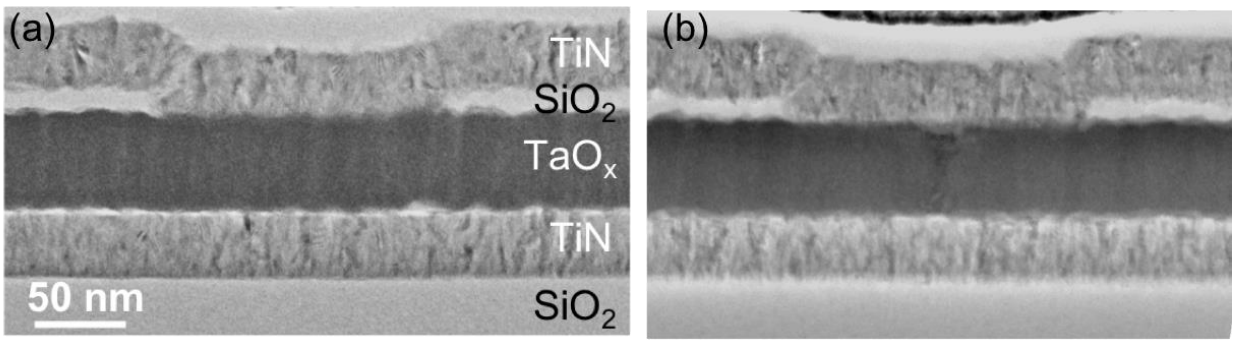

Figure 3.1 (a) Cross-sectional bright-field TEM of a pristine device. (b) Cross-sectional TEM of a device after forming in the positive polarity.

2) Electro-formation and switching

Devices have been formed while on Si chip (with a 1 μm thick thermal SiO_2 coating) in order to assure optimum heat sinking. A carefully designed forming procedure was used in order to avoid the artifacts associated with excessive temperature excursions and dissipating as little power as possible for this device design. The circuit used for formation and switching is shown in Figure 5 (a) with forming *I-V* in panel (b). The small lateral size (150×150 nm_2) limited the device self-capacitance and the on-chip load resistor (58 kΩ) reduced current spikes due to discharge of parasitic capacitances [19], [20], [88]. The devices were formed using a triangular source voltage sweep with the positive voltage applied to the top electrode (unless otherwise noted). The as-fabricated structures exhibit an S-type negative differential resistance (S-NDR)) extending between points A and C. In this range, there are no permanent changes of the device structure or electrical properties and the *I-V* can be retraced back and forth many times. The forming occurred when the current exceeded that at point C with the resistance decreasing with time once the process has started. Upon reaching point D, the time evolution stopped and the D→A→E part of the characteristics could be retraced repeatedly. The formation process was completed by increasing the negative source voltage with the device evolving from point E to F. The typical switching characteristics following formation are shown in Figure 3.2 (c). The cross-sectional view of a device that was electroformed under positive bias (top electrode) and left in the HRS is shown in Figure 3.1 (b).

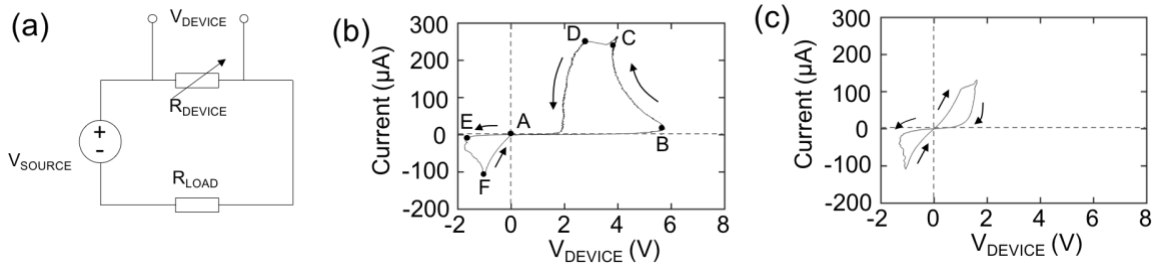

Figure 3.2 (a) Circuit used for electro-formation and switching. (b) *I-V* trace of the electro-formation process. (c) Switching *I-V* of fully formed device.

3) Device simulation

The device behavior was simulated using COMSOL Multiphysics, a finite element simulation software package, with the same model as that of the simulation presented in Chapter 2, but with different input parameters due to the different device structure, and a different composition of the TaO_x by changing the deposition conditions. The input parameters of the simulation were the material properties (density, specific heat capacity, thermal conductivity, electrical conductivity and the relative dielectric constant) and the circuit parameters including source voltage and load resistance. The materials properties are the same as in that was adopted by Goodwill *et al*. [18]. The devices had a square 150×150 nm^2 active area, which in the simulation was approximated with axisymmetric 2D device with radius of 170 nm. The temperature of the bottom surface of the Si was set to 300 K while the top and side surfaces of the structure are thermally insulated. Electrically, the bottom surface of the BE is at electrical ground, and the top surface of the TE is connected to an outside circuit with a load resistor and voltage source. In the simulation, the source voltage was increasing linearly with time over a 1 s period followed by a linear decrease to zero at the same rate.

4) FIB lamellae preparation

For the cross-sectional FIB lamellae, the entire active volume of each device was lifted out and thinned down by FIB (FEI NOVA Nanolab 600) leaving approximately the middle third of the device within the lamellae. During thinning, the Ga^+ FIB was kept at 30 keV, 1 nA – 50 pA, and 89 – 90˚ incident angle. Afterwards, Ga+ FIB at 5 keV, 70 pA, and 85˚ incident angle was used to polish the sidewalls of the FIB lamella on both sides to remove the surface damage. The filaments were captured in the cross-sectional TEM samples with about 80% success rate. For the plan-view FIB lamellae, the devices were lifted-out in plan-view with Xe^+ plasma focused ion beam (FIB) (FEI Helios) and thinned down to electron transparency using Ga+ FIB (FEI NOVA Nanolab 600). In order to leave the free-standing TaO_x functional layer for TEM analysis, the material beneath functional layer (Si substrate, 1 µm SiO_2 insulation layer, 5 nm Ta adhesion layer, and 20 nm TiN bottom electrode) was first removed, then the top electrode on the other side of the lamella was removed, both using Ga^+ FIB at 30 keV, 1 nA-50 pA, 89˚-90˚ incident angle. Afterwards, Ga^+

FIB at 5 keV, 70 pA, 85˚ incident angle was used to polish the side walls of the FIB lamella on both sides to remove the surface damage.

5) TEM/STEM analysis

The STEM HAADF images were collected with an FEI Titan 80-300 microscope operating at 300 keV with Fischione model 3000 HAADF detector. The camera length was 48 mm to keep the HAADF collection angle greater than 50 mrad for true Z-contrast. The STEM EELS data was collected also with an FEI Titan 80-300 microscope operating at 300 keV with Gatan Quantum ER detector.

6) Atom Probe Tomography

A Nova 200 dual beam FIB/SEM was used to lift-out and mount TiN/TaO_x/TiN structures to Si micro-posts, shape the layers into a needle shape with an approximately 50 nm radius, and clean the surface with a 2 kV final mill. The Si micro-post array was then transferred to a local electrode atom probe (CAMECA LEAP 4000 XHR), and run in laser mode with a 200 pJ laser power, 30K base temperature, a collection rate of 1 atom per 1,000 pulsed, and a pulse frequency of 125 kHz. The APT data was analyzed using CAMECA's integrated visualization and analysis software (IVAS) version 3.8.2. The APT experiments were performed by Jonathan D. Poplawsky from Oak Ridge National Laboratory.

3.3 Composition of the TaO_x and the TiN

An as-fabricated device was lifted out and thinned down to needles for Atom Probe Tomography (APT). Reconstruction of the atom probe is shown in Figure 3.3 (a). Quantification from the mass spectrum (Figure 3.3 (b)) shows that the [O]/[Ta] ratio is 1.9. Note that a small portion of O can be lost during the APT experiment due to the field evaporation of neutral O species. The [O]/[Ta] ratio calculated based on APT data is likely the lower bound of oxygen content. The composition of the TiN electrode was determined to be $TiN_{0.8}O_{0.2}$ in the same fashion. Figure 3.3(c) shows the interfacial concentration line profile at the TiN/TaO_x interface from TiN to TaO_x. $_{32}TiO^{++}$ and $_{16}Ti^{+++}$ mass spectral peaks were

deconvoluted from the $^{32}O_2^{2+}$ and $^{16}O^{+}$ peaks using the known isotopic abundancies of Ti and TiO for the interfacial profile.

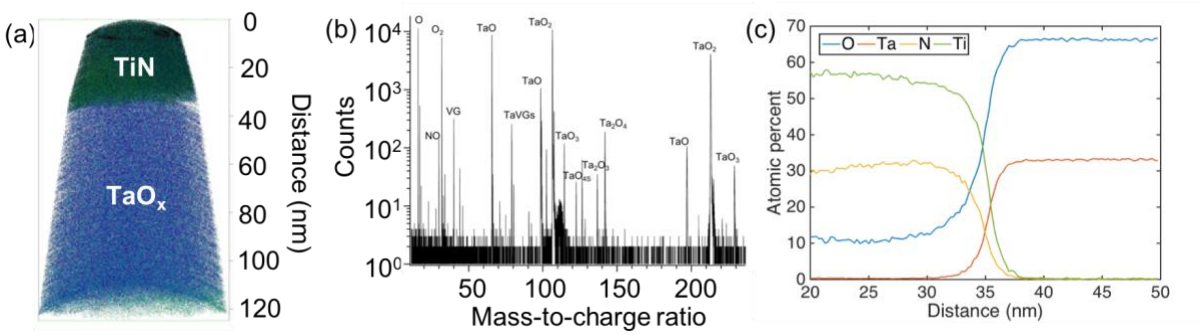

Figure 3.3 (a) Reconstruction of the atom probe sample prepared out of TiN/TaOx/TiN stack. (b) Mass spectrum in the TaOx section of the sample from Atom Probe Tomography with vacuum gases (VGs) labeled. (c) Concentration line profiles of O, Ta, N, and Ti versus distance.

3.4 Formation of the Filament

Figure 3.4 (a) shows a cross-sectional HAADF image of a device that was electroformed in positive polarity (positive voltage terminal connected to the top electrode). Formation resulted in the device in the HRS i.e. with the gap within the filament. HAADF images, it should be noted, exhibit higher intensity in areas of heavier atoms, assuming uniform thickness of the FIB lamellae. The dark areas at the top and bottom of the image correspond to TiN electrodes while the grey area in the middle corresponds to TaO_x. We have examined the entire cross-section of the active part of the structure of about 10 devices and found, in each case, only one area of contrast that could be associated with the filament. Each filament was located in the center of the device where the simulation predicted maximum temperature. As shown in Figure 3.4 (a), the filament displays a higher HAADF intensity compared to the surrounding TaO_x and has roughly an hourglass shape with a narrow ~8 nm diameter section in the middle and about 2-3 times as wide segments close to the electrodes. The overall shape and size are similar to that observed by atomic force microscopy (AFM) [129], [130] and TEM [89]. In addition to brighter contrast in the filament, there is also a visibly darker contrast close to the top electrode within the filament. All

STEM images did not show any protrusions, depressions, bubbles, voids, or craters in any of the devices.

All HAADF images were obtained using electrons scattered between 83 and 528 mrad from the optical axis of the microscope. The scattering mechanism resulting in such a high angle involves interaction of beam electrons with the atomic nuclei, in which the scattering strongly increases with the atomic number.[83] Because our structure in the proximity of the filament contains only nitrogen (Z=7), oxygen (Z=8), titanium (Z=22), and tantalum (Z=73), virtually all electrons (~99 %) forming the image have been scattered by Ta ions. All of the contrast is due to local changes of Ta concentration integrated over the thickness of the sample. The variation of HAADF intensity was translated into the quantitative maps of Ta content using established procedures [131].

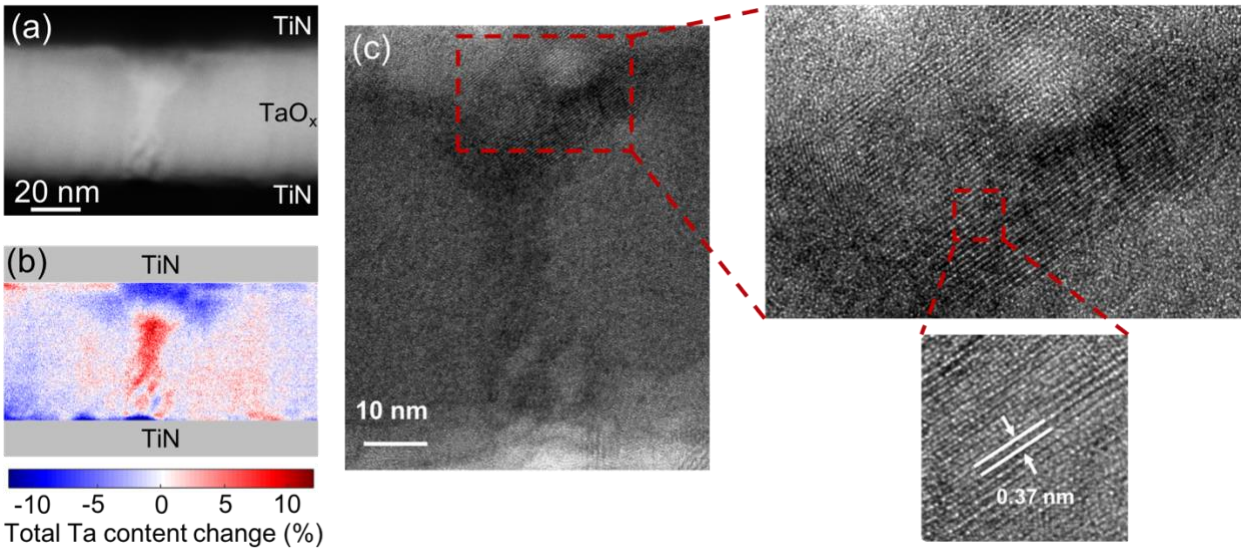

Figure 3.4 (a) High magnification HAADF-STEM image of the filament in TaOx device formed with the positive bias applied to the top electrode. (b) Corresponding map of total Ta content throughout the sample thickness. White color corresponds to the average Ta content far from the filament with red and blue representing an increase and decrease of Ta content. (c) Lattice image of a 10 nm crystallite found within the blue area at the top of the filament and close to the top electrode.

The local Ta content map integrated over the thickness of the specimen is shown in Figure 3.4 (b). The sample was ~70 nm thick as estimated by the log-ratio method [91]. The average scattering intensity in the figure is represented by the white color, with blue corresponding to the lower and red to higher integrated Ta content (the scale represents the percentage change of the average HAADF-STEM

intensity). The areas corresponding to TiN electrodes are marked by grey boxes, these areas did not contain any Ta either before or after forming/switching. The integral of the Ta content over the area 70 nm wide around the filament is the same as the integral away from the filament. This indicates that there is no significant net gain or loss of Ta in the area of the filament and the intensity changes are due only to redistribution of Ta within the oxide. It also indicates that the TEM sample was properly positioned and the entire filament was captured within the lamellae. Note that the red and blue features near the two interfaces between the functional layer and the electrodes from the HAADF image in Figure 3.4 (b) was caused by the interface roughness during the processing of the HAADF images for the quantitative analysis, therefore the contrast within ~3 nm of the interface should not be interpreted as the local variation of the Ta concentration. Also, in the TaO_x > 3 nm away from the interfaces, a large-scale (20-50 nm) mild contrast variation in addition to the filament can be observed as well, which is most likely caused by the slight thickness variation of the sample during the FIB thinning due to the uneven interfaces above the TaO_x, and should not be interpreted as a part of the local variation of the Ta concentration either. The interface contrast and the large-scale contrast are present in all of the processed HAADF images (red and blue maps), and will be disregarded in the rest of this thesis when during the analysis of the local variation of the Ta concentration.

The relative changes of Ta content have been recalculated to obtain the compositions within different segments of the filament. The procedure relied on the observation of a small (about 25 nm) crystallite within the blue area at the top of the filament (Figure 3.4 (c)). The lattice image of the crystallite (Figure 3.4 (c)) revealed set of planes with the separation of 0.37 nm corresponding to 0.369 nm interplanar distance of (010) planes in Ta_2O_5. Ta_2O_5 is the only compound in the Ta-O binary system [117] and has a narrow existence range. This fixed the stoichiometry of the blue area at $TaO_{2.5}$. Formation of the crystallite at this location was made possible not only by the proper stoichiometry of the oxide but also by the temperature distribution within the filament. Stoichiometric Ta_2O_5 is resistive while the oxygen deficient TaO_x is conductive. Accordingly, most of the voltage drop and power dissipation are expected within the gap in the filament [72], [100]. Resulting high temperature during formation allowed for crystallization of the material. The compositions elsewhere in the filament were obtained by normalizing

the integrated change of Ta content by the size of the filament. This gives the composition of the TaO_x in the functional layer as $TaO_{2.1 \pm 0.2}$ based on the change of Ta content and assuming the O density in the gap and the as-deposited TaO_x are the same. Independent estimates of functional layer composition by Atom Probe Tomography (Figure 3.3) gave the [O]/[Ta] ratio of 1.9 $\pm$ 0.2. We therefore assign the composition of the as-deposited functional layer to be $TaO_{2.0\pm0.2}$ using the average and standard deviation of the two compositional measurements. Given the thickness of the sample and size of the filament, we deconvoluted the contribution of the matrix to the contrast and estimated that Ta content within the filament is as high as 88 % relative to the surrounding matrix.

The plot of relative Ta content in Figure 3.4 (b) also draws out subtle features of the filament. In addition to the narrow core and the gap at the top, one can discern a light blue Ta-depleted region surrounding the core. The depletion is not uniform along the filament length. It is most pronounced in the upper 1/3 of the oxide layer and fades away in the lower part where the temperatures and diffusion coefficients were lower during formation. In our previous study, the depleted ring was observed in the plan view projection and interpreted as due to lateral motion of Ta ions driven by the temperature gradient (Soret effect) [50]. An important consequence of the Soret effect leading to accumulation of Ta is the stability of the filament. It is clear, that Ta is mobile in normal operating conditions. The local differences in composition would cause Fick's flux that eventually would homogenize the layer. This does not happen due to the balance between Fick's flux and Soret flux. In the bottom portion of the layer, the filament exhibits a fine structure with a length scale of ~5 nm. The overall accumulation of Ta in this region is much smaller than in the center of the layer and areas with enhanced Ta density are interspersed with areas of locally depleted Ta. Similar contrast was observed in plan-view and interpreted as due to phase separation of the Ta-rich oxide [50]. All features observed in the cross section are consistent with the observations made in plan-view.

The inspection of Figure 1 allows us to comment on the processes that led to the observed Ta distribution. The first question to answer is whether it is Ta, O, or both that are moving to induce such a distribution of Ta. The increased density of Ta in the core of the filament can only be interpreted as due

to Ta moving into this volume. Removal of oxygen would not produce a local increase in Ta and would create a vertical collapse of the structure. No surface corrugation was observed TEM/STEM images. Kumar *et al.* reported uniform Ta distribution in stressed devices fabricated on Si_3N_4 membranes [48]. The difference between their observations and the ones reported here could likely be due to a lower spatial resolution of the x-ray absorption technique compared to STEM. Since the HAADF is not sensitive to local O content, its mobility had to be assessed by other means.

In order to investigate the O content in the filament, we performed STEM electron energy-loss spectroscopy (EELS) analysis of plan-view samples similarly as in Chapter 2 [50]. In plan-view geometry, the electron beam is normal to the functional layer and parallel to the filament, thus maximizing the interaction volume with the filament. Figure 3.5 (a) shows a high magnification HAADF image of the filament after electroforming. Quantification on the HAADF images [131] shows a local increase of Ta content in the filament by a factor of 1.83 compared with the matrix, which is consistent with the results obtained in cross-section. Figure 3.5(b) shows the EELS spectra collected at the center of the filament (red trace) and away from the filament (blue trace), both after 5x5 binning. The log-ratio method for relative thickness estimate gives the thickness to electron mean-free-path (t/λ) of 0.37, which is acceptable for quantitative EELS analysis [91]. A normalization of the signal was applied to account for the loss of signal in the filament due to the increased scattering (same as the processing method of the EELS signal in Chapter 2). The normalized map of EELS O-K edge intensities in Figure 3.5 (c) shows ~57 % decrease in O counts in the filament compared with the matrix. An EELS measurement from another plan-view sample showed a 64 % decrease of oxygen giving an average of 60 $\pm$ 3 % decrease in O. Given the composition of the matrix and the change of both Ta (HAADF) and O (EELS), the composition of the filament was estimated to be $TaO_{0.4\pm0.1}$. These results so far do not indicate where the O goes after leaving the filament. Possible paths are: (i) moving towards the positively biased since O^{2-} is negatively charged [92]; or (ii) going horizontally away from the filament in the direction of the temperature gradient driven by the Soret effect [38], [132].

The EELS finding is different from what was observed in Chapter 2, where no visible variation of O in the filament was observed. This is most likely due to the fact that the device structure and testing methods between Chapter 2 and Chapter 3. The filaments formed in Chapter 2 had significantly larger diameters (~ 6 times larger) than those shown in Chapter 3, and the loss of O signal in the filaments from Chapter 2 can be significantly smaller than Chapter 3 assuming the amount of O loss in both device structures is the same.

The composition swing from $TaO_{2.0}$ to $TaO_{0.4}$ is large and poses the question of the build-up of local stresses and formation of protrusions or depressions on the surface of the device structure. One can notice, however, that Ta in the as deposited $TaO_{2.0}$ has a charge of either 4+ or 5+ with corresponding ionic radii of 0.066 nm and 0.064 nm.[109] The ionic radius of O_{2-} is 0.135 nm [109]. In this case, removing 60 % of O from the pristine $TaO_{2.0}$ results in a 57 % decrease of the volume. With a Ta-rich composition ($TaO_{0.4}$) [133], the radius of Ta could be approximated by that of Ta_{+} in metallic Ta (≈ 0.134 nm) and the volume occupied by Ta would increase by 71 %. Combining the changes caused by both O and Ta in the material can give a volume increase of 14 % between $TaO_{2.0}$ and $TaO_{0.4}$. It is a small change especially given the possible error of the Ta_{+} size estimate.

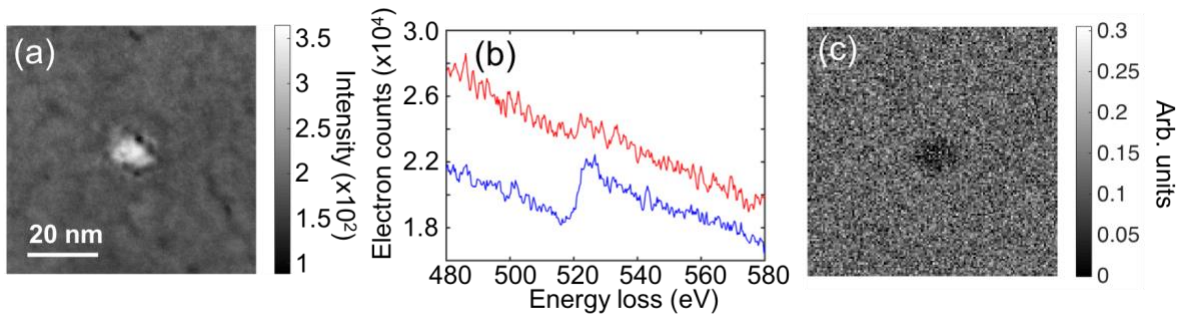

Figure 3.5 (a) HAADF reference image of a filament in plan-view. (b) Electron energy-loss spectroscopy (EELS) spectra collected from an area on the filament (red) and away from the filament (blue). Both spectra were after a 5x5 binning. (c) The O-K map after a normalization to account for the loss of signal in the filament due to the increase of Ta.

Cross sectional images contain information on Ta motion in both the vertical and lateral directions. From Figure 3.4 (b), it is clear that Ta ions moved downward away from the positively charged electrode. This motion created both the Ta depleted gap at the top of the filament and Ta accumulation in the lower

part of the TaO_x layer. The motion is consistent with the ion drift in the electric field with the velocity determined by the local value of the field and temperature.

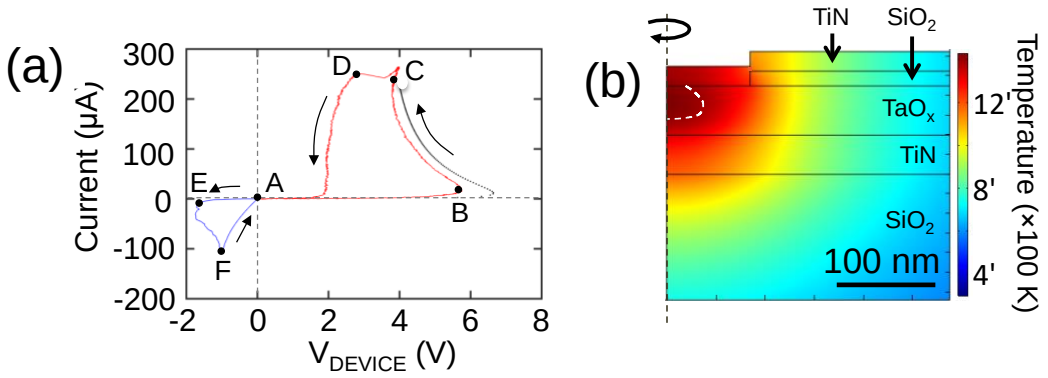

Figure 3.6 Experimental electro-forming *I-V* (red and blue traces) with the positive voltage values defined as positive bias on top electrode. The simulated *I-V* is shown as a series of black dots paralleling C to D experimental trace. (a) Temperature distribution within the device at the point of electro-formation (point C in Fig. 5 (b)).

The simulated temperature distribution at the beginning of the formation process is shown in Figure 3.6. The estimates of the temperature within NDR-type device are complicated by the expected electronic decomposition resulting in appearance of high current-density domains [113]. We have resolved this issue in our previous publications using a finite element electro-thermal model, which solved the charge and heat flow equations with an added restriction of the current flow in the circuit [18], [114], [134]. The simulation results reproduce the NDR shape of the *I-V* in Fig. 3.6 (a) and give the temperature distribution in Figure 3.6 (b) above. The maximum temperature within the functional layer at point C was 1480 K with the 1400 K isotherm marked with the white dashed line.

The maximum temperature is reached at about 1/3 of the functional layer thickness below the top electrode with the diameter of the 1400K isotherm being 72 nm wide at the widest point. For the purpose of illustrating the ion transport patterns in the device, one can consider a very simple model below. Let us assume that the diffusion coefficient of Ta is constant above a critical temperature of 1400K and zero below. All Ta ions within the isotherm would move downward with approximately the same velocity creating a gap against the top electrode and local enhancement of Ta density at the lower boundary of the isotherm replicating the two most prominent features of the experimental distribution. The new

distribution would result in an increase of the field and temperature within the gap driving more Ta downward. At the same time, the field and temperature in the accumulated region would decrease slowing the motion in this part of the filament. The argument presented here parallels the results of numerical simulations of Larentis *et al.* [72].

The drift of Ta ions in the electric field alone could not produce the observed changes of resistance. Increase of the Ta concentration in the middle and lower section of the functional layer would locally increase the electrical conductivity but the fully oxidized gap in the filament would more than compensate for this effect. Vertical motion could not explain the observed Ta distribution either. As can be seen from Figure 3.4 (b), the Ta depleted region close to the top electrode is much wider than the width of the filament core. Both observations find a natural explanation if Ta is also moving laterally due to the Soret effect. This interpretation was offered based on observation of the Ta-depleted cylinder around the filament core in plan-view images [50] and is in good agreement with the cross-sectional observations presented here.

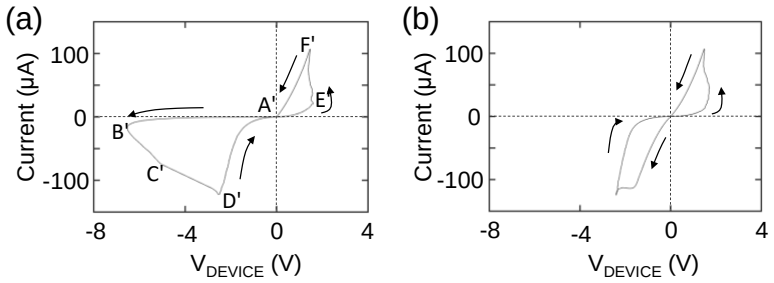

Figure 3.7 (a) *I-V* trace of the electro-formation process in the negative polarity. (b) Switching *I-V* of fully-formed device after forming in the negative polarity.

The model of formation presented above has also been tested using devices electro-formed with negative bias applied to the top electrode. The device formed into the HRS, similar to the case with positive bias formation, exhibited reversed polarity of switching (forming and switching *I-Vs* are shown in Figure 3.7). The forming process with negative bias applied to the top electrode (Figure 3.7) consists of two sweeps similarly as the one in positive bias (Figure 3.2 (b)). The *I-V* of as-fabricated *I-V* starts at point A and proceeds to threshold voltage at B and the point of electro-formation at C. A→B→C trace

is characteristic of S-NDR device with no permanent changes taking place. With current increasing past C, the device evolves along the load line to D. At this point, the device is partially formed to state close to the HRS in the switching characteristics (shown in panel (b)). To complete forming, a sweep in positive direction A→E→F is necessary. The resulting switching *I-V* and the direction of resistance changes are shown in panel (b). The SET process now occurred in positive bias with RESET under negative voltage in agreement with previously published data [135]. The corresponding cross-sectional HAADF-STEM image is shown in Figure 3.8 and colorization was generated in the same fashion as Figure 3.4 (b). In addition to the large scale color variation and the features within 3 nm next to the interfaces as discussed earlier, there is a general trend of the contrast in Figure 3.8 transitioning from red to blue from left to right in the image, which was also interpreted as an artifact the thickness difference in the sample because the front and back surfaces of the sample are not strictly parallel to each other. The results are consistent with those obtained in positive polarity in terms of filament diameter and Ta-rich composition of the core. The difference, however, is the location of the Ta-depleted gap, which in this device is formed at the bottom electrode. This terminal was positively charged during formation and the electric field induced Ta ions to move upwards. Similarly, as in Figure 3.4 (b), one can notice the Ta-depleted area around the Ta-enriched core indicating the inward lateral motion of Ta. Different switching polarities have been reported previously and interpreted as due to a gap appearing at the top and bottom electrodes [23], [136], [137] in agreement with observations here.

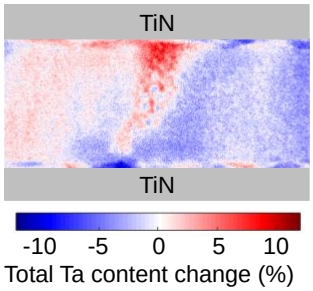

Figure 3.8 Cross-sectional HAADF-STEM image of device formed with negative voltage applied to the top electrode.

The distribution of Ta within the filament explains why the device forms to the HRS. For both polarities of applied bias during forming, Ta ions move away from the positively charged electrode creating a fully

oxidized gap in the filament. Hence, the high resistance state. The width of the gap is 5-10 nm and is much too wide to allow for direct tunneling of electrons. Highly nonlinear *I-V* curve after formation is similar to the negative differential resistance (NDR) part of an unformed device *I-V* curve and could be due to the Poole-Frenkel transport mechanism leading to thermal runaway. The difference between the *I-V* characteristics after electro-formation and that in as-fabricated devices is likely due to a higher electric field within the gap and a much smaller diameter hot spot induced by the small permanent Ta-rich filament.

3.5 Physical Changes in the Filament due to Switching

It is certainly of interest to assess the physical changes in the device structure resulting in the change of electrical characteristics. The HAADF-STEM of Ta distribution in a device that was formed in positive bias and switched to the LRS is shown in Figure 3.9 (a). Application of bias to the device in the HRS results in the SET process with increased conductance and the *I-V* becoming almost linear (Figure 3.2 (c)). The overall distribution is similar to that in device formed in positive bias but not switched (Figure 3.3 (b)): the filament is the same diameter and most of the Ta-depleted gap next to the top electrode remained unchanged. A closer inspection of Figure 3.9 (a) reveals a thin sub-filament connecting the filament core with the top electrode (marked by the arrow in the figure). Formation of the sub-filament in the gap has been considered before but lacked supporting evidence [137], [138]. The shape of the entire filament resembles a “martini glass” and is likely due to the crystallite present at the top of the filament (Figure 3.4 (c)). The diffusion coefficient in the crystalline Ta_2O_5 is expected to be much lower than that in amorphous material and ions should drift preferentially at the periphery of the crystallite where the temperature and diffusivities are high.

It is clear that when the device is switched from the HRS to the LRS, the Ta-deficient region (where the gap is located when the device is in HRS) remains depleted. Instead, the Ta content in the surrounding amorphous region around it increases. This is contrary to many publications which pictured the filament in the LRS as a uniformly doped cylinder of a semiconducting material or a metallic nanowire with a

constant diameter [72], [100]. This clearly cannot be the case as the gap in a symmetric filament could be produced in either polarity of bias. VCM devices, which TaO_x is a prime example of, undergo RESET preferentially in one polarity. This is an essential feature differentiating bipolar VCM devices [92] from unipolar ones operating in thermochemical memory (TCM) mode [42]. The underlying ion motion was suggested to be the oxygen vacancies moving along the electric field (VCM) or metal cations moving laterally and driven by the temperature gradient (TCM). Since we do find both driving forces and both types of atoms to be mobile in the prototypical VCM device, the difference between bipolar and unipolar switching is reduced to a difficulty in forming a strong filament in the gap. It is possible to do but requires higher voltages than the ones necessary for switching. The process leads to reversal of switching polarity [23], [136] and is discussed below.

The sub-filament shown in Fig. 3.9 (a) does not exhibit as strong contrast in HAADF-STEM images as the core of the filament, which is expected. Figures 3.4, 3.8, and 3.9 show the change of Ta content integrated over the thickness of the TEM lamellae. It is clear that the contrast is proportional to the diameter of the feature and the concentration within. Given the same composition of the sub-filament as that in the core, the contrast would be 5 times smaller than one seen in Figure 3.4. Mapping elemental distributions at this scale will require a much thinner specimen.

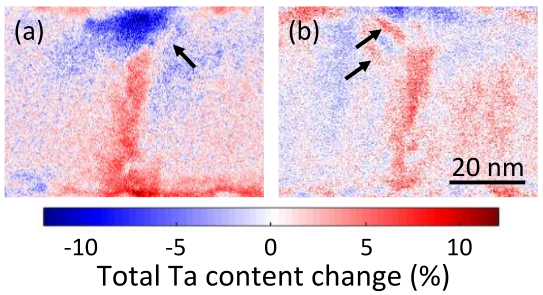

Figure 3.9 Cross-sectional HAADF-STEM images of devices that were (a) formed in positive bias and switched to LRS and (b) formed in positive bias and its polarity reversed in negative bias. The device was left in the LRS.

Figure 3.9 (b) shows the change of Ta in a device that has been formed under positive bias, followed by switching to the LRS in negative voltage. At this point, the device was subject to voltage pulses with negative polarity and increasing amplitude until device switching polarity was reversed. The gap in such

devices is expected to be located at the interface between the functional layer and the bottom electrode while there should be no gap at the other interface [23], [136]. The experimental observation of Ta distribution is somewhat different. The original Ta depleted region close to the top electrode still remained although the interconnecting segment at the periphery of the crystallite is now much stronger than in the LRS in Figure 3.9 (a). Also, the amount of the accumulated Ta close to the bottom interface is much lower, but the formation of a gap is not apparent.

Chapter 4 Quantification of Compositional Runaway during Electroformation in TaOx Resistive Switching Devices

4.1 Introduction

In Chapter 2 and 3, it was proposed that the temperature gradient created within NDR, drives the elemental redistribution to form the permanent filament and further lowers the device resistance. The typical forming *I-V* only shows one rapid transition in device resistance: the threshold switching is immediately followed by compositional change. In this chapter, we separate threshold switching from the compositional change and quantify the dynamics of the ion motion during for formation of the conductive filament.

4.2 Experimental

1) Device Fabrication

The devices used in this chapter are the same from those in Chapter 3 (Figure. 3.1 (a)).

2) Electrical Characterization

A series of constant source voltage pulses with amplitudes between 8 and 16 V was applied to a batch of as-fabricated devices as means of electroformation. An oscilloscope was connected in series to the device and the load resistor to monitor the current through the circuit. After the electroformation, all devices showed standard resistive switching characteristics. The integrated on-chip R_{LOAD} was 37 kΩ for all devices preventing the threshold switching and allowing for observation of the entire *I-V* characteristics. Note that the devices used in this study are from two different batches with slightly different R_{LOAD} and threshold voltage, but no significant difference of *I-V* behavior.

3) STEM analysis

The plan-view FIB procedure is the same as described in Chapter 3. The STEM HAADF images were collected with an FEI Titan 80-300 microscope operating at 300 keV with Fischione model 3000 HAADF detector. The camera length was 48 mm to keep the HAADF collection angle greater than 50 mrad for true Z-contrast.

4.3 The Compositional Runaway

Figure 4.1 (a) shows the current response caused by source voltage pulses (V_{SOURCE}) of 8, 10, 12, 14, and 16 V. After a rapid jump at the beginning of each pulse, the current gradually increases until a sudden rise followed by the saturation at a higher current level. At the end of the pulse, the device is formed and exhibits resistive switching with SET and RESET voltages below 1.6 V. The time from the beginning of the pulse to the step-like current increase will be referred to as the incubation time (τ). Fig. 4.1 (b) shows the forming *I-V*'s corresponding to the transients (colored lines corresponding to load lines at different V_{SOURCE}) and a forming curve obtained with the quasi-static voltage sweep (black curve). Let us consider the two traces corresponding to V_{SOURCE} =12 V (represented by the purple lines in Figure 4.1(a) and (b)). The initial current increase in (a) corresponds to 1 μs rise time of the voltage pulse with the current evolving along the intrinsic *I-V* to the point marked A in (b) (the *I-t* traces of V_{SOURCE} = 8 V and 10 V started at ~10^{-4} s because the oscilloscope loses time resolution when the pulse becomes very long (~100 s)). After the initial current increase, the current stabilizes within several thermal time constants of the device (estimated at ~2 μs [139]). Note that up to this point the changes of current are volatile. The NDR *I-V* characteristics were simulated by the finite-element electro-thermal model (dashed magenta curve in Figure 4.1 (b)) using experimental data for the conductivity as a function of temperature and field [18]. The resulting temperature distribution at point A is shown in Figure 4.1 (c) [18]. The temperature at the center is 860 K allowing for diffusion of ions. The highest temperature gradient in the structure is 3×10^{9} K/m.

After the first 10 μs and before the step-like current increase, the resistance of the device slowly decreases. This effect is easiest to see for transients collected at low pulse voltages where the current is low, and

resistance change higher than at high voltages. These changes are non-volatile, and we interpret them as due to ion motion. The likely mechanism driving the motion is thermophoresis (Soret effect) [42], [48]. We have recently obtained evidence of Ta ions moving toward the hot spot created by the current constriction within NDR [140]. As the Ta moves towards the hot spot, the material in the center of the forming filament becomes more conductive. The higher conductivity further increases the current flow and thus the temperature as well as the temperature gradient, which in turn provides a higher driving force for the diffusion of Ta. In other words, the temperature and the chemical composition of the filament form a positive feedback loop. In the beginning of this positive feedback loop, the motion is slow due to relatively low temperature and temperature gradient. The increase of device temperature and conductivity eventually leads to a point of compositional runaway. Positive feedback loops produce the characteristic slow-fast-slow dynamics reported in the case of thermal runaway [114] and evident in Figure 4.1 (a). The runaway can either lead to destruction of the device or, if current is limited by an appropriate load, a stable state where the effects of temperature gradient are balanced by the Fick's flux of Ta diffusing down the concentration gradient. In this case, with the protection of the load resistor, the compositional runaway caused the Ta to rapidly diffuse into the hot spot forming the conductive filament, and the concomitant fast decrease in device resistance gave rise to the step-like current increase observed from the *I-t* traces in Figure 4.1(a). When the current is stabilized again after this runaway, a Ta-rich filament is formed in the device.

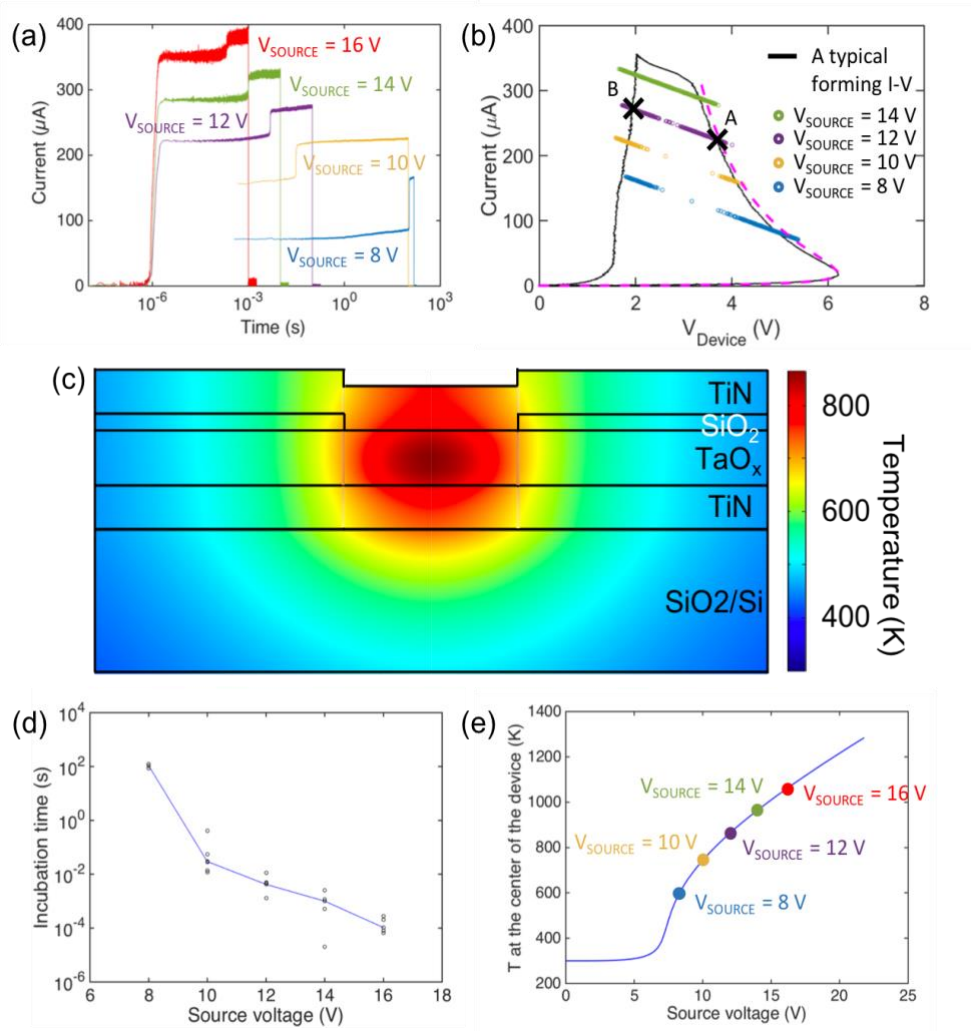

Figure 4.1 (a) Examples of electroformation using constant source voltage pulses with different amplitudes. (b) A typical forming device *I-V* (black solid line) and simulated S-NDR prior to the compositional runaway (magenta dashed line). The colored lines are the *I-V* replotted from the *I-t* plots from (a). The slope of the colored lines is the load resistance R_{LOAD} = 37 kΩ. (c) Simulated cross-sectional temperature distribution of the device in the NDR mark as A in (b). (d) Incubation time (τ) distribution of electroformation as a function of source voltage from all the devices tested. The blue line connects the median points of each source voltage. (e) Simulated steady-state maximum temperature (at the center of the device functional layer) as a function of source voltage.

One should also note that the voltage across the device corresponding to increasing source voltage decreases from 5 V at V_{SOURCE} = 8 V to 3.7 V at V_{SOURCE} = 14 V. The incubation time is shorter at lower electric filed. This is a direct evidence that electro-formation is not caused by the electric field as typically assumed [26]. The reason for shorter incubation is likely increasing dissipated power and temperature within the device. In order to better understand the role of temperature, we attempted to quantify the $\tau(T)$ dependence.

The *I-t* traces in Figure 4.1 (a) were used to determine τ. Figure 4.1 (d) shows the dependence of τ on V_{SOURCE}. For each of the starting points on the quasi-DC *I-V*, the temperature distribution within devices was simulated and the data were re-plotted as a function of T_{MAX} (maximum temperature in the device's functional layer) The dependence of T_{MAX} on V_{SOURCE} is shown in Figure 4.1 (e).

Figure 4.2 shows the relation between forming rate ($r = \tau^{-1}$) and $1000/T_{MAX}$ which appears as a straight line on semi-logarithmic plot. The dependence can be fitted by Arrhenius relation:

$$r = r_0 \exp\left(-\frac{E_a}{k_B T_{MAX}}\right),$$

with the activation energy of 1.4 ± 0.4 eV. The activation energy extracted here agrees with the 1.0 eV reported by Larentis *et al.* within the margin of error, which describes the SET/RESET process [72].

Previous studies suggested motion of either O or metal ions during the electroformation of metal-oxide based resistive switching devices [48], [140]. Gries *et al*. reported the diffusion activation energy for O and Nb tracers in crystalline Ta_2O_5 to be 1.6 ± 0.2 eV and 3.6 ± 0.4 eV [119]. These values are higher than our estimate, but the activation energies in amorphous metal-rich TaO_x are expected to be lower than in crystalline material.

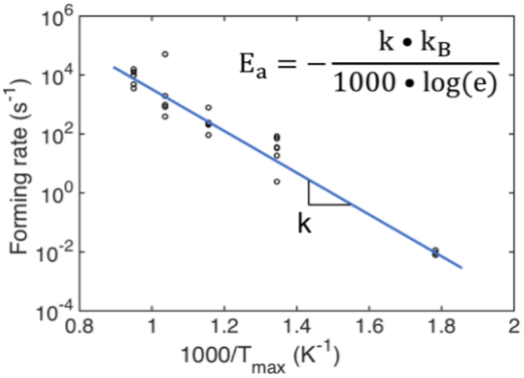

Figure 4.2 A plot of forming rate (τ^{-1}) versus the reciprocal of the maximum temperature within the device corresponding to different source voltages (black dots) and a fitted line by linear regression (blue solid line).

This result indicates that the formation of the filament requires elevated temperatures in order to provide activation energy. The mechanism proposed here would not result in formation of the filament at room temperature.

4.4 STEM Assessment of the Electroformation

To test our hypothesis of the two-step electroformation, plan-view High-Angle Annular Dark-Field (HAADF)-STEM images were collected from a device that was biased into NDR region but not allowed to form (*I-V* sweep up to 90 μA, Figure 4.3 (a)) and a device after electroformation (Figure 4.3 (b)). The two samples in Figure 4.3 were both thinned down to electron transparency leaving only the functional layer of the device. The STEM samples had area larger than the entire active area of the device. Comparing the two images, it is clear a bright spot of approximately 10 nm diameter appeared in the center of the device after the compositional runaway but was absent in the device biased into NDR for short time only. The dark wavy pattern in the functional layer throughout the TEM samples was attributed to the density variation due to the self-shadowing effect during the deposition of the film and is believed not caused by the forming event [107]. The HAADF signal intensity scales roughly with the square of the atomic number of the elements present in the material [83]. Since the region of interest in the sample contains mainly Ta and O, and the square atomic number of Ta (Z=73) is 11~108 times higher than the rest of the elements (Z=7, 8, 22), the contrast in the HAADF images corresponds to the distribution of Ta in the functional layer. The observation of the Ta-rich filament in formed device confirms that the fast decrease of device resistance during the second step of the electroformation is the compositional runaway of Ta forming the conductive filament.

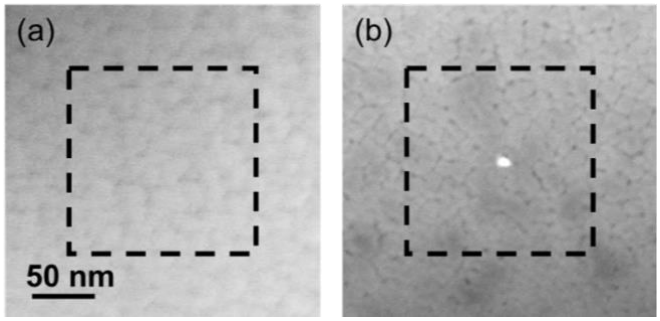

Figure 4.3 Plan-view images on the functional layer of (a) a device after NDR up to 90 µA, and (b) a device after electroformation with an *I-V* sweep. The black dashed boxes indicate where the device is defined.

Due to the limitation of the sample geometry and the imaging mechanism, the plan-view HAADF images cannot provide insights on the vertical driving forces (i.e. electric field) and the motion of O, which has been argued by many to be involved in the electroformation and resistive switching of metal-oxide resistive switching devices [26]. Our results do not exclude the role of electric field and O motion during the electroformation, but they demonstrate the importance of lateral motion induced by the temperature gradient.

4.5 Device Resistance after Electroformation

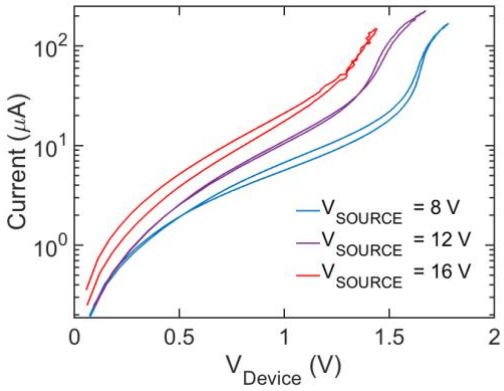

Figure 4.4 *I-V* sweeps of 3 devices measured after electro-formation performed at V_{SOURCE} = 8, 12, and 16 V, respectively.

At this point, it appears that the thermally activated process responsible for the electro-formation is the diffusion of Ta ions. It is possible, however, that the filament formed at different values of V_{SOURCE} is significantly different. This could affect the activation energy and possibly the interpretation of results. In order to test this hypothesis, we have performed series of post-forming *I-V* sweeps (note that devices were always formed to high-resistance state). Figure 4.4 shows the *I-V* curves from 3 devices directly after electroformation with V_{SOURCE} = 8, 12, and 16 V, corresponding to the forming power of 39, 84, and 112 μW, respectively. All three *I-V*'s have similar functional dependence and values within a factor of 2 of each other up to about 1.2 V. At this point, the current increases fast with voltage and the *I-V*'s diverge somewhat but do not differ more than an order of magnitude. Devices formed at higher power, are more conducting similarly as has been reported for conductance of the LRS. These, however, are relatively minor changes indicating the same operating mechanism. We interpret the difference as due to higher filament temperature and temperature gradient at higher forming power. This activates more Ta available for diffusion and also provides a higher driving force for the diffusion into the filament thus forming a stronger filament.

Chapter 5 Electrode/Functional Oxide Interdiffusion during Electroformation in TaOx Resistive Switching Devices

5.1 Introduction

Resistive switching in TaO_x is bipolar [14], [22], and, therefore, was thought to occur by the valence change mechanism. However, the experimental results obtained so far indicate that the distinction between VCM and TCM devices is not as clear as it was initially thought. The TaO_x devices were shown to exhibit attributes typically associated with TCM mechanism such as threshold switching and formation of the hot spot [18], [46]. In addition, tantalum ions were shown mobile [37], possibly accounting for higher fraction of composition change in the electro-formation process than the motion of oxygen [50], [51]. Scanning transmission electron microscopy (STEM) revealed that the Ta moves both laterally driven by the temperature gradient and vertically along the field direction [51]. On the other hand, STEM data on O distribution has not been as conclusive.

In this chapter, we set out to test whether there is an exchange of O between functional layer and the electrode during the electro-formation process in TaO_x-based devices as envisioned by the Valence Change Mechanism. The elemental maps after electro-formation have been collected by Energy Dispersive X-ray Spectroscopy (XEDS) mode of STEM on cross-sectional samples. The results allow for validation of a unified VCM/TCM model of electroformation.

5.2 Experimental

1) Device fabrication

The devices used in this chapter are the same as the ones used in Chapter 3 and 4.

2) Electrical characterization

The device corresponding to the forming *I-V* in Figure 5.3(a) and the STEM images in Figure 5.5 was formed by one *I-V* sweep from a voltage supply in positive polarity with a load resistor (R_{LOAD} = 58 kΩ) in series. The device in Figure 5.3(b) and Figure 5.7 was formed the same way as the one from Figure 5.3(a) and Figure 5.4, but with a negative forming polarity and R_{LOAD} = 60 kΩ.

3) Device simulation

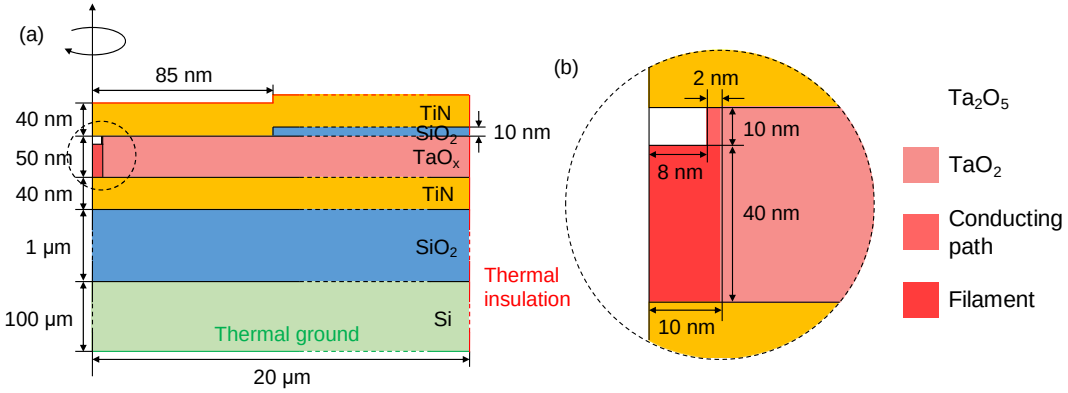

Figure 5.1 Schematic diagram of the device structure used for electro-thermal simulation.

The device simulation used in this chapter assumes the same charge transport and heat flow model as the previous chapters, but with a different geometry shown to account for the filament and the gap in the device after forming. Figure 5.1 shows the axially symmetric structure used in the simulation of the temperature distribution and *I-V* characteristics for the device in high resistance state. The bottom and top electrodes are both TiN (40 nm) sandwiching the TaO_2 functional layer (50 nm). The active area of the device is defined by the circular opening in the SiO_2 with 85 nm radius. The radius was chosen to give the area of the corresponding circle equal to an area of 150x150 nm_2 square. The device structure is heat sunk by the 1 µm SiO_2 and 100 µm thick Si substrate which replicates the heat flow of the actual structure. The gap in the filament is made, in part, of Ta_2O_5 plug (radius = 8 nm), which is assumed to be electrically insulating. The plug is intended to mimic the Ta_2O_5 crystallite seen in the STEM images of the filament [51]. The current was flowing around the crystallite between the top electrode and the trunk (conductive path, 2 nm thick in cross-section) with the geometry of the sub-filaments reported in the ON state. The trunk of the filament

was assumed to be 10 nm in radius, which is thicker than the TEM observation (radius = 4 nm on average). Since the filament is a lot more conductive than the gap and the surrounding sub-filament, its radius has little effect on the temperature, and it simplifies the geometry for the ease of the simulation. The electrical conductivity of the sub-filament and the TaO_2 was represented by the Poole-Frenkel formula the same as Equation 2.3 and 2.4. The values of the activation energy ($\frac{E_d+E_t}{2}$) and the pre-exponential factor ($q\mu N_c(\frac{N_d}{N_t})^2$) used in this work are listed in the table below:

Material	Pre-exponential factor (S/m)	Activation energy (eV)
TaO_2	$2.1x10_4$	0.13
Conductive path	$6x10_3$	0.30

Table 5.1 Parameters used in Poole-Frenkel conduction.

All the material parameters used in the simulation are listed in Table 5.2:

Material	Density ($kg \cdot m^{-3}$)	Thermal conductivity ($W \cdot m^{-1} \cdot K^{-1}$)	Electrical conductivity ($S \cdot m^{-1}$)	Heat capacity ($J \cdot kg^{-1} \cdot K^{-1}$)	Relative permittivity
Si	2329	130.0	N/A	700	N/A
SiO_2	2200	1.4	N/A	730	N/A
TiN	5210	5.0	5E6	545	4
Ta_2O_5	8200	4.0	0	174	22
TaO_2[141]	10000	4.0	See Table 5.1	174	22
Conductive path	10000	4.0	See Table 5.1	174	22
Filament	16650	57.5	1e6	140	1

Table 5.2 Material parameters used in this simulation.

The device in this simulation undergoes an *I-V* sweep from a voltage source with a load resistor in series (R_{LOAD} = 50 kΩ). The simulated *I-V* characteristics of the device before and after forming the filament (in positive polarity) are shown in Figure 5.2.

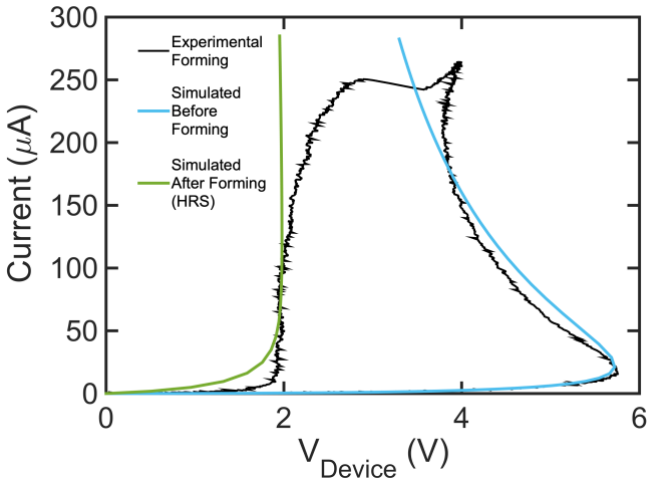

Figure 5.2 The experimental *I-V* of forming (black), simulated *I-V* of before (blue) and after (green) forming, respectively.

4) Materials Characterization

The FIB procedure for cross-sectional lift-out is the same as described from Chapter 3. The STEM images shown in this report were collected on a FEI Talos F200X S/TEM operating at 200 keV. The HAADF-STEM images were collected using with an inner and outer collection angle of 76.7 and 200 mrad, respectively. The EDS spectrum images were collected using the four detector "SuperX" energy dispersive X-ray spectrometer system with a probe current of 1nA.

5.3 Temperature Simulation of Devices before Forming

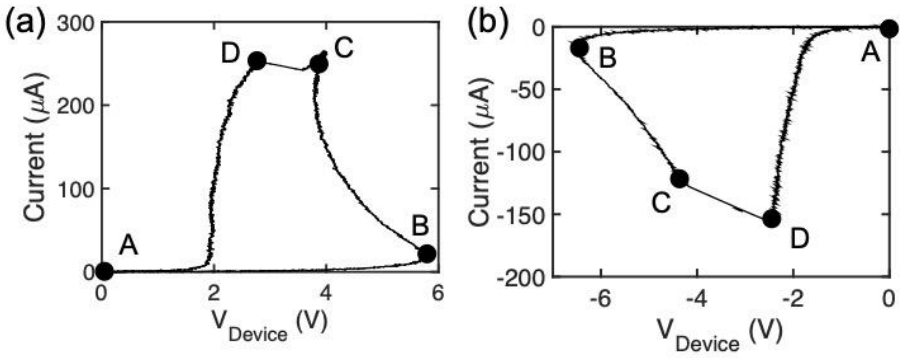

Figure 5.3 (a) Forming *I-V* of a device under positive bias with R_{LOAD}= 58 kΩ. (b) Forming *I-V* in negative polarity with R_{LOAD} = 60 kΩ.

The electro-formation of devices used for elemental mapping was performed with one voltage sweep with the resistor in series to the device. Figure 5.3 (a) shows a quasi-static forming *I-V* in positive polarity (defined as top electrode connected to the positive terminal). The A-C part of the curve corresponds to the S-NDR characteristics of a threshold switch and can be traced back and forth many times without any lasting changes to the device. Within NDR, the temperature and current distributions become constricted with the maximum temperature reaching ~920 K at the power dissipation of 1024 µW at point C (Figure 5.4 (a)). This allows for diffusion and permanent changes of *I-V* to occur as the device evolves between C and D. Upon reaching point D, the device is no longer evolving and the segment D to A is the result of downward sweep of the source voltage. At this point, the device is formed to HRS of the memory switching. Figure 5.3 (b) shows the forming *I-V* characteristics in negative polarity. The curve is quite similar to that in Figure 5.3 (a) only forming at a lower power dissipation with a maximum temperature ~620 K at the power dissipation of 533 µW at point C (Figure 5.4 (b)). Such behavior was consistently observed for many devices on multiple chips processed with slightly different conditions. The different power dissipation required for forming is likely due to the fact that the two interfaces were processed differently. The noise in the *I-V* shown in both Figure 5.3 (a) and (b) is in the direction of the load line imposed by the load resistance (parallel to the forming transition in the *I-V*), because the original electrical data was from the time trace collected by the oscilloscope whose noise is perpendicular to the time axis, and when the *I-t* trace was replotted into *I-V*, the data noise became parallel to the slope of the load line.

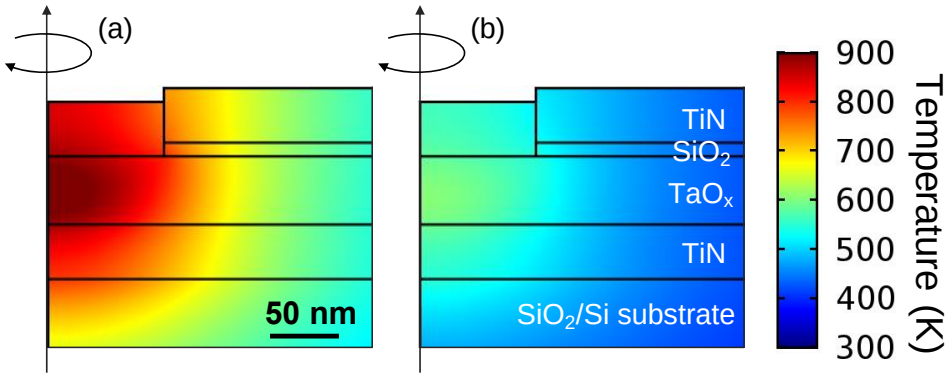

Figure 5.4 Simulated temperature distribution of a device under bias before the filament forms during (a) positive polarity forming and (b) negative polarity forming. Corresponding points on the *I-V* are marked by "C" in Figures 5.3(a) and (b).

5.4 Ionic Inter-Diffusion after Electroformation

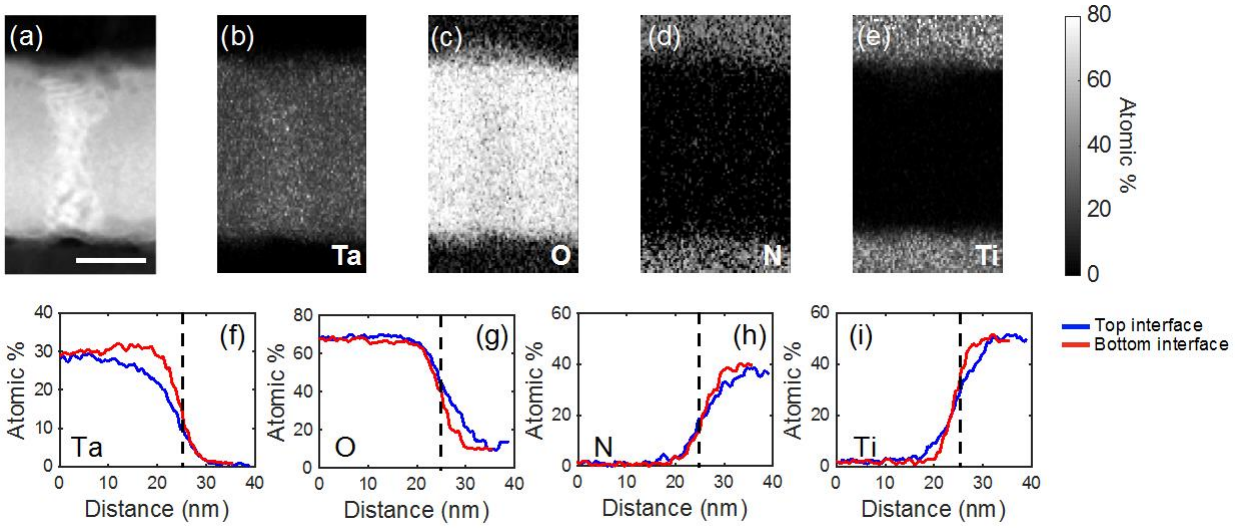

Figure 5.5 (a) Cross-sectional HAADF-STEM image of a filament in a device formed in positive polarity. The scale bar corresponds to 20 nm. (b)-(e) Ta, O, N, and Ti elemental maps collected by EDS from the same location as (a). The atomic percentage of the elements in (b)-(e) is indicated in the color scale. (f)-(i) Compositional line profiles of Ta, O, N, and Ti corresponding to images (b)-(e). The line profiles are integrated over ~20 nm laterally in the center of the images. The horizontal axis (distance) starts from the middle of the functional layer (distance=0 nm) to the top (distance=39 nm, blue line) and to the bottom (distance=35 nm, red line) of the images. The dashed vertical lines mark the position of the interfaces.

Figure 5.5 presents the cross-sectional EDS elemental maps and the corresponding line profiles of Ta, O, N, and Ti, as well as a reference High Angle Annular Dark Field (HAADF)-STEM image of the same area of the device formed in positive polarity. The dissipated power after filament formation was 750 µW (point D in Figure 5.3 (a)). The bright contrast in the middle of the HAADF-STEM image (Figure 5.5 (a)) corresponds to the filament, the width of which is ~6 nm in the middle and up to 20 nm near the two interfaces. HAADF images are formed by electrons scattered to high angles primarily by heavy ions. Accordingly, the HAADF intensity corresponds to the local Ta density since the light O atoms contribute less than 1% to the signal [83]. In other words, the HAADF images indicate that the filament has excess accumulated Ta. Our recent publication estimated that the filament contains 80% higher Ta and 60% less O compared with the as-deposited functional $TaO_{2.0\pm0.2}$ layer [51]. This finding was based on the

interpretation of HAADF-STEM intensity and Electron Energy Loss Spectroscopy (EELS) maps of oxygen distribution.

Figure 5.5 (b) shows the EDS map of Ta distribution in the proximity of the filament. The bright contrast in the center of the map indicates a local increase of Ta content at the location the filament. In addition, one can also note Ta depletion at the top of the column and close to the oxide/electrode interface. The differences of Ta distribution across both interfaces can be quantitatively assessed in the line profile in Figure 5.5 (f). The profiles have been obtained by integration of EDS signal along 20 nm wide vertical strip at the position of the filament with the blue line depicting the change of X-ray emission across the top interface and red line corresponding to the lower one. It is quite clear that the Ta content close to the anode is 15 at. % lower than that close to the cathode. One needs to note here that the filament takes up only a portion of the thickness of STEM sample and the intensity in the maps and line plots does not directly translate into composition of the filament.

The Ta EDS map agrees with the interpretation of the HAADF-STEM images in Chapter 3 (Figure 3.4 (a) and (b)) and the one shown in Figure 5.3 (a). Net transport of Ta downward is consistent with the electric field acting on positively charged Ta ions. Depletion of Ta close to the top electrode forms an insulating gap in the filament (referred to as the “gap” from now on). Since the gap is the most resistive segment of the current path through the device, most of the voltage drop and the power dissipation occurs there. After the formation, the temperature distribution along the filament becomes inhomogeneous with the high temperature region located next to the anode. This temperature distribution is discussed in more detail below.

The variation of oxygen EDS signal (if any) within the functional layer in the EDS O map (Figure 5.5 (c) and (g)) is at the noise level, although it was previously determined that the gap (Ta_2O_5) contains ~25 at. % more O than the matrix (TaO_2) [51]. This is likely due to the fact that the X-ray signal collected from the gap area is integrated throughout the thickness (~70 nm) of the sample over the gap area, including the gap (~20 nm) and the surrounding TaO_2 matrix (~50 nm), which produced an averaged signal with the O content

close to that of the matrix. Laterally there was no visible signal attenuation while vertical line profiles (red and blue lines in Figure 5.5 (g) are right on top of each other. However, one can notice a “dome” with increased oxygen content within the TiN top electrode directly above the filament (Figure 5.5 (c)). The width of the dome is ~35 nm, and the height is 8 nm.

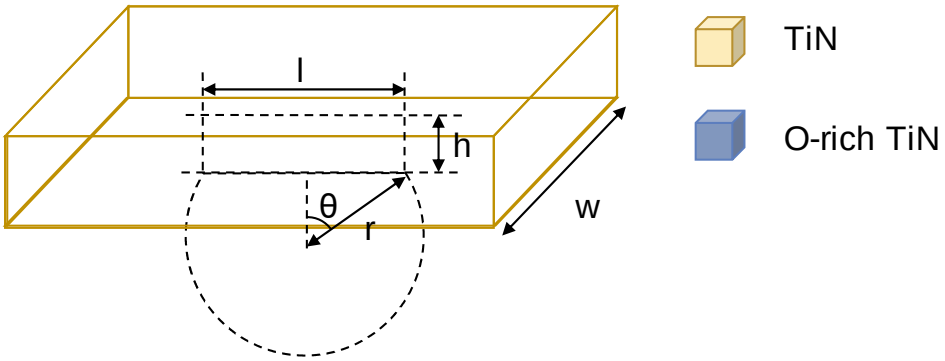

Figure 5.6 Schematic of the O-rich dome (blue) in the TiN electrode (yellow).

Figure 5.6 shows a schematic illustration of an O-rich dome in the TiN electrode that was shown in Figure 3 (c) in the main text. The dimensions of the dome have been estimated at l = 35 nm and h = 8 nm with the thickness of the sample w = 70 nm. According to the composition measured in Chapter 3, the composition of the core of the filament, the as-deposited TaO_x layer, and the as-deposited TiN electrode are $TaO_{0.4}$, $TaO_{2.0}$, and $TiN_{0.8}O_{0.2}$, respectively.

The shape of the dome was approximated as a spherical cap (Figure 5.6) with r=24 nm, and θ = 49˚. The volume of the O-rich dome is about 4300 nm^3 while the largest cross-sectional area of the done is 200 nm^2. The measured average O signal over the dome area is 15 at. % whereas elsewhere in the TiN is 11 at. %. However, this does not immediately translate into the O concentration in the O-rich dome since the dome fills only a fraction of the TEM sample. The O concentration in the dome can be calculated:

$$15\ \% \cdot w \cdot A_{dome} = V_{dome} \cdot [O]_{dome} + (w \cdot A_{dome} - V_{dome}) \cdot 11\ \%$$

$$\Rightarrow [\mathrm{O}]_{\mathrm{dome}} = 24\ \mathrm{at.}\,\%$$

Likewise, the $[N]_{dome}$ and $[Ti]_{dome}$ can be calculated as 27 at. % and 49 at. %, respectively, giving the composition of the O-rich dome as approximately $TiN_{0.5}O_{0.5}$.

The corresponding oxygen line profile (Figure 5.5 (g)) extends deeper into the anode than the one at the interface with the cathode. We interpret this oxygen-enriched dome as due to interdiffusion between TaO_x and TiN driven by the concentration gradient. The gap and its proximity are at the highest temperature within the filament with a large concentration change across the interface. The alternative interpretation with the oxygen motion driven by the electric field is questionable at least since the electric field inside the metallic electrode is negligibly small.

In order to confirms if the O motion is conserved in this process, an estimate of O gain and loss in different parts of the device active area is then calculated. To simplify the problem, the "trunk" of the filament is approximated as a cylinder with radius r_f = 4 nm, height h_f = 40 nm, and the gap is approximated as a cylinder with radius r_g = 10 nm, height h_g = 10 nm. Therefore, the volume of the filament V_f ~ 3500 nm_3, the volume of the gap V_g ~ 3100 nm_3. Note that changes of O in the trunk of the filament and the gap region are less visible in the cross-sectional maps of O shown in Figure 5.5(c) than what was measured in Chapter 3, and the calculation of the amount of O change in these two locations is based on the measurement from Chapter 3.

Material	Density (g/cm3)	Molar Mass (g/mol)
Ta_2O_5	8.2	441.9
TaO_2	10.0	212.9
TiN	5.2	61.9

Table 5.3 Density and molar mass of materials of interest [142].

- Oxygen loss from the trunk of the filament:

  The TaO_x composition originally was 1:2 and was reduced by 60 at. % (Chapter 3). The number of O atoms lost in the filament is: N_{Of} = 60 % • 2 • N_A • $V_f\rho_{TaO2}/M_{TaO2}$ = 6.8 x 10_4 atoms (N_A = 6.02 x 10_{23} is Avogadro's constant).

- Additional O that entered the gap region:

  At the gap, the composition changed from TaO_2 to Ta_2O_5 during electro-formation (Chapter 3). The O gain in the gap was $N_{Og} = (5 - 4) \cdot N_A \cdot V_g \rho_{Ta2O5}/M_{Ta2O5} = 3.5 \times 10^4$ atoms.

- Extra O that entered the TiN:

  The original composition of the TiN electrode was $TiN_{0.8}O_{0.2}$ (Chapter 3). After the filament was formed, the concentration of the O changed from 11 at % to 24 at. %. The amount of additional O in the dome: $N_{Og} = (0.24 – 0.11) \cdot N_A \cdot V_g \rho_{TiN}/M_{TiN} = 2.8 \times 10^4$ atoms.

This suggests that the gain of the O in the gap and the dome accounts for the O depleted from the filament. Therefore, one can conclude that during forming, most of the O redistribution occurred by vertical motion: O moved out of hot core of the filament driven by the electric field and accumulated in the gap and the anode. This result is quite different than the observation of a dark ring with the diameter of approximately 800 nm around a bright circle in X-ray absorption experiments by Kumar *et al*. [48]. The contrast in Kumar's work was interpreted as due to the out-diffusion of O from the hot core of the filament due to Soret effect. However, one needs to note that these experiments have been performed in conditions distinctly different from ours. The devices have been fabricated on Si_3N_4 membranes which thermally isolated the device structures [143]. This as well as an unusually high power dissipated during device switching (devices normally switched at 2V dissipating 500 μW but pulses of 5V have been used for stressing giving an estimated power of 5 mW) must have produced exceedingly high temperatures within the filament. This was also indicated by the electrode delamination reported by the authors. Lastly, one should note that, in Kumar's work, the first sign of the ring formation appeared after multiple switching cycles while our devices were examined only after electro-formation. The thermophoresis reported by Kumar *et al*. was an aftereffect of switching rather than the mechanism responsible for the formation of the filament.

In addition to O presence in TiN, one can note a matching N depleted volume at the location of the O-rich dome (Figure 5.5 (d) and (h)). This indicates that in the O-rich dome of the TiN, O replaced N in the TiN crystal lattice rather than forming interstitials or residing in grain boundaries [144]. Finally, a weak signal of

Ti in the gap of the filament was observed as well in Figure 5.5 (e) and (i), indicating that Ti diffused into TaO_x during the formation of the filament.

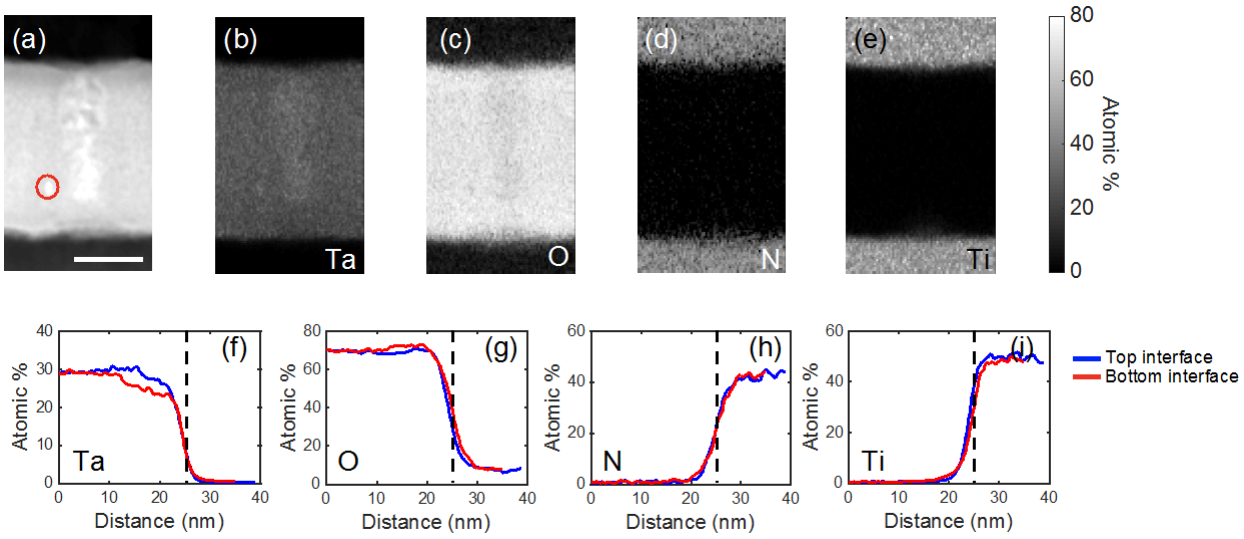

Figure 5.7 (a) Cross-sectional HAADF image of a filament in device formed in negative polarity. The scale bar corresponds to 20 nm. The feature marked by the red circle in (a) is caused by hydrocarbon contamination in STEM. (b)-(e) Ta, O, N, and Ti atomic maps collected by EDS from the same location as (a). The atomic percentage of the elements in (b)-(e) is indicated in the color scale. (f)-(i) Compositional line profiles of Ta, O, N, and Ti corresponding to images (a)-(e). The line profiles are integrated over ~20 nm laterally. The horizontal axis (distance) starts from the middle of the functional layer (distance=0 nm) to the top (distance=39 nm, blue plot) and to the bottom (distance=35 nm, red plot) of the images.

Figure 5.7 shows a set of images corresponding to the device used in Figure 5.3 (b) formed in negative polarity (380 µW power dissipation after the filament formation, point D in Figure 5.3 (b)). The elemental distributions are quite similar to those obtained in positive polarity and we will only point out the differences. The HAADF-STEM image in Figure 5.7 (a) shows similar Ta contrast within the filament and depleted region corresponding to the gap next to the bottom electrode. This is in line with the vertical motion of Ta being induced by the reversed electric field and consistent with the Ta distribution in Figure 5.5 (b) and (f). The Ta EDS map (Figure 5.7 (b)) and the line profile (Figure 5.7 (g)) correlate well with the HAADF-STEM. For the O distribution, one could expect O to have diffused into the TiN bottom electrode. However, no such dome could be seen in Figure 5.7 (c) and (g). This lack of interdiffusion is likely due to the lower power dissipation in negative polarity giving rise to the lower temperature at the gap and the adjacent interface. Likewise, the interdiffusion of N and Ti near the gap was too low to be detected.

5.5 Temperature Simulation in Devices Post-Forming

Comparing the two forming polarities, the observations from the EDS maps i.e. similar extent of Ta redistribution and much lower interdiffusion of all other elements at low power, can be interpreted in a following way. The hot spot created during threshold switching in TaO_x has a diameter of over 100 nm with the temperatures up to 920 K. The element with the highest diffusion coefficient is the Ta and it starts to move first accumulating within the hot spot. This process goes to completion approaching the steady state Ta distribution with the lateral Soret flux balanced by Fick's flux. During the process, the filament continuously collapses with the decreasing diameter and the appearance of the gap. The temperature distribution in the formed filament was assessed using a simple model shown in Figure 5.8 (a) and (b). The filament was assumed to be cylindrical with the diameter of 20 nm and metallic conductivity of 10_6 S/m. The diameter is larger than what was observed in the experiments to simplify the geometry. This does not affect the *I-V* behavior or the temperature distribution in the structure. In the gap region, filament contains an insulating plug with the diameter of 16 nm consisting of Ta_2O_5 where the conductivity was assumed to be 0. The plug leaves a thin layer of conducting material around its periphery. The geometry of the model reflects the geometry of the filament observed in the cross-sectional TEM with the plug corresponding to Ta_2O_5 crystallite [51]. The narrow channel around the gap defines the conducting path where the current flows around the gap corresponding to the HRS of the device after forming. Figure 5.8 (a) and (b) show the simulated temperature distribution within the filament and the gap after forming in both polarities and their corresponding dissipated power. The temperature at the top and bottom interfaces near the hot-spot for the two cases are 1310 and 920 K, respectively. The activation energy of O diffusion in TiN [145] gives the ratio of diffusion coefficients of about 1,000 or the diffusion length ratio of 30 between the positive and the negative polarities, making the interdiffusion in negatively formed device practically undetectable. Likewise, the lack of diffusion of N and Ti in the negatively formed device can be also explained as due to the lower temperature and hence the diffusion coefficient. Interdiffusion can occur when the collapse of the filament's size results in sufficient increase of the temperature at the anode/oxide interfaces as seen in Figure 5.8 (a).

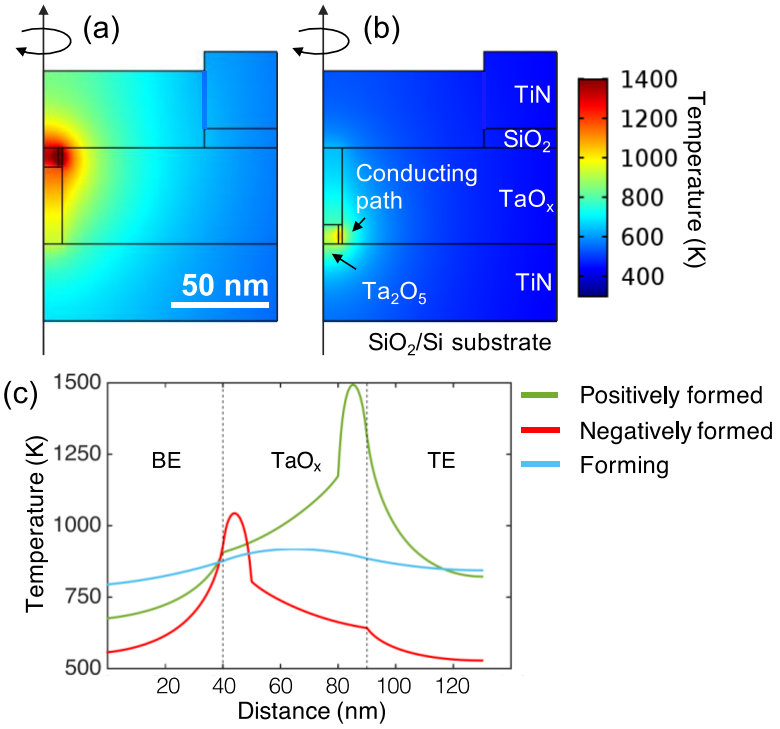

Figure 5.8 Simulated axi-symmetric cross-sectional temperature distribution of devices assuming a 20 nm wide filament is formed with a 16 nm thick gap on the (a) top (at power dissipation 750 μW) and (b) bottom (at power dissipation 380 μW), respectively. The 2 nm thick path conducts current since the gap is made of Ta_2O_5, which is an insulator. The views in (a) and (b) show only the magnified portion of the entire structure with an efficient heat sink (see Supporting Information for more details). (c) The temperature line profiles of positively (green) and negatively (red) formed devices along the vertical direction 9 nm away from the rotational axis (through the conductive path) at the power dissipation of 750 μW and 380 μW, respectively. The blue curve is the temperature line profile along the center of the device before the formation of the filament (point C of Figure 1(b)). The line profiles start at the bottom of the bottom electrode (BE, distance = 0), and ends at the top of the top electrode (TE, distance = 130 nm).

5.6 Discussion

Based on the interpretation from the STEM observation and temperature simulation, the behavior of our TaO_x devices is different from the original VCM model as well as most of the publications on this subject which interpreted electro-formation and switching as due to oxygen motion driven by the imposed electric field, the argument of which needs much more significant support in experimental observations. Even a brief inspection of *I-V* forming characteristics (Figure 5.3 (a) and (b)) indicates that forming does not occur at the highest electric field in the device which corresponds to the threshold voltage, or the start point of

NDR. Instead, it occurs when the dissipated power and temperature within the device is highest. The process is thermally activated with an activation energy of 1.4 eV rather than field-induced (Chapter 4). This eliminates the original VCM model as responsible for electro-formation in our structures. Whether it is O or Ta that moves first can be determined by comparing the temperature distribution before (Figure 5.2) and after forming (Figure 5.8) and the elemental maps in Figures 5.3 and 5.5. It is apparent that the heated area of the electrode post-forming is much smaller than before forming. If O was to initiate the electro-formation at point C in Figure 5.3 (a), the diameter of the O-rich dome should correspond to the temperature distribution in the threshold switch and be much larger than the size of the dome seen in Figure 5.5 (c). Also, the diameter of the O-depleted core in the functional layer seen in plan-view EELS in our recent publication [51] should be much bigger than the diameter of the final filament, which is in fact the same as that of the filament.

The interpretation in the paragraph above does not consider the lateral motion of O. We feel justified in doing so as no evidence of such motion was detected in our experiments. This is also in agreement with observations of Kumar *et al*. [48] These authors have reported oxygen redistribution at temperatures significantly higher than created during electro-forming processes used here, and no lateral O transport was detected in devices with low power dissipation.

It is unlikely that the oxygen ions within the anode can go back to the oxide during subsequent switching. The electric field at the interface does not appear to be the driving force for the initial exchange and should remain so. Moreover, there is no significant electric field in the metal electrodes. This indicates that during the lifespan of a device, the filament can continue to lose oxygen gradually changing switching *I-V* characteristics eventually making the filament too conductive to reset, or the electrode becoming too resistive inducing failure. Likewise, the diffusion of Ti into the Ta-rich filament is also of concern and should be properly controlled.

Chapter 6 Evolution and Failure of Conductive Filament due to Stressing

6.1 Introduction

So far, we know that the TaOx-based RRAM works as a hybrid mechanism that have both the characteristics of VCM and TCM due to its bipolar nature and the Soret effect during forming. From electroformation, a Ta-rich and O-poor filament together with an Ta_2O_5 gap form caused by the temperature gradient and electric field, and Ta-rich sub-filaments form and disappear under the imposed field during resistive switching. The O diffusion into the electrode during forming has been argued to be due to the temperature activated Fick diffusion rather than driven by the electric field and considered as a byproduct of forming rather than an essential process of the filament formation and the resistive switching.

Having properly understood the forming and switching of the TaO_x RRAM cell by S/TEM analysis of the conductive filament, we come to the last chapter of this dissertation to understand the evolution of the conductive filament and its endurance failure mechanism.

6.2 Experimental

1) Device Fabrication

The devices used in this study are the same as those from Chapter 3-5.

2) Electrical Characterization

Devices were formed with a 12 V constant voltage source with an R_{LOAD} = 37 kΩ in series. Figure 6.1 presents a typical forming *I-t* trace that shows an abrupt current increase at ~10 ms after the beginning of the pulse, which corresponds to the formation of the filament (Chapter 4). Figure 6.2 shows an example

typical *I-t* traces of SET and RESET pulse operations (pulse width = 10 ns). For endurance cycling, ±2 V pulses (pulse width = 10 µs, duty cycle = 10 %) were applied with an on chip R_{LOAD} = 10 kΩ in series. Between each programming pulse, an *I-V* current sweep (up to 10 µA) was applied to read the device resistance. For continued programming and read, reed relay switches the device under test between the pulse generator/oscilloscope and the DC source meter alternately (Figure 6.3).

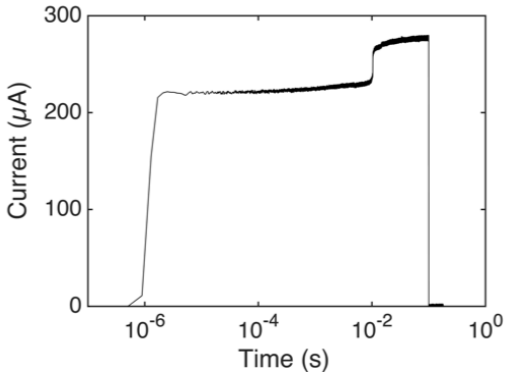

Figure 6.1 An *I-t* trace of the electroformation with a constant voltage source.

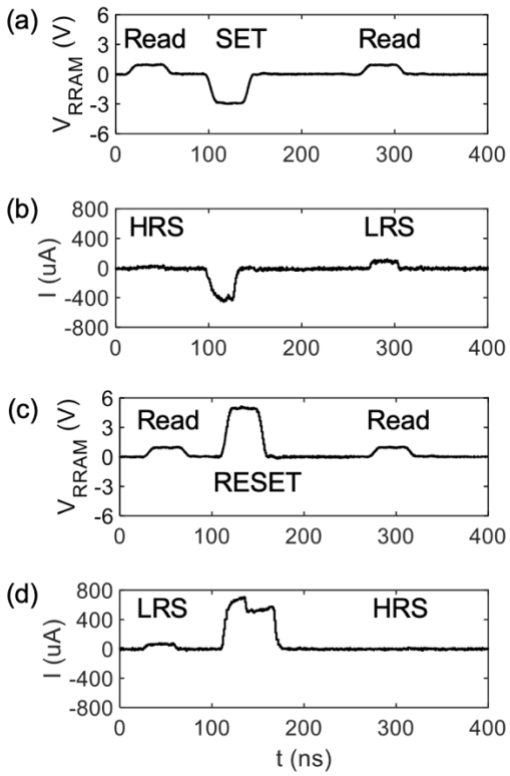

Figure 6.2 V-t and *I-t* traces of (a-b) SET and (c-d) RESET pulses with alternate read pulses showing the device resistance states before and afterwards.

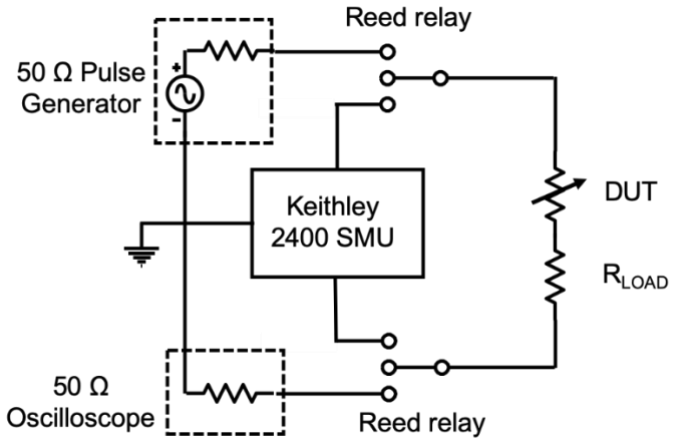

Figure 6.3 Circuit configuration used in device testing. The device under test (DUT) is switched between the pulse generator/oscilloscope and the DC source meter by Reed relay.

3) S/TEM Analysis

The FIB lift-out procedure is the same as described in Chapter 3 and 5. The STEM images were taken by FEI Themis S/TEM operating at 200 keV. The HAADF-STEM images were collected with Fischione model M3000 ADF detector at the inner and outer collection angle of 76.7 and 200 mrad, respectively. The EDS spectrum images were collected using the four detector "SuperX" energy dispersive X-ray spectrometer system with a probe current of 1 nA.

6.3 Evolution of the Conductive Filament due to Electrical Stressing

A series of devices were first put under pulse stressing with the same testing condition but varying the number of pulses in order to observe if and how the conductive filament evolves due to electrical stressing. Figure 6.4 (a) and (b) show the resistance plots of two devices under pulse stressing for 10^3 and 10^4 cycles, respectively. The resistance in the plots are obtained from the current sweeps after SET and RESET pulses after a sequence of exponentially increasing number of pulses (1, 10, 10^2, …). During Read operations, the current was kept under 10 μA which is below the typical current at which the device resistance switches to the other state (Figure 3.2 (c)). As shown in Figure 6.4 (a) and (b), with the testing condition for forming and endurance cycling fixed, the device resistance is relatively consistent and stable throughout the electrical

stressing. During the stressing, HRS mostly falls within 80 ± 10 kΩ and the LRS within 10 ± 2 kΩ, giving an ON/OFF ratio ≈ 8 (the ratio between the HRS and the LRS resistance).

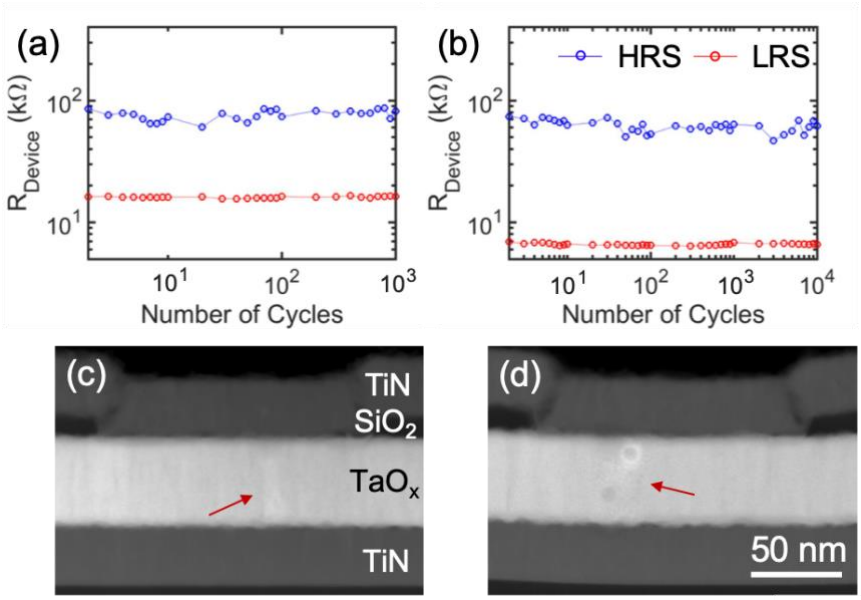

Figure 6.4 The resistance plots of two devices that were put under pulse cycling for (a) 10_3 cycles, and (b) 10_4 cycles, respectively. (c) and (d) are the HAADF images of the devices after 10_3 and 10_4 cycles, corresponding to the devices from (a) and (b), respectively. The red arrows in (c) and (d) mark where the filaments are located. The scale bar applies to both (c) and (d).

The devices after the electrical stressing were then lifted-out by FIB and characterized by STEM. The HAADF images of the two devices from Figure 6.4 (a) and (b) are shown in Figure 6.4 (c) and (d). The conductive filaments in the two devices are less visible than previously observed (Figure 3.4, 3.8, and 3.9), likely due to the relatively higher thickness of the TEM sample reducing the contrast of the filament buried within the functional layer. Nevertheless, in Figure 6.4 (c), a pillar shaped filament with a similar shape and dimension as those observed in Figure 3.4, 3.8, and 3.9 can be observed, indicating the filament after 10_3 switching cycles, the device behavior remains the same and the conductive filament also remains the same as just formed or switched for several cycles.

On the other hand, the conductive filament after 10_4 cycles (Figure 6.4 (d)) does not relate to anything observed thus far. Instead of a ~ 40 nm long and ~10 nm wide pillar, it has a mottled contrast that resembles the contrast of the filament observed in plan-view from Figure 2.4 and 2.8. More interestingly, two circular dark features can also be seen in the filament area which suggests void formation in this process. In the

upper void, one can see brighter contrast surrounding the void indicating Ta enrichment around this void. It is important to note that, even though the filament has changed into a completely different structure and possibly composition as well, the device behavior has remained the same, which suggests that the structure of the conductive filament does not necessarily dictate the electrical behavior of the device.

Void formation in room-temperature sputtered amorphous oxides have been observed after annealing [146], or a one-time catastrophic high current induced Joule heating in the device [70], [147], [148]. Since the current overshoot due to parasitic capacitive discharge was limited by the on-chip resistor, and that the void formation is within a small scale compared with the catastrophic explosion from the current overshoot induced Joule heating reported by others [70], [147], [148], it is more likely that the repeated heating from the electrical stressing that caused the voids to form: every time the device is switched, the switching pulse causes Joule heating, which could resemble a mild annealing effect when repeated long enough (10^4 times), as suggested by Chen *et al*. [149]. After repeated heating from the stressing, the TaO_x should gradually relax the structure and approach the thermodynamically favorable phases (metallic Ta and crystalline Ta_2O_5) [117] and become denser than the original film. Therefore, the formation of voids is mostly likely due to the density difference between the as-deposited TaO_x and that after stressing. The repeated heating due to the electrical stressing can also explain the mottled contrast in the filament from Figure 6.4 (d). After repeated heating, the originally amorphous TaO_2 can separate into metallic Ta and stoichiometric Ta_2O_5, similar to the phase separation discussed on the filament observed in plan-view from Chapter 2, Section 2.4 (Figure 2.8).

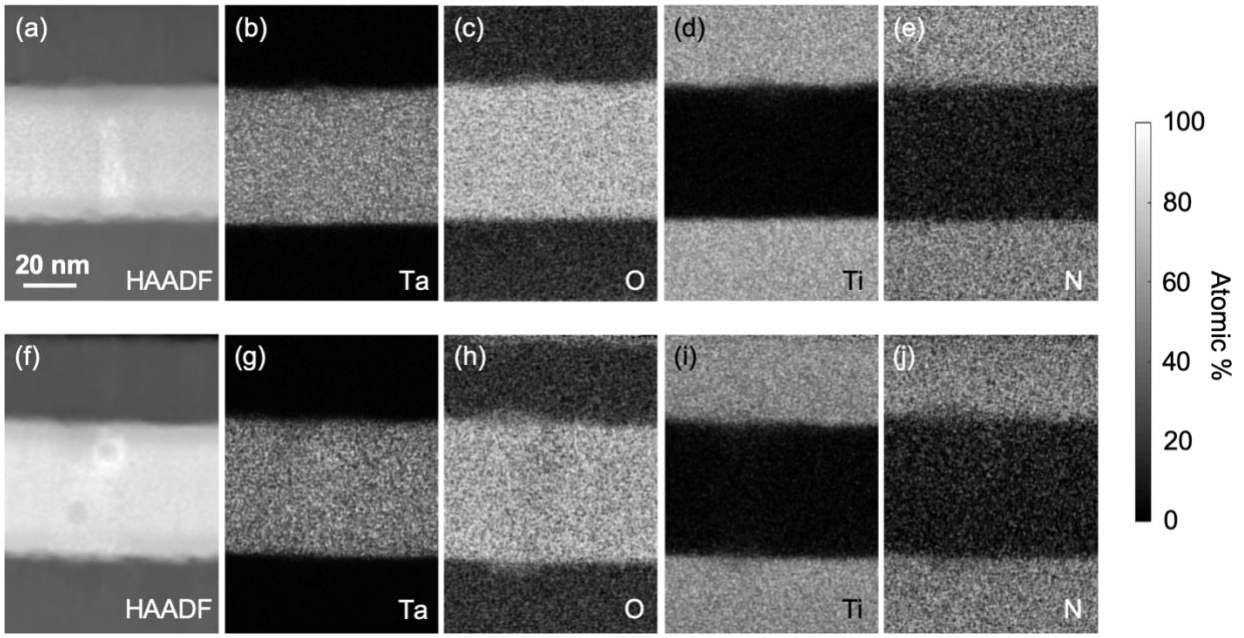

Figure 6.5 (a) HAADF reference image of the EDS maps collected from the same device as that shown in Figure 6.4 (c). (b-e) EDS elemental maps of Ta, O, Ti, and N over the area the same as (a). (f) HAADF reference image of the EDS maps collected from the same device as that shown in Figure 6.4 (d). (g-j) EDS elemental maps of Ta, O, Ti, and N over the area the same as (f). The scale bar applies to all images.

Figure 6.5 shows the EDS elemental maps collected from the devices shown in Figure 6.4. For the device that was stressed for 10_3 cycles (Figure 6.5 (a-e)), no clear ionic interdiffusion happened other than a small O dome in the TiN at the top interface and a corresponding N depletion. The less visible ionic interdiffusion in this device (compared with Figure 5.5) is likely due to the lower forming power (≈ 830 μW, Figure 6.1 and Figure 4.1 (b)), and thus the lower temperature and diffusion coefficient during forming.

The device after 10_4 cycles (Figure 6.5 (f-j)), however, does show a more significant interdiffusion at both interfaces between the TaO_x and the TiN compared with the one after 10_3 cycles. Since all devices in this chapter were formed with the same testing condition, it is reasonable that the higher extent of interdiffusion in the device after 10_4 cycles is caused by the longer pulsing on the device. From Chapter 5, we concluded that the ionic interdiffusion is mainly caused by the Fick diffusion activated at high temperature. This suggests that with continued Joule heating from the stressing, the elements indeed can keep diffusing down the concentration gradient where the temperature is high enough. Chen *et al*. argued that during electrical stressing, O could keep penetrating into the TiN electrodes and form resistive TiON resulting in device

failure [32]. Our observation contradicts their argument because although O has penetrated into the TiN within the 10_4 cycles, the device's electrical characteristics remained as good as new, which suggests the O enrichment in the TiN is not sufficient for device failure.

Comparing the two devices observed in Figure 6.4 and 6.5, it is safe to argue that the filament remains more or less stable both electrically and structurally in the earlier stage (e.g. first 10_3 cycles), but as the device continues to be under electrical stressing, the original filament structure with the Ta-rich O-poor core and the Ta_2O_5 gradually evolves into the mottled contrast caused by phase segregation due to prolonged heating from the stressing. Moreover, the prolonged heating in the device also drives continued ionic interdiffusion at the interfaces between the functional layer and the electrodes.

6.4 Understanding Endurance Failure due to Electrical Stressing

Knowing how the conductive filament evolves due to electrical stressing, we further move on to put devices under electrical stressing with the same condition until failure to understand the failure mechanism and how it is related to the filament evolution. Within the 41 devices cycled for endurance, 21 experienced SET failure (stuck at HRS), and 20 experienced RESET failure (stuck at LRS), rendering an almost even chance for device to fail with either of the failure modes. Figure 6.6 (a) and (b) show typical resistance plots of two devices that ended with SET failure and RESET failure, respectively. In the device that experienced SET failure in the end (Figure 6.6 (a)), the HRS ranged within 40-100 kΩ but remained relatively consistent with its original value, whereas the LRS gradually increased resistance over the entire life time of the device and eventually fell within the resistance range of HRS i.e. the SET failure after 10_5 cycles. On the other hand, for the device that ended with RESET failure (Figure 6.6 (b)), the LRS has remained stable (≈ 5 kΩ) while the HRS ranged from 10 kΩ to 80 kΩ corresponding to an ON/OFF ratio within a range of 2-20 before failure, and after 10_5 cycles the HRS dropped to the level of LRS and could not be switched back to the original HRS anymore. In both devices shown in Figure 6.6 (a) and (b), the resistance states have varied along the stressing and are less stable compared what was previously observed in Figure 6.4 (a) and (b).

This can be explained by the intrinsic cycle-to-cycle variability of the RRAM cells due to the random nature of the ion-motion induced sub-filament behavior which is responsible for the device resistance during resistive switching [150], [151]. Since the ion motion is somewhat stochastic even with the general direction imposed by the electric field, the sub-filament can be stronger or weaker from the same switching operation resulting in varied resistance states.

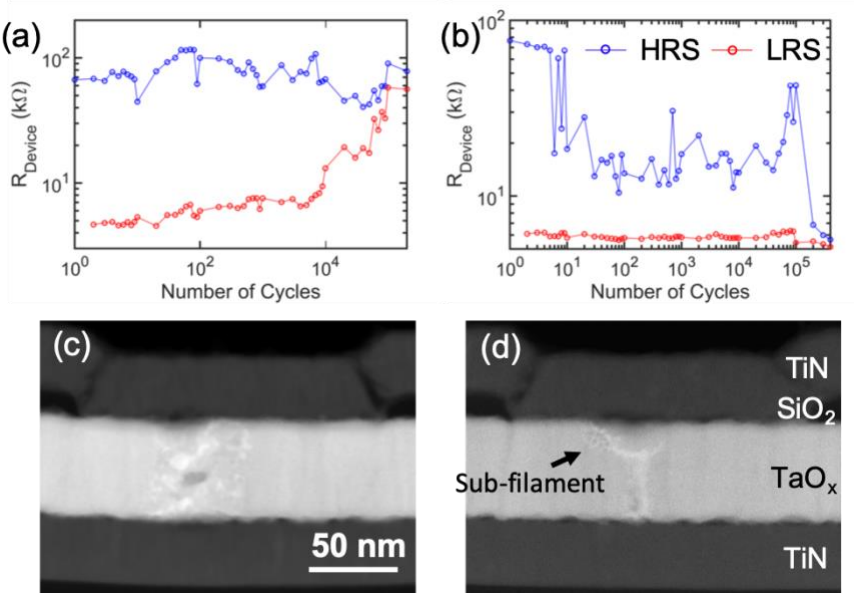

Figure 6.6 The resistance plots of two devices that were put under pulse cycling until (a) SET failure (stuck at HRS), and (b) RESET failure (stuck at LRS), respectively. (c) and (d) are the HAADF images of the devices after SET and RESET failure, corresponding to the devices from (a) and (b), respectively. The scale bar applies to both (c) and (d).

HAADF-STEM inspection on the two devices described above showed major changes in the structure of the conductive filament (Figure 6.6 (c) and (d)) compared with the still functioning ones such as freshly after forming (Figure 3.4 and 3.8), switched several times (Figure 3.9), and stressed for 10_3 cycles (Figure 6.4).

The device after SET failure (Figure 6.6 (c)) shows somewhat similar and consistent contrast as the one after 10_4 cycles (Figure 6.4 (d)). Instead of the original structure, the conductive filament became bright Ta-rich clusters distributing across the thickness of the functional layer with a diameter of 50 nm. The TaO_x surrounding the Ta-rich clusters is slightly darker than the unaffected TaO_x away from the filament, suggesting Ta depletion and possible composition close to Ta_2O_5 (denoted as Ta_2O_{5-x}). In addition, a dark

oval feature is also visible suggesting possible void formation. The observation in Figure 6.6 (c) indicates that with even longer stressing, the filament can keep segregating into larger metallic Ta clusters surrounded in the Ta_2O_{5-x} matrix. As the phase segregation goes on, the Ta rich clusters can grow larger to reduce the surface energy and the number of clusters becomes less. Eventually, the Ta clusters can be completely isolated by the resistive Ta_2O_{5-x} matrix and make the device's resistance permanently stuck at high resistance state, which is the cause of the SET failure. One can see that this failure mode is intrinsic to the thermodynamic properties of the material, and can likely be delayed but not avoided. To increase the endurance of the RRAM cells, one can also potentially select a different material system with a higher eutectic/eutectoid reaction temperature to delay the phase segregation.

The device with RESET failure from Figure 6.6 (d) shows a sub-filament with a high contrast and larger width (5-10 nm), which suggests after prolonged cycling, the sub-filament became more Ta concentrated and became almost as strong as the main body of the filament. At the top of the sub-filament next to the top electrode, there is mild mottled contrast indicating the sub-filament has started the phase segregation due to stressing after the total amount of 4×10^5 cycles. Even though this device has lasted longer than 10^4 cycles, it does not show the mottled contrast as much as the one after 10^4 cycles (Figure 6.4 (d)). One possible explanation is that due to the decreased resistance of HRS after 10 cycles, the load resistor takes over larger voltage drop from the switching pulses, and the power dissipation of the device during cycling is reduced due to the reduced voltage drop on the device, which lowers the temperature in the device during switching and delays the phase segregation. However, it is believed that it is a general trend with longer stressing, the conductive filament endures phase separation caused by repeated heating.

Recall from Chapter 3 that the resistive switching happens by the sub-filaments connecting and disconnecting between the filament core and the electrode adjacent to the gap driven by the high temperature and large electric field at the gap. Due to the intrinsic cycle-to-cycle variability of the filamentary RRAM, chances are that the sub-filament can grow stronger from the SET pulses. Once the sub-filament becomes too conductive, the gap will no longer undertake most of the voltage drop across the device, but rather, the voltage drop will distribute more evenly across the filament core and the sub-filament. The

weakened electric field lowers the driving force for the sub-filament to be disconnected from the electrode, which leads to the RESET failure.

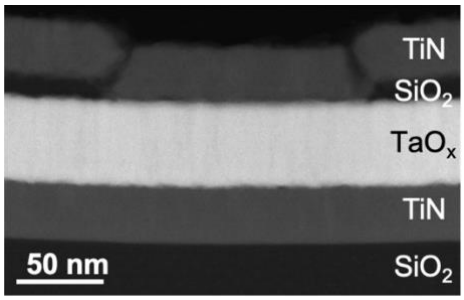

Figure 6.7 High-Angle Annular Dark-Field (HAADF) of an as-fabricated device.

A close inspection of the failed cells from Figure 6.6 (c) and (d) reveals additional dark areas at the interfaces between the functional layer and the electrodes compared with an as-fabricated cell (Figure 6.7). These features appear to be due to nanoscale delamination between TaOx and TiN after the cell was programmed, possibly due to the thermal stress. Note that the interfaces are not delaminated where the filament comes in contact with the electrodes, which suggests the delamination does not necessarily lead to device failure.

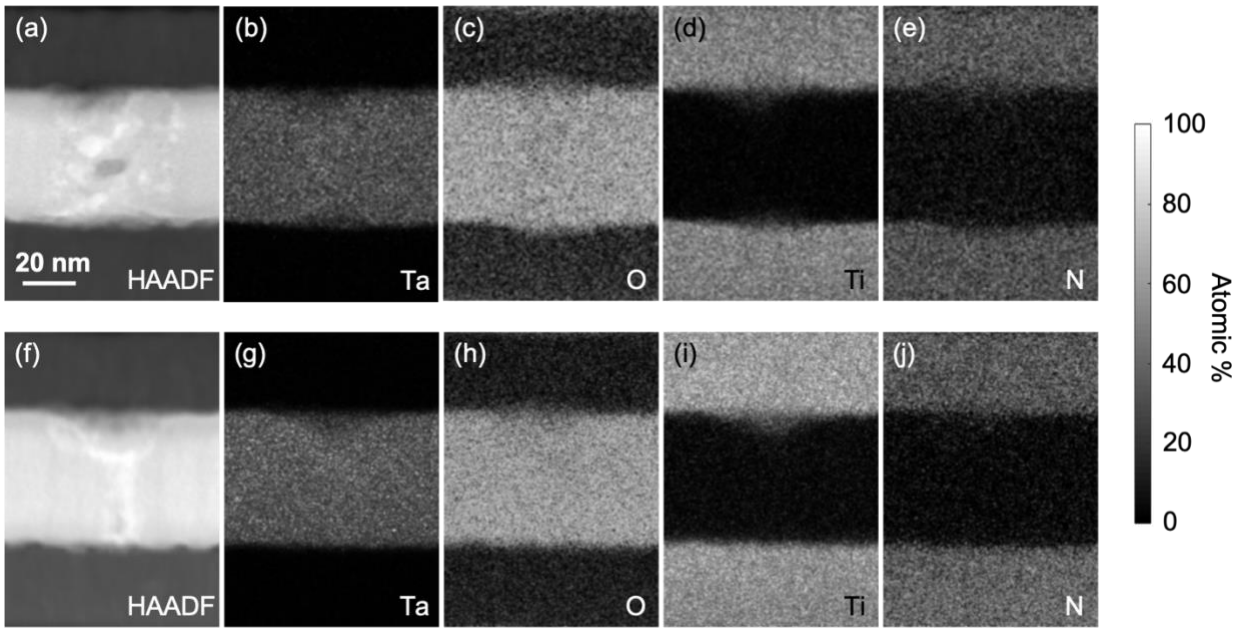

Figure 6.8 (a) HAADF reference image of the EDS maps collected from the device after SET failure, the same as that shown in Figure 6.6 (c). (b-e) EDS elemental maps of Ta, O, Ti, and N over the area the same as (a). (f) HAADF reference image of the EDS maps collected from the device after RESET failure, the same as that shown in Figure 6.6 (d). (g-j) EDS elemental maps of Ta, O, Ti, and N over the area the same as (f). The scale bar applies to all images.

EDS elemental maps of the device after SET failure (Figure 6.8 (b-e)) shows clear interdiffusion at both interfaces of O and Ti which aggress with our observation and interpretation from Figure 6.5 (g-j). The limited O interdiffusion in the RESET failed device (Figure 6.8 (h)) aggress with the argument earlier that the lower HRS in this device resulted in less voltage drop, which leads to lower power dissipation and therefore lower temperature in the cell rendering the limited O diffusion into the TiN. From the discussion in Chapter 5.6, there was a concern that with the continued diffusion of O from the functional layer into the electrode, it could eventually become too conductive and lead to RESET failure. Our observation above suggests that O diffusion out of the oxide does not necessarily result in RESET failure since more O interdiffusion was observed in the device after SET failure, whereas only a little amount of O interdiffusion was observed in the device after RESET failure.

6.6 The Influence of Testing Condition on Device Endurance

After the research on the origin of endurance failure in the TaO_x-based RRAM cells, and knowing that the Joule heating is one of the major reasons that lead to device failure, the devices were further tested endurance with varying programming pulse width keeping all the other testing conditions the same, in order to better understand how the endurance is affected by testing condition.

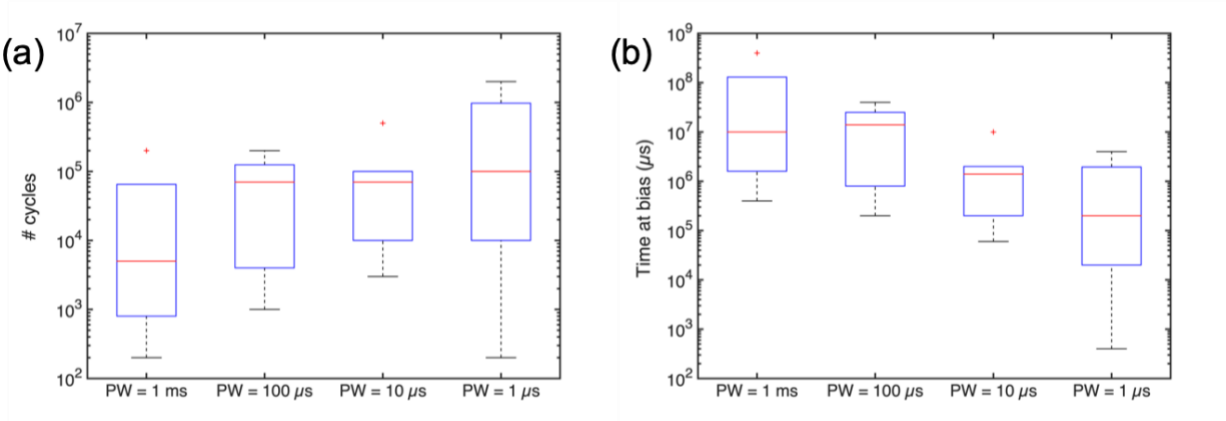

Figure 6.9 Distribution of device endurance in (a) number of cycles, and (b) total time at bias, by varying the pulse width (PW) during stressing.

Figure 6.9 (a) shows a distribution of device endurance (in number of cycles before failure) as a function of pulse width during stressing. One can see that with shorter programming pulses, the device endurance tends to increase. This agrees with our previous discussion since shorter pulses reduce the time at high temperature due to Joule heating for each switching cycle, therefore allowing the devices to last more switching cycles.

The same thing cannot be said for Figure 6.9 (b), however, which shows the total time at bias during the lifespan of the devices decreases with the decreasing pulse width, despite the increasing number of endurance cycles. The counterintuitive observation from Figure 6.9 (b) can be likely explained as due to a mild parasitic capacitive discharge from each switching operation. Although the on-chip load resistor is believed to limit the current overshoot due to parasitic capacitive discharge [20], and has been included in all testing on our memory cells, it only reduces the current overshoot but cannot have it completely eliminated. Therefore, even with the on-chip resistor, with the reduced capacitive discharge induced current

overshoot repeated for thousands of times, the lifespan of the devices can still be affected as seen in Figure 6.9 (b).

6.7 Conclusion and Discussion

As a device goes under electrical stressing, the original structure with the filament core and sub-filament gradually evolves into a widened filament with mottled contrast due to phase separation caused by repeated Joule heating from the switching operations. With continued stressing, the separated phases can keep growing larger and sparser to reduce the surface energy. Due to the intrinsic cycle-to-cycle variability in RRAM, the sub-filament can grow stronger before the phase separation starts, making it harder for RESET, and leads to RESET failure. With the prolonged stressing, the phase segregation can continue so much so that the conductive parts of the filament are isolated from each other leading to SET failure.

Both RESET and SET failures are intrinsic to RRAM where RESET failure is induced by its cycle-to-cycle variability, and SET failure is due to the material's thermodynamic properties, therefore cannot be avoided but can likely be delayed to increase RRAM endurance. Although shorter programming pulses do decrease the total time at bias of the memory cells during their lifespan, they do increase the number of switching cycles meaning the memory cells can be programmed and accessed more times with shorter pulses. Also, integrating an access device that can better control current, such as a transistor in series can significantly help control the current overshoot due to the capacitive discharge by directly limiting the current therefore help increase the endurance of the memory cells.

Chapter 7 Summary and Conclusions

This thesis has been focused on understanding the operation mechanism of TaO_x-based filamentary RRAM by using electron microscopy techniques to study the conductive filament in the memory cells.

The work started off in Chapter 2, where the conductive filament post-forming was investigated in plan-view geometry where the possibility of visualization of the conductive filament is maximized by inspecting the entire functional layer of the device. Analytical STEM techniques were used to identify the structure and the composition of the filament. Also, Soret effect in the lateral direction of the functional layer was identified as a driving force causing the formation of the conductive filament.

In Chapter 3, the S/TEM inspection of conductive filaments was continued in cross-sectional geometry with a new device structure so as to understand the vertical structure of the filament without artifacts introduced from device fabrication or electrical testing. It was found that in addition to Soret effect, the electric field across the device also serves as a driving force during forming. The device after forming contains a pillar shaped Ta-rich O-poor filament core and a Ta_2O_5 resistive gap. Moreover, the conductive filaments were also studied from devices after resistive switching, which revealed the mechanism of resistive switching to be the sub-filament around the resistive gap connecting and disconnecting between the filament core and the electrode under applied electric field resulting in the change in device resistance.

The forming dynamics was studied in Chapter 4. Formation by different voltages revealed an Arrhenius relationship between the speed of forming and the temperature inside the device, from which the activation energy of the forming process was extracted. Plan-view STEM investigation also showed that the conductive filament only appeared after the permanent resistance drop of the forming but not before. The forming power was also found to affect the device resistance after forming.

Chapter 5 re-evaluated the operation mechanism of TaO_x-based RRAM by X-ray compositional analysis in STEM. Temperature activated ionic interdiffusion of elements across interfaces after filament formation was revealed. The O was found to have diffused into the TiN electrode by Fick diffusion rather than driven by the electric field. The findings suggest the operation mechanism of TaO_x-based RRAM is a mixture of VCM (bipolar, switching driven by electric field) and TCM (Soret effect, temperature activated interdiffusion). Concerns were also raised on device endurance which can be potentially affected by the interdiffusion.

Finally, the evolution of conductive filament due to electrical stressing and the failure mechanism of TaO_x-based RRAM were explored in Chapter 6. It was found that the continued stressing on the memory cell can cause repeated Joule heating which can be considered as virtual annealing, and, if long enough, can cause the filament to separate into more thermodynamically favorable phases (Ta + Ta_2O_5). Furthermore, the SET failure was concluded to be the consequence of continued stressing induced annealing, and the RESET failure was argued to be due to the intrinsic cycle-to-cycle variation of metal-oxide RRAM. Several approaches were proposed to increase the device endurance according to the findings of this Chapter.

The conclusions of this thesis are believed to be able to explain the forming and resistive switching in RRAM cells with inert electrodes and thick amorphous metal sub-oxide as the functional layer based on Poole-Frenkel charge transport mechanism. For devices with an electrochemically active metal electrode, the CBM applies rather than the combination of VCM and TCM argued in this thesis. Moreover, if the functional layer is thin (< 20 nm), the electric field across the functional layer during the device operation can be significantly higher which can potentially change the balance between the Soret effect and electric field as observed in this work, and lead to behaviors with a different mechanism. The devices with functional layers whose charge transport mechanism is not based on Poole-Frenkel may also behave differently since the different electrical behavior can likely cause a different temperature distribution in the memory cell. For fully oxidized metal oxide as the functional layer, if the composition is stable at elevated temperatures due to Joule heating from the programming, the resistive switching can also happen with a different mechanism

since there will be no phase segregation, and it will likely be harder for the metal ions to diffuse. If the functional material is crystalline, the diffusion of particles can be anisotropic and add another factor into the ionic motion during the resistive switching therefore would also need to be discussed differently.

Moving forward, devices with the same structure but with a different metal sub-oxide that is amorphous and based on Poole-Frenkel conduction as the functional layer, such as NbO_x, should be investigated to test for the universality of the proposed model in this thesis. Additionally, devices with the structure and functional layer kept the same, but with a thinner functional oxide should also be studied in order to explore if the resistive switching behavior can be affected with a stronger electrical field in the functional layer. When studying the devices with thinner oxides, the interfaces between the oxide and the electrodes should be kept as flat as possible to avoid the ununiform electric field distribution in the functional layer due to interface unevenness.

References

[1] M. A. Zidan, J. P. Strachan, and W. D. Lu, “The future of electronics based on memristive systems,” *Nat. Electron.*, vol. 1, no. 1, pp. 22–29, 2018.

[2] S. Lee, “What’s Wrong With 3D NAND?” [Online]. Available: https://www.eetimes.com/author.asp?section_id=36&doc_id=1331957.

[3] S. W. Fong, C. M. Neumann, and H. S. P. Wong, “Phase-Change Memory - Towards a Storage-Class Memory,” *IEEE Trans. Electron Devices*, vol. 64, no. 11, pp. 4374–4385, 2017.

[4] J. Choe, “Intel 3D XPoint Memory Die Removed from Intel Optane™ PCM (Phase Change Memory).” [Online]. Available: https://www.techinsights.com/about-techinsights/overview/blog/intel-3D-xpoint-memory-die-removed-from-intel-optane-pcm/.

[5] D. Kau *et al.*, “A stackable cross point phase change memory,” *Int. Electron Devices Meet. IEDM*, pp. 27.1.1-27.1.4, 2009.

[6] Y. J. Song *et al.*, “Highly functional and reliable 8Mb STT-MRAM embedded in 28nm logic,” *Tech. Dig. - Int. Electron Devices Meet. IEDM*, pp. 27.2.1-27.2.4, 2017.

[7] G. W. Burr *et al.*, “Neuromorphic computing using non-volatile memory,” *Adv. Phys. X*, vol. 2, no. 1, pp. 89–124, 2017.

[8] S. Yu *et al.*, “Binary neural network with 16 Mb RRAM macro chip for classification and online training,” in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 16.2.1-16.2.4.

[9] S. Ambrogio *et al.*, “Equivalent-accuracy accelerated neural-network training using analogue memory,” *Nature*, vol. 558, no. 7708, pp. 60–67, 2018.

[10] A. Chen, “A review of emerging non-volatile memory (NVM) technologies and applications,” *Solid. State. Electron.*, vol. 125, pp. 25–38, Nov. 2016.

[11] H.-S. P. Wong *et al.*, “Metal–Oxide RRAM,” *Proc. IEEE*, vol. 100, no. 6, pp. 1951–1970, Jun. 2012.

[12] I. G. G. Baek *et al.*, “Multi-layer cross-point binary oxide resistive memory (OxRRAM) for post-NAND storage application,” *IEEE Int. Devices Meet. 2005. IEDM Tech. Dig.*, vol. 00, no. c, pp. 750–753, 2005.

[13] H. Y. Lee *et al.*, “Evidence and solution of over-RESET problem for HfOX based resistive memory with sub-ns switching speed and high endurance,” *Tech. Dig. - Int. Electron Devices Meet. IEDM*, pp. 7–10, 2010.

[14] M.-J. Lee *et al.*, “A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta2O(5-x)/TaO(2-x) bilayer structures.,” *Nat. Mater.*, vol. 10, no. 8, pp. 625–630, Aug. 2011.

[15] T. Y. Liu *et al.*, “A 130.7mm2 2-layer 32Gb ReRAM memory device in 24nm technology,” *Dig. Tech. Pap. - IEEE Int. Solid-State Circuits Conf.*, vol. 56, pp. 210–211, 2013.

[16] D. Ielmini, “Brain-inspired computing with resistive switching memory (RRAM): Devices, synapses and neural networks,” *Microelectron. Eng.*, vol. 190, pp. 44–53, 2018.

[17] Y. J. Jeong, S. Kim, and W. D. Lu, “Utilizing multiple state variables to improve the dynamic range of analog switching in a memristor,” *Appl. Phys. Lett.*, vol. 107, no. 17, 2015.

[18] J. M. Goodwill, A. A. Sharma, D. Li, J. A. Bain, and M. Skowronski, “Electro-Thermal Model of Threshold Switching in TaOx-Based Devices,” *ACS Appl. Mater. Interfaces*, vol. 9, no. 13, pp. 11704–11710, 2017.

[19] K. Kinoshita *et al.*, “Reduction in the reset current in a resistive random access memory consisting of Ni Ox brought about by reducing a parasitic capacitance,” *Appl. Phys. Lett.*, vol. 93, no. 3, p. 033506, Jul. 2008.

[20] Y. Fan, L. Zhang, D. Crotti, T. Witters, M. Jurczak, and B. Govoreanu, “Direct Evidence of the Overshoot Suppression in Ta 2 O 5 -Based Resistive Switching Memory With an Integrated Access Resistor,” *IEEE Electron Device Lett.*, vol. 36, no. 10, pp. 1027–1029, 2015.

[21] Y. M. Lu, M. Noman, W. Chen, P. A. Salvador, J. A. Bain, and M. Skowronski, “Elimination of high transient currents and electrode damage during electroformation of TiO 2-based resistive switching devices,” *J. Phys. D. Appl. Phys.*, vol. 45, no. 39, 2012.

[22] Y. Yang, P. Sheridan, and W. Lu, “Complementary resistive switching in tantalum oxide-based resistive memory devices,” *Appl. Phys. Lett.*, vol. 100, no. 20, p. 203112, 2012.

[23] F. Nardi, S. Balatti, S. Larentis, D. C. Gilmer, and D. Ielmini, “Complementary Switching in Oxide-Based Bipolar Resistive-Switching Random Memory,” *IEEE Trans. Electron Devices*, vol. 60, no. 1,

pp. 70–77, Jan. 2013.

[24] Y. Hirose and H. Hirose, “Polarity-dependent memory switching and behavior of Ag dendrite in Ag-photodoped amorphous As2S3films,” *J. Appl. Phys.*, vol. 47, no. 6, pp. 2767–2772, 1976.

[25] D. Jana *et al.*, “Conductive-bridging random access memory: challenges and opportunity for 3D architecture,” *Nanoscale Res. Lett.*, vol. 10, no. 1, 2015.

[26] R. Waser, R. Dittmann, C. Staikov, and K. Szot, “Redox-based resistive switching memories nanoionic mechanisms, prospects, and challenges,” *Adv. Mater.*, vol. 21, no. 25–26, pp. 2632–2663, 2009.

[27] I. Valov and M. N. Kozicki, “Cation-based resistance change memory,” *J. Phys. D. Appl. Phys.*, vol. 46, no. 8, 2013.

[28] S. P. Thermadam, S. K. Bhagat, T. L. Alford, Y. Sakaguchi, M. N. Kozicki, and M. Mitkova, “Influence of Cu diffusion conditions on the switching of Cu-SiO2-based resistive memory devices,” *Thin Solid Films*, vol. 518, no. 12, pp. 3293–3298, 2010.

[29] B. Govoreanu *et al.*, “10x10nm2 Hf/HfOx Crossbar Resistive RAM with Excellent Performance, Reliability and Low-Energy Operation,” in *International Electron Devices Meeting*, 2011, pp. 729–732.

[30] J. Kwon, A. a. Sharma, J. a. Bain, Y. N. Picard, and M. Skowronski, “Oxygen Vacancy Creation, Drift, and Aggregation in TiO2-Based Resistive Switches at Low Temperature and Voltage,” *Adv. Funct. Mater.*, vol. 25, pp. 2876–2883, 2015.

[31] D. S. Jeong, H. Schroeder, U. Breuer, and R. Waser, “Characteristic electroforming behavior in Pt/TiO2/Pt resistive switching cells depending on atmosphere,” *J. Appl. Phys.*, vol. 104, no. 12, p. 123716, 2008.

[32] C. Y. Chen *et al.*, “Understanding the Impact of Programming Pulses and Electrode Materials on the Endurance Properties of Scaled Ta 2 O 5 RRAM Cells,” in *IEDM 2014*, pp. 355–358.

[33] B. Traore *et al.*, “Impact of electrode nature on the filament formation and variability in HfO2RRAM,” *IEEE Int. Reliab. Phys. Symp. Proc.*, pp. 5E.2.1-5E.2.5, 2014.

[34] B. Traoré, P. Blaise, E. Vianello, L. Perniola, B. De Salvo, and Y. Nishi, “HfO2-based RRAM: Electrode effects, Ti/HfO2 interface, charge injection, and oxygen (O) Defects diffusion through

experiment and Ab initio calculations," *IEEE Trans. Electron Devices*, vol. 63, no. 1, pp. 360–368, 2016.

[35] K. Szot, W. Speier, G. Bihlmayer, and R. Waser, "Switching the electrical resistance of individual dislocations in single-crystalline SrTiO3.," *Nat. Mater.*, vol. 5, no. 4, pp. 312–320, 2006.

[36] J. Joshua Yang *et al.*, "The mechanism of electroforming of metal oxide memristive switches," *Nanotechnology*, vol. 20, no. 21, p. 215201, May 2009.

[37] A. Wedig *et al.*, "Nanoscale cation motion in TaOx, HfOx and TiOx memristive systems," *Nat. Nanotechnol.*, vol. 11, no. 1, pp. 67–74, 2015.

[38] D. B. Strukov, F. Alibart, and R. Stanley Williams, "Thermophoresis/diffusion as a plausible mechanism for unipolar resistive switching in metal-oxide-metal memristors," *Appl. Phys. A Mater. Sci. Process.*, vol. 107, no. 3, pp. 509–518, 2012.

[39] I. G. Baek *et al.*, "Highly Scalable Non-volatile Resistive Memory using Simple Binary Oxide Driven by Asymmetric Unipolar Voltage Pulses," *IEDM Tech. Dig.*, pp. 587–590, 2004.

[40] J. Yun *et al.*, "Random and localized resistive switching observation in Pt / NiO / Pt pss," vol. 282, no. 6, pp. 280–282, 2007.

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

[42] D. Ielmini, R. Bruchhaus, and R. Waser, "Thermochemical resistive switching: materials, mechanisms, and scaling projections," *Phase Transitions*, vol. 84, no. 7, pp. 570–602, Jul. 2011.

[43] G. A. Gibson, "Designing Negative Differential Resistance Devices Based on Self-Heating," *Adv. Funct. Mater.*, vol. 28, no. 22, p. 1704175, 2018.

[44] A. A. Sharma, M. Noman, M. Abdelmoula, M. Skowronski, and J. A. Bain, "Electronic Instabilities Leading to Electroformation of Binary Metal Oxide-based Resistive Switches," *Adv. Funct. Mater.*, vol. 24, pp. 5522–5529, Jul. 2014.

[45] J. M. Goodwill, A. A. Sharma, D. Li, J. A. Bain, and M. Skowronski, "Electro-Thermal Model of Threshold Switching in TaOx-Based Devices," *ACS Appl. Mater. Interfaces*, vol. 9, no. 13, pp. 11704–11710, 2017.

[46] J. M. Goodwill *et al.*, “Spontaneous current constriction in threshold switching devices,” *Nat. Commun.*, vol. 10, no. 1, p. 1628, 2019.

[47] H. Schmalzried and A. Navrotsky, *Festkörperthermodynamik*. Weinheim: Verlag Chemie, 2004.

[48] S. Kumar *et al.*, “Direct Observation of Localized Radial Oxygen Migration in Functioning Tantalum Oxide Memristors,” *Adv. Mater.*, vol. 28, no. 14, pp. 2772–2776, 2016.

[49] S. Kumar *et al.*, “Oxygen migration during resistance switching and failure of hafnium oxide memristors Oxygen migration during resistance switching and failure of hafnium oxide memristors,” vol. 103503, pp. 1–5, 2017.

[50] Y. Ma *et al.*, “Formation of the Conducting Filament in TaOx-Resistive Switching Devices by Thermal-Gradient-Induced Cation Accumulation,” *ACS Appl. Mater. Interfaces*, vol. 10, no. 27, pp. 23187–23197, 2018.

[51] Y. Ma *et al.*, “Stable Metallic Enrichment in Conductive Filaments in TaO x -Based Resistive Switches Arising from Competing Diffusive Fluxes,” *Adv. Electron. Mater.*, vol. 1800954, pp. 1–8, 2019.

[52] A. R. Allnatt and A. V. Chadwick, “Thermal Diffusion in Crystalline Solids,” *Chem. Rev.*, vol. 67, no. 6, pp. 681–705, 1967.

[53] H. Timm and J. Janek, “On the Soret effect in binary nonstoichiometric oxides-kinetic demixing of cuprite in a temperature gradient,” *Solid State Ionics*, vol. 176, no. 11–12, pp. 1131–1143, 2005.

[54] J. Janek, J. Sann, B. Mogwitz, M. Rohnke, and M. Kleine-Boymann, “Degradation of functional materials in temperature gradients - Thermodiffusion and the soret effect,” *J. Korean Ceram. Soc.*, vol. 49, no. 1, pp. 56–65, 2012.

[55] J. A. Brinkman, “The effect of temperature gradients on diffusion in crystals,” *Physical Review*, vol. 93, no. 2. p. 345, 1954.

[56] A. D. Le Claire, “Some predicted effects of temperature gradients on diffusion in crystals,” *Phys. Rev.*, vol. 93, no. 2, p. 344, 1954.

[57] W. Shockley, “Some Predicted Effects of Temperature Gradients on Diffusion in Crystals,” *Phys. Rev.*, vol. 93, no. 2, pp. 345–346, 1954.

[58] J. Janek and H. Timm, “Thermal diffusion and Soret effect in (U, Me) O2+δ : the heat of transport of

oxygen," *J. Nucl. Mater.*, vol. 255, no. 2–3, pp. 116–127, 1998.

[59] J. P. Strachan *et al.*, "Direct identification of the conducting channels in a functioning memristive device.," *Adv. Mater.*, vol. 22, no. 32, pp. 3573–3577, Aug. 2010.

[60] D.-H. Kwon *et al.*, "Atomic structure of conducting nanofilaments in TiO2 resistive switching memory.," *Nat. Nanotechnol.*, vol. 5, no. 2, pp. 148–153, 2010.

[61] F. Miao *et al.*, "Anatomy of a nanoscale conduction channel reveals the mechanism of a high-performance memristor.," *Adv. Mater.*, vol. 23, no. 47, pp. 5633–5640, 2011.

[62] J. Kwon *et al.*, "Transient Thermometry and High-Resolution Transmission Electron Microscopy Analysis of Filamentary Resistive Switches," *ACS Appl. Mater. Interfaces*, vol. 8, no. 31, pp. 20176–20184, 2016.

[63] H. Jiang *et al.*, "Sub-10 nm Ta Channel Responsible for Superior Performance of a HfO2 Memristor.," *Sci. Rep.*, vol. 6, p. 28525, 2016.

[64] Y. Yang, X. Zhang, L. Qin, Q. Zeng, X. Qiu, and R. Huang, "Probing nanoscale oxygen ion motion in memristive systems," *Nat. Commun.*, vol. 8, no. May, p. 15173, 2017.

[65] S. Kumar *et al.*, "Conduction Channel Formation and Dissolution Due to Oxygen Thermophoresis/Diffusion in Hafnium Oxide Memristors," *ACS Nano*, vol. 10, no. 12, pp. 11205–11210, Dec. 2016.

[66] R. Muenstermann, T. Menke, R. Dittmann, and R. Waser, "Coexistence of filamentary and homogeneous resistive switching in Fe-doped SrTiO3 thin-film memristive devices.," *Adv. Mater.*, vol. 22, no. 43, pp. 4819–4822, Nov. 2010.

[67] K. Skaja *et al.*, "Avalanche-Discharge-Induced Electrical Forming in Tantalum Oxide-Based Metal-Insulator-Metal Structures," *Adv. Funct. Mater.*, vol. 25, no. 46, pp. 7154–7162, Dec. 2015.

[68] H. Schroeder and D. S. Jeong, "Resistive switching in a Pt/TiO2/Pt thin film stack - a candidate for a non-volatile ReRAM," *Microelectron. Eng.*, vol. 84, no. 9–10, pp. 1982–1985, 2007.

[69] Deok-Kee Kim, Dong-Seok Suh, and Jucheol Park, "Pulse-Programming Instabilities of Unipolar-Type NiOx," *IEEE Electron Device Lett.*, vol. 31, no. 6, pp. 600–602, Jun. 2010.

[70] Y. Meng Lu, M. Noman, Y. N. Picard, J. A. Bain, P. A. Salvador, and M. Skowronski, "Impact of Joule heating on the microstructure of nanoscale TiO2 resistive switching devices," *J. Appl. Phys.*,

vol. 113, no. 16, p. 163703, 2013.

[71] Z. Wang, S. Kumar, H. S. P. Wong, and Y. Nishi, “Effect of thermal insulation on the electrical characteristics of NbOxthreshold switches,” *Appl. Phys. Lett.*, vol. 112, no. 7, pp. 1–6, 2018.

[72] S. Larentis, F. Nardi, S. Balatti, D. C. Gilmer, and D. Ielmini, “Resistive Switching by Voltage-Driven Ion Migration in Bipolar RRAM — Part II : Modeling,” *IEEE Trans. Electron Devices*, vol. 59, pp. 2468–2475, 2012.

[73] D. Ielmini, “Modeling the Universal Set / Reset Characteristics of Filament Growth,” *IEEE Trans. Electron Devices*, vol. 58, pp. 4309–4317, 2011.

[74] A. Marchewka *et al.*, “Nanoionic Resistive Switching Memories: On the Physical Nature of the Dynamic Reset Process,” *Adv. Electron. Mater.*, vol. 2, no. 1, p. 1500233, Jan. 2016.

[75] C. Li *et al.*, “Direct Observations of Nanofilament Evolution in Switching Processes in HfO2-Based Resistive Random Access Memory by In Situ TEM Studies,” *Adv. Mater.*, vol. 29, no. 10, p. 1602976, 2017.

[76] J. Chen, C. Hsin, C. Huang, C. Chiu, and Y. Huang, “Dynamic Evolution of Conducting Nanofilament in Resistive Switching Memories,” *Nano Lett.*, vol. 13, pp. 3671–3677, 2013.

[77] H. Lv *et al.*, “Evolution of conductive filament and its impact on reliability issues in oxide-electrolyte based resistive random access memory,” *Sci. Rep.*, vol. 5, pp. 1–6, 2015.

[78] A. A. Sharma, M. Noman, M. Abdelmoula, M. Skowronski, and J. A. Bain, “Electronic Instabilities Leading to Electroformation of Binary Metal Oxide-based Resistive Switches,” *Adv. Funct. Mater.*, vol. 24, pp. 5522–5529, 2014.

[79] S. Yu, H. Y. Chen, B. Gao, J. Kang, and H. S. P. Wong, “HfOx-based vertical resistive switching random access memory suitable for bit-cost-effective three-dimensional cross-point architecture,” *ACS Nano*, vol. 7, no. 3, pp. 2320–2325, 2013.

[80] Z. Jiang *et al.*, “Microsecond transient thermal behavior of HfOx-based resistive random access memory using a micro thermal stage (MTS),” in *Technical Digest - International Electron Devices Meeting, IEDM*, 2017, pp. 21.3.1-21.3.4.

[81] H.-S. P. Wong *et al.*, “Phase Change Memory,” *Proc. IEEE*, vol. 98, no. 12, pp. 2201–2227, 2010.

[82] D. Li *et al.*, “Joule Heating-Induced Metal-Insulator Transition in Epitaxial VO2/TiO2Devices,” *ACS*

Appl. Mater. Interfaces, vol. 8, no. 20, pp. 12908–12914, 2016.

[83] S. J. Pennycook, “Z-contrast stem for materials science,” *Ultramicroscopy*, vol. 30, no. 1–2, pp. 58–69, Jun. 1989.

[84] L. A. Giannuzzi, R. Geurts, and J. Ringnalda, “2 keV Ga+ FIB Milling for Reducing Amorphous Damage in Silicon,” *Microsc. Microanal.*, vol. 11, no. S02, pp. 2004–2006, 2005.

[85] J. Mayer, L. A. Giannuzzi, T. Kamino, and J. Michael, “T EM Sample Preparation and Damage,” no. May, pp. 400–407, 2007.

[86] C. A. Volkert and A. M. Minor, “Focused Ion Beam Microscopy and Micromachining,” *MRS Bull.*, vol. 32, no. 05, pp. 389–399, 2007.

[87] A. Chen, “Area and thickness scaling of forming voltage of resistive switching memories,” *IEEE Electron Device Lett.*, vol. 35, no. 1, pp. 57–59, 2014.

[88] Y. M. Lu, M. Noman, W. Chen, P. A. Salvador, J. A. Bain, and M. Skowronski, “Elimination of high transient currents and electrode damage during electroformation of TiO 2 -based resistive switching devices,” *J. Phys. D. Appl. Phys.*, vol. 45, no. 39, p. 395101, 2012.

[89] Z. Wei *et al.*, “Demonstration of High-density ReRAM Ensuring 10-year Retention at 85 o C Based on a Newly Developed Reliability Model,” pp. 721–724, 2011.

[90] G.-S. Park *et al.*, “In situ observation of filamentary conducting channels in an asymmetric Ta_2O5-x/TaO2-x bilayer structure.,” *Nat. Commun.*, vol. 4, p. 2382, Jan. 2013.

[91] R. F. Egerton, “Electron energy-loss spectroscopy in the TEM,” *Reports Prog. Phys.*, vol. 72, no. 1, p. 16502, Jan. 2009.

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

[93] J. P. Strachan *et al.*, “Direct identification of the conducting channels in a functioning memristive device.,” *Adv. Mater.*, vol. 22, no. 32, pp. 3573–3577, 2010.

[94] J. P. Strachan, D. B. Strukov, J. Borghetti, J. J. Yang, G. Medeiros-Ribeiro, and R. S. Williams, “The switching location of a bipolar memristor: chemical, thermal and structural mapping.,” *Nanotechnology*, vol. 22, no. 25, p. 254015, Jun. 2011.

[95] J. P. Strachan *et al.*, “Structural and chemical characterization of TiO2 memristive devices by spatially-resolved NEXAFS.,” *Nanotechnology*, vol. 20, no. 48, p. 485701, 2009.

[96] J. P. Strachan *et al.*, “Spectromicroscopy of tantalum oxide memristors,” *Appl. Phys. Lett.*, vol. 98, no. 24, p. 242114, 2011.

[97] S. Yu and H. S. P. Wong, “A phenomenological model for the reset mechanism of metal oxide RRAM,” *IEEE Electron Device Lett.*, vol. 31, no. 12, pp. 1455–1457, 2010.

[98] R. Degraeve *et al.*, “Dynamic ‘Hour Glass’ Model for SET and RESET in HfO2 RRAM,” *2012 Symp. VLSI Technol.*, no. 1, pp. 75–76, Jun. 2012.

[99] S. Kim *et al.*, “Physical electro-thermal model of resistive switching in bi-layered resistance-change memory,” *Sci. Rep.*, vol. 3, p. 1680, 2013.

[100] S. Kim, S. Choi, and W. Lu, “Comprehensive Physical Model of Dynamic Resistive Switching in an Oxide Memristor,” *ACS Nano*, vol. 8, pp. 2369–2376, 2014.

[101] K. Xue *et al.*, “A Combined Ab Initio and Experimental Study on the Nature of Conductive Filaments in Pt / HfO 2 / Pt Resistive Random Access Memory,” vol. 61, no. 5, pp. 1394–1402, 2014.

[102] P. R. Mickel, A. J. Lohn, C. D. James, and M. J. Marinella, “Isothermal switching and detailed filament evolution in memristive systems,” *Adv. Mater.*, vol. 26, no. 26, pp. 4486–4490, 2014.

[103] L. Zhu, J. Zhou, Z. Guo, and Z. Sun, “Synergistic Resistive Switching Mechanism of Oxygen Vacancies and Metal Interstitials in Ta 2 O 5,” *J. Phys. Chem. C*, vol. 120, no. 4, pp. 2456–2463, 2016.

[104] J. A. Davies, B. Domeij, J. P. S. Pringle, and F. Brown, “The Migration of Metal and Oxygen during Anodic Film Formation,” *J. Electrochem. Soc.*, vol. 112, no. 7, p. 675, 1965.

[105] J. L. Whitton, “The Measurement of Ionic Mobilities in the Anodic Oxides of Tantalum and Zirconium by a Precision Sectioning Technique,” *J. Electrochem. Soc.*, vol. 115, no. 1, p. 58, 1968.

[106] M. Lubben, P. Karakolis, V. Ioannou-Sougleridis, P. Normand, P. Dimitrakis, and I. Valov, “Graphene-Modified Interface Controls Transition from VCM to ECM Switching Modes in Ta/TaOx Based Memristive Devices,” *Adv. Mater.*, vol. 27, no. 40, pp. 6202–6207, 2015.

[107] Q. Xu, Y. Ma, and M. Skowronski, “Nanoscale density variations in sputtered amorphous TaOx functional layers in resistive switching devices,” *J. Appl. Phys. (Under Review)*, 2019.

[108] S. Kumar *et al.*, “Oxygen migration during resistance switching and failure of hafnium oxide memristors,” *Appl. Phys. Lett.*, vol. 110, no. 10, pp. 1–5, 2017.

[109] R. D. Shannon and C. T. Prewitt, “Effective ionic radii in oxides and fluorides,” *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.*, vol. 25, no. 5, pp. 925–946, 1969.

[110] I. P. Radu *et al.*, “Switching mechanism in two-terminal vanadium dioxide devices,” *Nanotechnology*, vol. 26, no. 16, 2015.

[111] S. Slesazeck *et al.*, “Physical model of threshold switching in NbO_2 based memristors,” *RSC Adv.*, vol. 5, no. 124, pp. 102318–102322, 2015.

[112] G. A. Gibson *et al.*, “An accurate locally active memristor model for S-type negative differential resistance in NbO_x,” *Appl. Phys. Lett.*, vol. 108, no. 2, p. 023505, Jan. 2016.

[113] B. K. Ridley, “Specific Negative Resistance in Solids,” *Proc. Phys. Soc.*, vol. 82, no. 6, p. 954, 1963.

[114] J. M. Goodwill, D. K. Gala, J. A. Bain, and M. Skowronski, “Switching dynamics of TaOx-based threshold switching devices,” *J. Appl. Phys.*, vol. 123, no. 11, p. 115105, 2018.

[115] L. Crespi *et al.*, “Modeling of Atomic Migration Phenomena in Phase Change Memory Devices,” *2015 IEEE 7th Int. Mem. Work. IMW 2015*, pp. 6–9, 2015.

[116] G. Novielli, A. Ghetti, E. Varesi, A. Mauri, and R. Sacco, “Atomic Migration in Phase Change Materials,” *2013 Int. Electron Devices Meet.*, pp. 589–592, 2013.

[117] S. P. Garg, N. Krishnamurthy, A. Awasthi, and M. Venkatraman, “The O-Ta (Oxygen-Tantalum) system,” *J. Phase Equilibria*, vol. 18, no. 4, pp. 407–407, 1997.

[118] S. Wu, H. M. Chan, and M. P. Harmer, “Compositional tailoring of the thermal expansion coefficient of tantalum (V) oxide,” *J. Mater. Sci.*, vol. 41, no. 3, pp. 689–695, 2006.

[119] U. N. Gries *et al.*, “A SIMS study of cation and anion diffusion in tantalum oxide,” *Phys. Chem. Chem. Phys.*, vol. 20, no. 2, pp. 989–996, 2018.

[120] D. S. Jeong, H. Schroeder, U. Breuer, and R. Waser, “Characteristic electroforming behavior in Pt/ TiO2/Pt resistive switching cells depending on atmosphere,” *J. Appl. Phys.*, vol. 104, no. 12, 2008.

[121] R. Münstermann, J. J. Yang, R. Strachan, John Paul Medeiros-Ribeiro, Gilberto Dittmann, and R. Waser, “Morphological and electrical changes in TiO2 memristive devices induced by electroforming and switching,” *Phys. status solidi – Rapid Res. Lett.*, vol. 4, pp. 16–18, 2010.

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

[123] Y. Yang *et al.*, “Electrochemical dynamics of nanoscale metallic inclusions in dielectrics.,” *Nat. Commun.*, vol. 5, p. 4232, 2014.

[124] H. Sun *et al.*, “Direct observation of conversion between threshold switching and memory switching induced by conductive filament morphology,” *Adv. Funct. Mater.*, pp. 5679–5686, 2014.

[125] M.-J. Lee *et al.*, “A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta2O5−x/TaO2−x bilayer structures,” *Nat. Mater.*, vol. 10, no. 8, pp. 625–630, 2011.

[126] S. Menzel, M. Waters, A. Marchewka, U. Böttger, R. Dittmann, and R. Waser, “Origin of the Ultra-nonlinear Switching Kinetics in Oxide-Based Resistive Switches,” *Adv. Funct. Mater.*, vol. 21, no. 23, pp. 4487–4492, Dec. 2011.

[127] E. Yalon, A. A. Sharma, M. Skowronski, J. A. Bain, D. Ritter, and I. V Karpov, “Thermometry of Filamentary RRAM Devices,” *IEEE Trans. Electron Devices*, vol. 62, pp. 2972–2977, 2015.

[128] S. Tirano *et al.*, “Microelectronic Engineering Accurate analysis of parasitic current overshoot during forming operation in RRAMs,” *Microelectron. Eng.*, vol. 88, no. 7, pp. 1129–1132, 2011.

[129] U. Celano *et al.*, “Three-dimensional observation of the conductive filament in nanoscaled resistive memory devices,” *Nano Lett.*, vol. 14, no. 5, pp. 2401–2406, 2014.

[130] U. Celano *et al.*, “Imaging the three-dimensional conductive channel in filamentary-based oxide resistive switching memory,” *Nano Lett.*, vol. 15, no. 12, pp. 7970–7975, 2015.

[131] J. M. LeBeau and S. Stemmer, “Experimental quantification of annular dark-field images in scanning transmission electron microscopy,” *Ultramicroscopy*, vol. 108, no. 12, pp. 1653–1658, 2008.

[132] S. Kumar *et al.*, “Physical origins of current and temperature controlled negative differential resistances in NbO2,” *Nat. Commun.*, vol. 8, no. 1, p. 658, Dec. 2017.

[133] L. Pauling, “Atomic Radii and Interatomic Distances in Metals,” *J. Am. Chem. Soc.*, vol. 69, no. 3, pp. 542–553, Mar. 1947.

[134] D. Li, J. M. Goodwill, J. A. Bain, and M. Skowronski, “Scaling behavior of oxide-based electrothermal threshold switching devices,” *Nanoscale*, vol. 9, no. 37, pp. 14139–14148, 2017.

[135] H. Jiang and Q. Xia, “Effect of voltage polarity and amplitude on electroforming of TiO2 based

memristive devices," *Nanoscale*, vol. 5, no. 8, p. 3257, Mar. 2013.

[136] S. Balatti, S. Larentis, D. C. Gilmer, and D. Ielmini, "Multiple Memory States in Resistive Switching Devices Through Controlled Size and Orientation of the Conductive Filament," *Adv. Mater.*, vol. 25, no. 10, pp. 1474–1478, Mar. 2013.

[137] S. Ambrogio, S. Balatti, D. C. Gilmer, and D. Ielmini, "Analytical modeling of oxide-based bipolar resistive memories and complementary resistive switches," *IEEE Trans. Electron Devices*, vol. 61, no. 7, pp. 2378–2386, Jul. 2014.

[138] P. Huang *et al.*, "A physics-based compact model of metal-oxide-based RRAM DC and AC operations," *IEEE Trans. Electron Devices*, vol. 60, no. 12, pp. 4090–4097, Dec. 2013.

[139] M. Noman, A. a. Sharma, Y. Meng Lu, M. Skowronski, P. a. Salvador, and J. a. Bain, "Transient characterization of the electroforming process in TiO2 based resistive switching devices," *Appl. Phys. Lett.*, vol. 102, no. 2, p. 23507, 2013.

[140] Y. Ma *et al.*, "Formation of the Conducting Filament in TaOx-Resistive Switching Devices by Thermal-Gradient-Induced Cation Accumulation," *ACS Appl. Mater. Interfaces*, vol. 10, no. 27, pp. 23187–23197, 2018.

[141] D. R. Lide, *CRC Handbook of Chemistry and Physics*, 2005th ed. Boca Raton, FL: CRC Press, Boca Raton, FL, 2005.

[142] D. R. Lide, *CRC Handbook of Chemistry and Physics*, Internet V. CRC Press, Boca Raton, FL, 2005, 2005.

[143] Z. Wang, S. Kumar, H.-S. P. Wong, and Y. Nishi, "Effect of thermal insulation on the electrical characteristics of NbOx threshold switches," *Appl. Phys. Lett.*, vol. 112, no. 7, p. 73102, Feb. 2018.

[144] Y. R. Luo, *Comprehensive handbook of chemical bond energies*. Boca Raton, FL: CRC Press, 2007.

[145] J. Desmaison, P. Lefort, and M. Billy, "Oxidation of titanium nitride in oxygen: Behavior of TiN0.83 and TiN0.79 plates," *Oxid. Met.*, vol. 13, no. 3, pp. 203–222, 1979.

[146] A. Obstarczyk, D. Kaczmarek, M. Mazur, D. Wojcieszak, and J. Domaradzki, "The effect of post-process annealing on optical and electrical properties of mixed - HfO 2 – TiO 2 thin film coatings," *J. Mater. Sci. Mater. Electron.*, vol. 30, no. 7, pp. 6358–6369, 2019.

[147] G. Martín, M. B. González, F. Campabadal, F. Peiró, A. Cornet, and S. Estradé, "Transmission

electron microscopy assessment of conductive-filament formation in Ni-HfO2-Si resistive-switching operational devices," *Appl. Phys. Express*, vol. 11, no. 1, 2018.

[148] D. Carta *et al.*, "X-ray spectromicroscopy investigation of soft and hard breakdown in RRAM devices," *Nanotechnology*, vol. 27, p. 345705, 2016.

[149] Y. Y. Chen *et al.*, "Understanding of the Endurance Failure in Scaled HfO 2 -based 1T1R RRAM through Vacancy Mobility Degradation," in *IEDM 2012*, 2012, pp. 10–13.

[150] C. Baeumer *et al.*, "Subfilamentary Networks Cause Cycle-to-Cycle Variability in Memristive Devices," *ACS Nano*, vol. 11, no. 7, pp. 6921–6929, 2017.

[151] M. B. Gonzalez, J. M. Rafí, O. Beldarrain, M. Zabala, and F. Campabadal, "Analysis of the Switching Variability in Ni/ HfO2-Based RRAM Devices," *IEEE Trans. Device Mater. Reliab.*, vol. 14, no. 2, pp. 769–771, 2014.