Redox-based memristive devices : towards highly scalable synaptic electronics

Putu Andhita Dananjaya

2020

Putu Andhita Dananjaya. (2020). Redox-based memristive devices : towards highly scalable synaptic electronics. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/146143>

<https://doi.org/10.32657/10356/146143>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 20:41:32 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**Redox-based Memristive Devices:
Towards Highly Scalable Synaptic Electronics**

Putu Andhita Dananjaya

SCHOOL OF PHYSICAL AND MATHEMATICAL SCIENCES

2020

**Redox-based Memristive Devices:
Towards Highly Scalable Synaptic Electronics**

Putu Andhita Dananjaya

SCHOOL OF PHYSICAL AND MATHEMATICAL
SCIENCES

A thesis submitted to the Nanyang Technological
University in partial fulfilment of the requirement for the
degree of Doctor of Philosophy

2020

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

15 January 2020

Date

Putu Andhita Dananjaya

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

15 January 2020
.....

Date

.....
Prof Lew Wen Siang

Authorship Attribution Statement

This thesis contains material from 1 publication in peer-reviewed journal and 2 on-going research. I am the first author in all of the abovementioned articles. Other articles those are not included in the thesis in which I am one of the co-authors will also be listed below.

Chapter 5 is published as P. A. Dananjaya, D. J. J. Loy, S. C. W. Chow, and W. S. Lew, "Unidirectional Threshold Switching Induced by Cu Migration with High Selectivity and Ultralow OFF Current under Gradual Electroforming Treatment," *ACS Applied Electronic Materials*, vol. 1, no. 10, pp. 2076-2085, 2019/10/22 2019. The contributions of the co-authors are as follows:

- Prof Lew Wen Siang provided preliminary directions to the project
- I co-designed and fabricated the devices investigated in the project with the help of Samuel Chow
- Desmond, Samuel, and I performed the electrical and material characterization on different samples prepared
- I conducted the data analysis and theoretical framework development under the supervision of Prof Lew Wen Siang
- I prepared the manuscript drafts, which were further edited by Desmond and Samuel and revised by Prof Lew Wen Siang.

Chapter 3 is an on-going research work that will be presented as P. A. Dananjaya, W.S. Lew, "Trap-controlled Space-Charge-Limited Switching Dynamics in Pt/HfO_x/Ti Memristive Devices". The contributions of the co-authors are as follows:

- I fabricated the devices and performed the device characterizations under the supervision of Prof Lew Wen Siang.
- I conducted data analysis and prepared the manuscript drafts, which were revised by Prof Lew Wen Siang

Chapter 4 is an on-going research work that will be presented as P. A. Dananjaya, W.S. Lew, "Compliance-Free Anion-based Pt/HfO_x/Ti Memristive Devices for Analog Synaptic Device Applications". The contributions of the co-authors are as follows:

- I fabricated the devices and performed the device characterizations under the supervision of Prof Lew Wen Siang.
- I conducted data analysis and prepared the manuscript drafts, which were revised by Prof Lew Wen Siang

Other articles those are not included in this thesis are the following:

1. X. L. Hong, D. J. J. Loy, P. A. Dananjaya, F.N. Tan, C.M. Ng, and W.S. Lew*, "Oxide-based RRAM materials for neuromorphic computing", Journal of Materials Science, 53, 8720 (20 18).
2. D. J. J. Loy, P. A. Dananjaya, X. L. Hong, D. P. Shum, W.S. Lew*, "Conduction Mechanisms on High Retention Annealed MgO-based Resistive Switching Memory Devices", Scientific Reports, 8, 14774 (2018).

16 January 2020

Date

Putu Andhita Dananjaya

Abstract

Complimentary Metal-Oxide Semiconductor (CMOS)-based systems have been the core elements of the semiconductor technology for decades. With the predicted CMOS scaling limit and the increasing amount of data in today's technology, researchers around the world have started looking for emerging electronics to keep up with the hardware requirements and new radical computing paradigm, e.g., quantum and neuromorphic computing, to further lower the computational cost, especially in handling unstructured data set where the conventional von Neumann architecture struggles to strike a balance between power cost and space trade-off.

Redox-based memristive devices emerge as one of the promising candidates to fulfil the hardware requirements of the emerging neuromorphic computing systems, e.g., as a synaptic device element. The highly scalable nature of the device along with its analog characteristic have been the focus of the research in the field. However, the inherent stochasticity, non-linearity, and symmetry of the device conductance switching behaviour hinder its progress in synaptic device applications. Fortunately, the synaptic device requirements are highly dependent on the target applications. Thus, systematic and thorough understanding upon the device physics involve during the switching operation is required to have full control on the performance at the system level and how to further improve it.

This thesis focuses on the development of redox-based memristive devices governed by different underlying physical mechanisms, i.e., anion and cation-based system, to facilitate different device applications. The anion-based devices were operated under different mode of programming to investigate its potential application in different synaptic array architectures. The switching dynamics, under trap-controlled space-charge-limited mechanism, and its correlation with the linearity and symmetry of the device conductance response are extensively discussed. On the other hand, the cation-based devices were operated under volatile switching regime to investigate its unique switching dynamics for highly scalable select devices. The device temporal response to external voltage applied was used to understand the device switching behaviour under the theoretical framework of field-induced nucleation theory and Rayleigh instability.

Acknowledgements

First, I would like to express my gratitude to my supervisor Prof Lew Wen Siang for all his advises and guidance throughout my PhD programme. I would also like to thank my friends in the spintronic device group for their help and endless support for me in carrying out my research work.

I thank my brother Wiswa and friend Tanjung for their close company here in Singapore. I also thank Mary, Steven, and Nathaniel for their hospitality while having me staying at their home.

And my special thanks to all my beloved support system in Bali, my mother Ida Ayu Nyoman Tirta, my father I Wayan Jajus Parwata, and the love of my life Anak Agung Putri Satwika, who have always been there for me, understanding and supporting me throughout this journey.

Table of Contents

Abstract	1
Acknowledgements.....	2
Table of Contents	3
CHAPTER 1 Introduction	6
1.1. Beyond CMOS Technology and Von Neumann Architecture	7
1.2. Hardware Requirements for Emerging Neuromorphic Systems	10
1.3. Redox-Based Memristive Devices in Neuromorphic Computing Platforms	12
1.4. Research Objectives	27
1.5. Thesis Organization.....	29
1.6. References	30
CHAPTER 2 Experimental Techniques	38
2.1. Device Fabrication	39
2.2. Material Characterizations	43
2.3. Electrical Characterizations	47
CHAPTER 3 Switching Dynamics of Pt/HfO_x/Ti Anion-Based Memristive Devices	50
3.1. Different Electroforming Treatments and The Corresponding IV Characteristics	51
3.2. Multilevel Conductance States	54
3.3. The Conduction and Switching Mechanisms Under LCF and SCF Treatments	62
3.4. Summary	70

3.5. References	71
CHAPTER 4 Synaptic Behaviour of Pt/HfO_x/Ti Anion-based Memristive Devices	75
1.1. Different Learning Approaches and Corresponding Synaptic Device Requirements	76
1.2. Trade-Off Between Dynamic Ratio and Progressive Conductance Switching Behaviour of The Devices	78
1.3. Identical Pulse Programming Operation of LCF And SCF Devices	82
1.4. Correlation Between Dynamic Ratio and Asymmetric Non-Linearity Factor	87
1.5. Reliability Aspect of The Devices	91
1.6. Summary	93
1.7. References	94
CHAPTER 5 Cation-Based Diffusive Memristor	97
1.1. Diffusive Memristor High Density Crossbar Array and Other Emerging Systems	98
1.2. Gradual Forming Process and Threshold Switching Characteristics of DM	99
1.3. Temporal Response of DM Devices	107
1.4. Reliability Aspects of DM: Device Endurance and Observation of Random Telegraph Signal (RTS)	115
1.5. Summary	117
1.6. References	119
CHAPTER 6 Conclusion and Future Work	126
1.1. Conclusion	127
1.2. Future Work	129
1.3. References	134

CHAPTER 1

Introduction

This chapter presents the research motivation, followed by a thorough review on the state-of-the art of redox-based synaptic devices, research objectives, and thesis organization. It starts with the significance of the research beyond CMOS technology and new computational paradigm. It also specifically addresses the hardware requirements and challenges in realizing neuromorphic computing systems with the emerging redox-based memristive devices. Then, it ends with the research objectives to be achieved and the thesis organization.

1.1. BEYOND CMOS TECHNOLOGY AND VON NEUMANN ARCHITECTURE

The building blocks of today's technology are heavily reliant on Complimentary Metal-Oxide Semiconductor (CMOS) based systems. CMOS has served the semiconductor industry well for decades. It has been the very core element in the manufacturing of the Integrated Circuits (ICs). It is frequently correlated with Very Large-Scale Integration (VLSI) process in which a considerably large number of MOS transistors are highly connected within a single compact chip or die. In fact, VLSI was only made possible with the discovery of the MOS transistor technology. In order to meet the scaling requirements of the main drivers of the current technology such as big data, cloud computing, blockchain, autonomous vehicle, augmented reality, and artificial intelligence (AI), various CMOS scaling techniques have been implemented to further push its physical limit.

Moore's Law has been used as a gauge of where the technological era stands over the years. It correlates the number of transistors packed within a single microchip with the course of time. It suggests that the transistors per chip would double every 2 years [1, 2]. In early 2000s, the transistor gate thickness of ~ 1.2 nm SiO_2 had been achieved [3]. This was the fundamental limit of the conventional scaling technique because further scaling would result in predominant electron tunnelling effect contribution on the total transistor leakage current. Thus, this marked the start of the transistor scaling innovations era. The first breakthrough was achieved by Intel in 2003, when strained silicon (Si) was introduced for NMOS and PMOS transistors in 90 nm technology node [4]. The strain technology on Si enables the atoms to stretch apart by $\sim 1\%$. In NMOS transistors, strain was promoted by contact etch-stop layer (CESL) process in which high-stress layer surrounding the transistor was introduced, while PMOS strain introduction utilised embedded Silicon Germanium (Si-Ge) process, where strained epitaxial Si-Ge was used to replace the conventional source-drain region. The strained Si technology successfully improved the transistor performance in terms of drive current and channel (electrons/holes) mobility. While it continues to be implemented in 10-nm technologies, this scaling technique was superseded by the introduction of high κ -dielectric hafnium oxide (HfO_2) taking over SiO_2 place as the gate oxide and the use of metal electrodes to replace doped-polysilicon in 2007 Intel's 45 nm technology [5]. This approach was able to improve the transistor drive current while maintaining considerably low leakage current. It was then scaled further down to 32 nm technology node with the use of immersion lithography process [6]. The subsequent significant achievement was the establishment of tri-gate transistors (Si-FinFET) at 22 nm node with significant channel electrostatics improvement [7].

It is able to facilitate the gate length scaling down to 15 nm for 7 nm node (Si/Ge-FinFET). There are several pathways currently being pursued in order to facilitate scaling of 5 nm and beyond, i.e., III-V FinFET [8], nanosheet gate all around (GAA) transistor [9], vertical tunnel field-effect transistor (TFET) [10, 11], negative capacitance FET (NCFET) [12], and carbon nanotube FET (CNFET) [13]. Moore's law is expected to end around 2025, but the real outcome is remained to be seen in near future. And even if Moore's law is able to continue what has been predicted, the scaling of the transistors comes with diminishing return in performance.

Based on IBM study in 2017, an astonishing 90% of the data recorded at the time was created within 2 years (from 2015). This enormous growth rate was attributed to the advancement of the internet. Several estimations and projections on this enormous growth rate have been made. The study of "The Digital Universe in 2020" conducted by IDC predicted the amount of data from digital world would double every 2 years, resulting in the projection of 38.5 zettabytes based on 1.2 zettabytes of data recorded back in 2010. Thus, in parallel with the view of CMOS physical scaling limit, other efforts in the field of memory technology and computing architecture to reduce the systems' reliance on CMOS technology have been made to cope with the enormous growth of data.

In the search of a single universal memory device, several technologies have emerged as potential candidates. An ideal memory device requires excellent scalability (< 10 nm) while maintaining high performances. Due to the trade-off among these performance parameters, achieving all requirements within a single memory cell becomes a highly challenging task. However, the progress made in the emerging memory technologies have shown the capability of different technologies to achieve parts of the characteristics required in an ideal memory and to fill the latency gap in memory hierarchy. The front runners of emerging memory candidates are phase change memory (PCM), spin transfer torque magnetoresistive random-access memory (STT-MRAM), and resistive switching memory (also known as redox-based memristive device or memristor) due to their highly scalable nature, i.e., two-terminal device footprint, and multibit per cell capability. These devices store their information based on the change in device conductance under external electric field. PCM device works based on the phase transition between amorphous (low conductance) and crystalline (high conductance) of chalcogenide materials induced by Joule heating generated by external electric field. When the amorphous chalcogenide is heated at the temperature between crystallization and melting point, the device will switch to its crystalline state resulting in high conductance. While the low

conductance state is realized by Joule heating above the melting point, leaving the chalcogenide layer in disordered (amorphous) state once it is cooled. The basic STT-MRAM structure consists of two magnetic layers, i.e., fixed (pinned or reference) layer and free layer, separated by an oxide tunnel barrier. It utilises the transfer of spin-angular momentum from electrons polarized by the fixed layer to alter the free layer magnetization. This enables parallel and anti-parallel magnetization between the free and fixed layer that results in device high and low conductivity. Redox-based memristive device consists of a switching host (oxide insulator or chalcogenide) sandwiched between two electrodes employing cations and/or anions migration within the structure under external electric field. Cations migration induces different conductivity by the presence of metallic species in the switching layer, while anions migration causes intrinsic defects modulation in the form of oxygen vacancy generation and recombination under external electric field. Different underlying switching mechanisms of these devices lead to different macroscopic performances, such as analog conductance switching behaviour, programming voltage and energy, device latency, endurance, and retention. Thus, some devices might be more suitable for certain applications than the other.

Along with CMOS-based systems, the conventional von Neumann computing paradigm has been the foundation that powers all the conventional computers and smart devices available today. They have been an excellent aggregate in handling structured well defined data sets. However, when it comes to performing complex tasks that involved unstructured data sets with imprecise input and output specifications and real-time data processing, such as sound classification and image recognition, the conventional approach face power cost and space trade-off issue. Thus, new computing paradigms with emerging technologies have been widely investigated to mitigate those bottlenecks. Inspired by massively parallel operations with low power consumption capability of human brain, neuromorphic computing has been considered as one of the most promising computing paradigms. In order to realize the high connectivity among around 100 billion of neurons, i.e., each neuron can be connected to up to 100,000 other neurons [14-16], highly scalable and high performance synaptic hardware is required.

Due to the desired low power consumption, highly scalable footprint, and analog programmable behaviour, memristive devices have been widely investigated as one of the hardware elements for artificial neural network (ANN), i.e., not only as a memory element, but also as a computing unit. From algorithm viewpoint, there are two ways of looking at memristive-based neuromorphic systems. Deep learning provides a plausible inference only design, i.e., direct mapping of pre-trained deep learning models within hardware constraints

onto memristive based neuromorphic hardware without any further training. On the other hand, memristive devices can also enable on-chip training capability in the system, where additional interface circuitry required for the algorithm implemented. Inference alone requires the conversion of existing pre-trained deep learning algorithms in high precision digital domain to the binary event-based (or spiking) domain so as, to be able to be mapped onto memristive based neuromorphic hardware. Whereas, on-chip training may be implemented at the memristive synaptic array in the neuromorphic hardware by emulating local spike timing-based algorithms such as spike timing dependent plasticity (STDP) or its variants. These two methods belong to a new computational paradigm known as spiking deep neural network (SDNN).

1.2. HARDWARE REQUIREMENTS FOR EMERGING NEUROMORPHIC SYSTEMS

An ideal synaptic device is one of the major elements required to realize a robust neuromorphic computing platform with high learning accuracy. It plays a major role in determining the interconnectivity strength among neurons in the system by storing the weight values, i.e., usually in the form of the device conductance. These values are updated according to the learning rules implemented during the training process. In the case of spiking neural network (SNN), the tunable conductance state of memristive-based synapses is analogous to the synaptic plasticity of the brain. The electrical connection between a presynaptic neuron and a postsynaptic neuron changes, strengthening or weakening the synaptic impulses, thus mimicking brain-like functionalities. With the excellent device and array level scalability of the memristive system, highly connected crossbar architecture can potentially be implemented in the large neural network.

From neural network (NN) accuracy and robustness viewpoint, the most important requirements for the synaptic device are deterministic switching with symmetrical and linear weight update, as depicted on **Figure 1.1**. Ideally, each synaptic device should exhibit non-overlapping multilevel conductance characteristics of at least 32 levels (5-bit). However, due to inherent cycle-to-cycle and device-to-device variation of memristive devices, trade-off in the number of bits per cell might be required to accommodate the state variation allowing sufficient read margin in between the states. Tighter distribution of the states can be achieved by implementation of write-verify scheme at the expense of programming energy and overall speed. The more conductance levels can be obtained within a single synapse will enhance the network immunity towards input noise, thus realizing higher learning and test accuracy. Despite that, the number of bits required per cell is still subjected to the network architecture

and algorithms implemented. Linearity of the weight update is associated with the relationship between the change of weight value for every programming cycle, while the symmetry is referring to the change of weight value during potentiation and depression mode. Symmetric linear weight update feature will allow convenient direct mapping of the device conductance and the algorithm weight values. Furthermore, it will enable more efficient training process through state-independent weight update. However, due to the two-terminal nature of memristive devices, asymmetric nonlinear change of conductance is a huge challenge. This undesired feature has been shown to significantly reduce the network learning accuracy. Thus, different techniques from materials and circuits perspective as well as hardware-algorithm co-optimizations have been investigated.

Other requirements from key device performance parameters consist of endurance characteristics of $\geq 10^9$, long data retention of ≥ 10 years, low programming energy of ≤ 10 fJ, high scalability of ≤ 10 nm, and maximum dynamic ratio of ≥ 100 . High endurance capability is required to allow more training cycles for the network. This is especially important for on-chip learning implementation. On the other hand, long retention is important to accommodate more inference processes, in which the weight values are read with minimum read disturbance. If the data retention of the device is poor, the number maximum number of inferences can be performed without refreshing the weight value will be relatively low. Trade-off between endurance and retention in memristive devices has been reported, thus optimization from materials and programming point of view must be thoroughly considered. In order to not only mimic human brain functionalities but also its efficiency, the device must be able to operate in the order of ~ 10 fJ per synaptic event. This is one of the most challenging aspects in synaptic device engineering, especially in highly scalable two-terminal devices since programming and reading of the states are done through the same terminals. This leads to another trade-off with device retention. A long data retention requires high state energy barrier to reduce the effect of external disturbance such as heat and electric field, however at the same time this energy barrier must be sufficiently low to achieve low programming energy requirement. Like high density storage devices, highly scalable device footprint is also desired for synaptic device applications to enable large scale neural network within compact chip dimension. In order to fully utilize the high scalability of the memristive devices, a two-terminal select device is required to facilitate pure crossbar array implementation. This can potentially add on to the challenging task of achieving linear and symmetrical weight update. Dynamic ratio is defined as the ratio of the highest conductance value to the lowest one. Higher dynamic ratio can be translated into

more superior mapping capability of the network. It also enables larger network connectivity in which maintaining sufficient read margin is crucial. These correlations among the device parameters post an enormous challenge in finding a reliable device that can provide excellent scalability while maintaining high synaptic performances. Thus, device, circuit, and algorithm-level co-optimization is needed.

Figure 1.1. Ideal analog synapse properties with gradual, linear, and symmetrical weight modulation under identical programming pulse condition with sufficient margin between the states and large dynamic ratio.

1.3. REDOX-BASED MEMRISTIVE DEVICES IN NEUROMORPHIC COMPUTING PLATFORMS

Redox-based memristive devices can be classified into two major groups, i.e., anion and cation-based devices. Their promising performances as synaptic devices have been widely investigated and demonstrated on different neuromorphic computing platforms. The underlying mechanism of different redox-based memristive structures might lead to a huge difference in macroscopic behavior of the device. Through structural engineering and rigorous optimization of device programming schemes, redox-based memristive devices have demonstrated highly stochastic memory behavior to significantly more deterministic features. To accommodate the different synaptic behaviors of these devices, various learning rules have also been implemented. In this section, synaptic properties of different redox-based memristive

systems under various learning rules with several system-level simulations are discussed. For simplicity, the redox-based memristive devices will be simply addressed as memristive devices or more specific as cation or anion devices.

1.1.1 Anion-based Synaptic Devices

The underlying mechanism of anion devices is based on the oxygen vacancy defects movement within the oxide layer under external electric field. The fundamental structure of an anion device consists of an oxide switching layer coupled with an inert electrode on one side and oxygen reservoir system on the other side, which can be in the form of reactive electrode (Ti, Hf, Ta, etc) or oxygen-deficient oxide layer. Anion devices have been reported to have high scalability of sub-10 nm [17-19], excellent reliability (endurance as high as 10^{12} and retention of more than 10 years) [20-22], multibit per cell capability, and low energy consumption. Anion devices initially emerged as one of the most promising candidates in non-volatile memory technology as both embedded memory and standalone memory for high density storage applications. In recent years, these devices have also attracted interest from neuromorphic computing and engineering community due to their desired characteristics. They have then been extensively studied and implemented as synaptic device for various neural network (NN) applications, mainly taking advantage of their high scalability and analog memory characteristic. Anion-based devices can be categorized into two major classes based on the switching nature of the device, i.e., localized (filamentary) and non-localized (non-filamentary) switching class. The difference between these two device classes is mainly on the active area involved during switching operation, with the former involves significantly smaller area than the latter.

Filamentary Devices

In general, the filamentary anion devices have an abrupt SET process, i.e., transition from low to high conductance state, while having a gradual RESET process, i.e., transition from high to low conductance state. The gradual RESET process is the main advantage of anion device over its cation counterpart to achieve gradual depression in synaptic device implementation. This is because achieving a gradual SET process during potentiation can be performed by controlling the compliance current level in 1-transistor-1-redox-memristor (1T1R) structure during the weight update. Thus, both gradual potentiation and depression can be achieved in anion devices. However, this approach still requires non-identical programming pulses and resistance state verification before the programming step, thus overhead on circuitry

is needed during the learning process. Furthermore, the variation in the amount of oxygen vacancies involved during the switching makes achieving symmetrical and linear weight update with sufficient read margin remains a huge challenge in these devices. Several structures, i.e., AlO_x , HfO_x and TaO_x -based structures, have been comprehensively investigated, improved, and implemented as synaptic devices to meet the requirements of an ideal synaptic device.

Aluminum Oxide (AlO_x)-based Devices

AlO_x -based anion devices have been investigated as both digital and analog memory devices under different systems, i.e., $\text{Ti}/\text{Al}_2\text{O}_3/\text{Pt}$ [23], $\text{TiN}/\text{Al}_2\text{O}_3/\text{Pt}$ [24], $\text{Ni}/\text{Al}_2\text{O}_3/\text{Pt}$ [25], $\text{CNT}/\text{AlO}_x/\text{CNT}$ [26], $\text{Ti}/\text{AlO}_x/\text{TiN}$ [27], and $\text{Al}/\text{AlO}_x/\text{Pt}$ [28]. In general, the reported AlO_x -based devices have high dynamic ratio (ranging from 10 to 1000), high scalability (down to 36 nm^2 device active area) [26], and low switching energy (below 2pJ) [24, 27]. The potentiation and depression characteristic of the AlO_x -based structure was experimentally tested in $\text{Ti}/\text{AlO}_x/\text{TiN}$ [27]. The linear gradual conductance change in both directions was achieved under non-identical pulses scheme. Different compliance currents (CCs) from 50 μA to 900 μA under 1.5 V, 500 μs voltage pulse were imposed during potentiation while different pulse amplitudes from -1V to -1.6V with 500 μs duration were used in depression mode. The device was able to achieve an average of 1.2% and 1.7% conductance change per programming pulse with 85 potentiation and 60 depression steps while maintaining ~ 10 dynamic ratio. However, due to non-identical pulses scheme required during the operation for both SET and RESET, significant overhead must be implemented on the peripheral circuit, which is not ideal for on-chip learning application.

AlO_x has relatively high oxygen scavenging immunity, which is reported to result in significantly smaller filament dimension [29]. While it is a desired property to achieve high ON/OFF ratio, high speed, and excellent uniformity, it also raises a challenge in achieving linear and symmetrical weight update. Thus, rather than being implemented as the main switching layer, AlO_x has been more widely used as an insertion layer interfacing the main switching layer to improve the synaptic performance of the memristive devices in terms of uniformity and linearity of the conductance update [30-37].

Hafnium Oxide (HfO_x)-based Devices

One of the first structures explored for synaptic device applications is HfO_x-based device with Ti oxygen reservoir electrode. Different weight update schemes have been demonstrated for this system, i.e., identical and non-identical pulses. Under identical pulses scheme, the device can only achieve gradual depression while still having abrupt potentiation. The gradual potentiation can be achieved under non-identical pulse condition in which the current flowing through the 1T1R device is closely controlled by pulsing the transistor's gate, resulting in a well-control filament formation. Several approaches have been implemented to enable gradual weight update in both directions under identical pulse condition, i.e., insertion of an oxide layer with less defect's mobility and thermal enhancement layer (TEL).[35, 38]

The first approach interfaced HfO₂ layer with AlO_x at the inert electrode side of the structure [35]. AlO_x layer has higher oxygen vacancy diffusion barrier compared to HfO₂ layer, which induced filament constriction at the AlO_x/HfO₂ interface. This promoted lateral filament modulation during potentiation and depression process. The devices are able to achieve gradual conductance change in both directions and further improved their linearity. In comparison with HfO₂/Ti structure, the improvement on the linearity, α , ($\alpha = 0$ represents the ideal linear and symmetrical update) of the potentiation in AlO_x/HfO₂/Ti system, i.e., from α value of 16.53 to -0.01, carried a noticeable trade-off in dynamic range of the device conductance (reduced from 10 to 3). The AlO_x/HfO₂/Ti synaptic device properties were simulated into NN implementing multilayer perceptron algorithm. It was evaluated under the Mixed National Institute of Standards and Technology (MNIST) dataset to test the NN accuracy in performing pattern recognition. It was shown that the improvement in linearity of the conductance change of the synapse was translated into significant increase in pattern recognition accuracy, i.e., from ~10% for HfO₂/Ti to ~90% for AlO_x/HfO₂/Ti structure.

The insertion of TEL in HfO_x-based device was designed based on gradual SET process observed during high temperature programming of the device [38]. HfO_x/Ti anion device was observed to exhibit abrupt SET under room temperature programming condition and gradual SET during cell programming at 150 °C. In order to obtain gradual SET process at room temperature operation, oxygen deficient TaO_x layer was introduced as TEL and oxygen reservoir in the structure replacing Ti electrode. This layer has significantly lower thermal conductivity compared to Ti electrode, which induced localized Joule heating effect across the active filament region during the switching process. This shifted the device switching property

from predominantly electric field to thermally induced switching. This resulted in the formation of multiple weak filaments instead of single filament switching, converting the abrupt into analog SET process, while maintaining dynamic ratio of 10. While it provides a promising solution to mitigate abrupt potentiation issue, the multiple weak filaments system has a trade-off in read disturb and retention of the conductance state. This will have negative impact on the amount of inferences the NN can perform while maintaining the weight values within acceptable deviation. This synaptic device has been experimentally demonstrated on a 1k-bit 1T1R array to carry out human face classification.

Tantalum Oxide (TaO_x)-based Devices

Another oxide system that has been widely investigated for synaptic device applications is TaO_x -based devices. One of the first reports on TaO_x -based devices was Ta_2O_{5-x}/TaO_{2-x} structure that demonstrated an excellent digital memory endurance capability of 10^{12} cycles under 10 ns operating speed [21]. Ta_2O_{5-x} was implemented as the oxide switching layer with an oxygen deficient TaO_{2-x} acting as the oxygen reservoir in the structure. Multilevel cell capability of this structure was demonstrated with an improved ON/OFF ratio of ~ 1000 with well separated 4 conductance levels (2 bits/cell) and 10 years extrapolated retention [39]. The application of this structure as synaptic device was demonstrated through rigorous optimization of pulse amplitude, pulse width, and the interval between subsequent pulses during operation to achieve gradual potentiation and depression [40]. This system was demonstrated as 2nd order memristor to realize synaptic plasticity in which different parameters involved during switching operation were considered. The modulation of conductive filament dimension, w , that directly results in the device conductance change was referred as 1st order parameter, utilizing memristor as a simple programmable memory device. On the other hand, the 2nd order memristor uses the local temperature, T , within the active switching region. T governs the evolution of the 1st order parameter, w , capturing the dynamics aspect of the device. T provides time-dependent variable that abruptly increases with the applied pulse and spontaneously decays after its removal. T enables the system to bio-realistically demonstrate activity-dependent plasticity, which is analogous to Ca^{+} concentration that regulates the weight-state variable.

Other TaO_x -based devices used Ti or TiO_x layer as oxygen reservoir in the system. Compared to the HfO_x -based devices, the proposed switching mechanism is based on predominant lateral modulation of filament width instead of vertical modulation of filament

gap connecting the electrodes [41]. This has been proven to be crucial in achieving gradual and linear weight update. Ta₂O₅/TiO_x synaptic device was implemented on simulated multilayer perceptron neural network under on-chip training condition by back-propagation algorithm. Even with dynamic ratio of ~5, the system was able to achieve almost 90% recognition accuracy using MNIST training data set [41].

Various efforts to improve the desired synaptic device characteristics of anion filamentary devices have been implemented from material and programming perspective. These devices tend to exhibit either abrupt potentiation and gradual depression with excellent dynamic ratio or gradual potentiation and depression with trade-off in dynamic ratio, as depicted in **Figure 1.2**.

Figure 1.2. General synaptic behavior of conventional filamentary anion devices under identical pulse condition, abrupt potentiation and gradual depression. (b) Gradual potentiation can be achieved through structural engineering in the expense of the device dynamic ratio

Non-filamentary Devices

Non-filamentary anion device utilizes interfacial defects movement between two layers of material, i.e., oxides and/or metals, which uniformly occurs across the entire device area. The change in the structural defects configuration under external electric field modulates the Schottky barrier at the interface causing significant change in device conductance, i.e., interfacial type switching. During the SET or RESET operation, it also alters the thickness of tunneling gap in the system, allowing higher or lower number of electrons flowing through the device. These enable gradual conductance change during the switching operation, mitigating

the issue encountered in most of the filamentary synaptic devices. Non-filamentary switching has been widely reported in many oxide structures, e.g., TiO_x , TaO_x , and WO_x .

Titanium Oxide (TiO_x)-based Devices

Analog characteristic in TiO_x systems have been demonstrated by the tuning the oxide layer stoichiometry and the use of oxygen gathering electrode. One of the early structures investigated for synaptic device was $\text{TiO}_x/\text{TiO}_y$ bilayer oxides system [42]. It was composed of ~50 nm sol-gel TiO_x layer grown on top of 6 nm TiO_y layer with defect ratio of ~0.23 and ~0.17, respectively. This created active interface between the two oxide layers in which the exchange of oxygen content occurred under external electric field. Gradual potentiation (4 MVcm^{-1} , 10 ms) and depression (-2 MVcm^{-1} , 10 ms) were obtained with dynamic ratio of ~10. The excellent device characteristics enabled its implementation on weight change, Spike-timing-dependent plasticity (STDP), and STDP triple model.

Engineering at the interface of the between the electrode and the active TiO_x switching layer have been reported to successfully improve the dynamic ratio as well as reduce the switching current of the TiO_x -based devices. Insertion of thin Al_2O_3 layer (~2 nm) at the interface between TiO_2 and TiN electrode could achieve memory window of >100 with <10 μA switching current [43]. Further improvement of switching current, i.e., down to ~1 μA was demonstrated by replacing Al_2O_3 with a-Si layer [44, 45]. a-Si plays a role as an oxygen gathering layer facilitating the movement of oxygen ions at the interface. The semi-insulating property of a-Si layer enabled nonlinear IV cell characteristics, which caused amplification of the energy barrier modulation leading to the large dynamic ratio of the device. The synaptic characteristics of a-Si/ TiO_2 -based devices were input into a simulated 3-layer ANN and the pattern recognition accuracy of the NN was tested using MNIST database. The focus of the demonstration was on investigating the effect of read noise, i.e., random telegraph noise (RTN), on the pattern recognition accuracy. a-Si/ TiO_2 -based device achieved much better accuracy compared to filamentary TaO_x -based devices due to lower RTN amplitude value and distribution with much less noise occurrence rate [46].

TiO_x has also been implemented on various bilayer oxide systems [47], i.e., AlO_x , TaO_x , WO_x , HfO_x , ZnO_x , and SiO_x . Different dynamic ratio and multibit capability performances were achieved with oxides paired with TiO_2 . The most promising multilevel conductance states

property was found in $\text{AlO}_y/\text{TiO}_x$ bilayer oxide structure, which was able to achieve non-overlapping conductance states of 6.5 bit per cell despite of less than 10 dynamic ratio. This was attributed to AlO_x property being the oxide with lowest oxygen ions mobility among the pairing oxide layers tested. Although no synaptic characteristics were especially discussed, well control conductance update can be achieved with different pulse schemes. Specific conductance level can be achieved from the same starting value using train of identical pulses or single pulse with optimized amplitudes and durations. Together with the non-overlapping conductance states achieved, this showed a promising feature towards an ideal analog synaptic device characteristic.

The electrode engineering in TiO_x system was associated with the symmetry characteristics of the system. $\text{TiN}/\text{TiO}_x/\text{Mo}$ system was found to improve the symmetry of the system as compared to $\text{TiN}/\text{TiO}_x/\text{Pt}$ [48]. It was due to the work function difference between the corresponding two electrodes. 64 conductance levels with excellent distribution was achieved. The device potentiation exhibited more linear conductance change as compared to its depression. In order to improve the linearity of the device depression and thus achieving more symmetrical weight update, current pulse scheme was adopted. The hybrid scheme of voltage (potentiation) and current (depression) pulse was able to improve the pattern recognition accuracy by around 10%.

Tantalum Oxide (TaO_x)-based Devices

$\text{Ta}/\text{TaO}_x/\text{TiO}_2/\text{Ti}$ structure was initially proposed as 3D-integrated storage class memory [49, 50]. It has high endurance of 10^{12} , forming free, self-compliant, and self-rectifying characteristics that significantly simplify the peripheral circuit required during operation. It works based on homogenous Schottky barrier modulation due to oxygen vacancy defects migration at Ta/TaO_x interface under external electric field. In this structure, TiO_2 layer provided diode-like effect that resulted in self-rectifying characteristics in the structure with rectification ratio of $\sim 10^5$. The switching mechanism was confirmed by simulation to accurately reproduce experimentally obtained DC and AC characteristics of the device [51]. Moreover, its synaptic characteristics, i.e., long term potentiation (LTP), long term depression (LTD), STDP and paired-pulse facilitation (PPF), have also been experimentally investigated [51, 52]. The structure exhibited non-linear gradual potentiation and depression with dynamic ratio of >2 under identical pulse scheme (LTP: +3 V/5ms, LTD: -3 V /5 ms, and read: -1.5 V /1 ms). The training pulse duration was found to linearly scale with pulse amplitude required

to maintain similar synaptic plasticity [52]. Extremely low <10 fJ per synaptic event was experimentally recorded [53]. The nonlinearity of the weight update could be improved under two different pulse schemes, i.e., state-independent unipolar pulse scheme (UPS) and bipolar pulse scheme (BPS) [52]. UPS used single pulse (positive or negative) to move the weight value up or down, while BPS utilized a pair of pulses of different polarities (positive-high, negative-low or negative-high, positive low) to run one cycle of weight update. The linearity of the weight update was improved from 0.6-0.81 (UPS) to 0.42-0.54 (BPS) with $\sim 50\%$ trade-off in the dynamic ratio of the weights. This device characteristics were implemented in the simulation of the training evolution of 8×8 binary pattern. BPS achieved $\sim 90\%$ accuracy, which was significantly higher than $\sim 75\%$ accuracy attained under UPS. This showed the importance of weight update linearity in the long run to provide more immunity to input noise.

Other than insertion of TiO_2 layer, non-filamentary TaO_x devices have also been paired with Al_2O_3 barrier layer. Different deposition techniques were used. i.e., electron beam evaporation, post-rapid thermal annealing in O_2 ambient, and ALD. Different deposition techniques resulted in significantly different initial resistance values. When both films are deposited via e-beam evaporation (AlO_x 3 nm and TaO_x 5 nm), the initial resistance of the structure was found to be around 200Ω , which is extremely low. This was due to the loss of oxygen content during the deposition. On the other hand, the films deposited via post-rapid thermal annealing in O_2 ambient and ALD consistently started from highly insulating state. This is in line in which the two techniques tend to result in stoichiometric films. Other than RESET process required to initiate the switching operation in the non-stoichiometric structures, the switching polarity and mechanisms involved during the operation remained the same. From current-voltage characteristics, area-dependent LRS, and elemental analysis, the change in conductance during operation was attributed to the tunneling barrier modulation induced by oxygen ions migration across the whole area of TaO_x /electrode interface. With the tunable gradual SET/RESET feature of the device, LTP/LTD, PPF, and STDP were demonstrated. LTP and LTD were characterized under identical pulse scheme (50 pulses) with different pulse amplitude (4.5 to 5.5 V) and duration (1 μs to 100 μs). Estimated energy of 50 pJ per spike of programming pulse was recorded. Increase in linearity of conductance change was observed with the decrease in both pulse amplitude and duration at the expense of the weight dynamic ratio. Improvement in linearity and dynamic ratio could be achieved under non-identical training pulse scheme with increasing pulse amplitude (2 V to 6 V of 100 μs pulse). Under this

scheme, rough and fine tuning to achieve certain weight value from any randomly chosen weight with excellent <1% variation was also shown.

Tungsten Oxide (WO_x)-based Devices

Another extensively studied oxide structure with underlying mechanism of homogenous anions migration across device active area is WO_x. The migration of the oxygen ions enables the system to tune the interchanging role of Schottky barrier emission and tunneling as predominant conduction mechanism during the operation. In Pt/WO_x/Ti structure [54], the Schottky barrier at the interface of Pt/WO_x is formed due to the higher work function of Pt compared to Ti. During the SET process in which the Pt electrode was positively biased, the oxygen ions migrated towards the Pt electrode and got accumulated at the Pt/WO_x interface. This reduced the Fermi level near the WO_x surface and at the same time decreased the Schottky barrier height between Pt and WO_x, resulting in the increase of device conductance. This specific structure was demonstrated on a flexible substrate. The synaptic properties of the device, i.e., excitatory postsynaptic current (EPSC), PPF, STP/LTP, and STDP were characterized and no performance degradation occurred under large angle bending or 100 times bending tests. EPSC property of the device was experimentally obtained through device dynamic response upon receiving 2V, 50 ms programming pulse. Immediately after the removal of the electric field, the conductance of the device started to drop and eventually relaxed back to the initial conductance value after ~400 ms. PPF was determined through the ratio of EPSC peaks obtained by sending two identical pulses (2 V, 50 ms). The correlation of the PPF and the interval between the subsequent pulses was recorded up to 1 s. It was well fitted with double exponential function containing the initial facilitation magnitudes and characteristic relaxation times of the PPF. The same function can also be used to correlate the retention characteristics of the device. The extracted time constants described the transition between STP to LTP under different number of subsequent programming pulses. STDP function of the structure was characterized by sending a pair of pulses with opposite polarity (+2 V and -2V, 50 ms) to top and bottom electrode as pre- and post- synaptic spike. The relative change in weight value was recorded under different interval of the pulse pairs.

Another structure that has been investigated as synaptic device even earlier than Pt/WO_x/Ti was Pd/WO_x/W [55, 56]. Despite the difference in the electrodes implemented, similar homogenous switching and conduction mechanisms were obtained. However, under lower programming voltage of 1.3 V with shorter 1 ms duration, this structure was able to

achieve better retention characteristics. This could be attributed to the smaller difference in electrode work functions of Pd-W as compared to Pt-Ti pair [57]. Thus, the choice of electrodes used in the structure plays critical role in determining the operating voltage and the temporal dynamics of the device.

Despite the promising performance in terms of gradual weight update symmetry and linearity, non-filamentary anion devices tend to have high programming voltages, which might not be suitable for 1T1R integration. This currently limits the implementation of the various non-filamentary devices only on small scale neural network. Furthermore, the devices have significant trade-off between device latency and retention capability. Thus, more optimization is still required to get closer to ideal synaptic device characteristics.

Figure 1.3. Gradual potentiation and depression with more deterministic nature can be achieved in most of the reported non-filamentary anion devices. However, the excellent synaptic properties in terms of weight update linearity and symmetry are usually accompanied by low dynamic ratio, extremely low device conductance, and high switching voltages.

Comparing anion non-filamentary device to its filamentary counterpart, it can achieve a more deterministic behaviour enabling gradual potentiation and depression. However, it tends to have low dynamic ratio. Even more so with the programming approaches required to improve the linearity of the weight updates, as shown in **Figure 1.3**.

1.1.2 Cation-based Synaptic Devices

Cation-based devices work based on the formation and dissolution of metallic filaments within switching layer under external electric field. These devices are also known as conductive-bridge RAM (CBRAM) or electrochemical-metallization memory (ECM). The most commonly used active metal electrodes are Ag and Cu with electrochemically inert electrodes such as Au, Pt, and Ir [58]. Wide variety of compounds have been investigated as switching layer, which can be classified into three major groups, i.e., solid electrolytes, oxides, and nitrides [58]. They have been known to have promising characteristics in terms of scalability, switching speed, and programming power. In general, they also have lower operating voltage compared to their anion devices counterpart. These desirable properties are due to the high mobility of Cu and Ag ions within the switching layer. While having high ions mobility is beneficial in terms of programming speed and power, it also raises challenges in device reliability, i.e., achieving high endurance and long retention. The device failure has been reported to mainly due to excessive amount of metal species residing inside the switching host. Furthermore, it also leads to generally abrupt and stochastic switching operation. These challenges have especially been hindering the cation-based devices application as artificial synapses in NN.

Based on the amount of metal cations involved during the switching operation, the cation-based devices can be divided into two categories, i.e., infinite and finite cations source devices.

Infinite Cations Source

Infinite cations source devices refer to devices that rely on active metal electrodes as the source of the cations to facilitate the switching operation, as depicted on **Figure 1.4**. This configuration virtually enables infinite amounts of cations responsible for the conductance change during the device operation. In agreement with the aforementioned challenges, the amount of metal species migrating within the switching layer in this type of devices plays a critical role in the uniformity and reliability of the device, especially in obtaining multilevel conductance characteristics for analog synaptic device applications.

Figure 1.4. The switching operation of filamentary cation-based devices with infinite cation supply.

From the device programming viewpoint, multilevel conductance switching has been demonstrated in cation-based devices by implementing different compliance current values during the device operation [58-60]. Different compliance currents lead to different amount of active metal ions injected and different conductive filament dimensions, allowing the device to have different values of conductance. This operating scheme requires the use of a transistor to work in tandem with the memristive element to provide a precise current control through the device. Thus, it limits the array level implementation to active array (1T1R) in which the footprint of a single synapse will be limited by the transistor size. To achieve multibit per cell capability in the device, constant drain to source voltage is required, while different voltage pulse amplitudes are implemented to allow different current level flowing through the memristive device. This weight update scheme will require prior reading of the conductance state before moving upward or downward on the weight level. This will significantly slow down the training process and increase the amount of programming energy due to additional overhead on the network circuitry. While this architecture provides solution to achieve gradual long-term potentiation behavior during SET process, emulating the same characteristics for long-term depression during RESET process remains a challenge.

Different approaches from materials design and engineering perspective have also been investigated to achieve a better control over the amount of the metal species driven under

external electric field to mitigate the stochastic switching nature of the device as well as abrupt RESET process. This is extremely important towards realizing the ideal analog deterministic synapse characteristic. The first approach is done by scaling down the active device area involved during the switching operation. This can significantly reduce the amount of active metal species injected into the switching layer under external electric field. The use of plug structure to scale down the electrode to sub-20 nm area has been evidently improved the switching uniformity and reliability [58, 61-63]. The scaling was further extended to switching layer area of the device to sub-30 nm dimension [64]. With smaller switching area, the electrochemical reaction and the movement of the active metal species becomes more restricted, which resulted in improved uniformity and data retention [64].

The second approach utilizes a thin insertion layer to either prevent unwanted oxide formation at the active electrode/oxide interface or to obtain a better control of cations injection and filament formation during device operation. Insertion of Ti at the interface of Cu/TaO_x-based devices reduced the cycle-to-cycle and device-to-device variation with significant improvement on device dynamic ratio (from ~10 to ~100) [65]. This was attributed to the formation of TiO_x instead of CuO_x at the interface of Cu/Ti/TaO_x structure. Insertion of thin TiW layer at the interface of Cu/AlO_x has also been shown to improve the overall performance of the device [66]. This barrier layer helped to maintain the cell structural integrity up to BEOL processing temperature of 400 °C. It also prevented gradual drifting of conductance states due to parasitic diffusion effects, resulted in excellent cycling control. The W\Al₂O₃\TiW\Cu cell fabricated on 90 nm W plug exhibited high voltage-disturb immunity with high dynamic ratio of >100 and fast switching operation of ~10 ns with <3 V pulse amplitude. The dynamic ratio was further enhanced to ~1000 by the insertion of WO_x by thermal oxidation of the W plug at 500 °C. It was ascribed to a filament constriction at WO_x/Al₂O₃ interface obtaining an hourglass conductive filament shape that enabled deeper RESET process. In Al/Cu/GeSe_x/TaO_x/W [67], TaO_x insertion layer at the inert electrode side of the GeSe_x switching layer provided an additional layer with lower Cu mobility to alter the filament shape and dimension during the switching. Improvement in switching stability was attributed to nanofilament confinement within TaO_x layer.

The third approach uses a mixture in the form of metal alloy as the source of cations. The first example is copper tellurium (Cu_xTe_{1-x}) as active ions source. It was first demonstrated on 180 nm CMOS technology in CuTe/GdO_x/W structure [60]. It was able to achieve excellent 2-bit memory property with excellent retention under different compliance current levels.

Dynamic ratio of ~ 1000 ($10\text{ M}\Omega/10\text{ k}\Omega$) was achieved under programming parameters of 3 V, 110 μA , and 5 ns for SET and -1.7 V, 125 μA , and 1 ns for RESET. The device also showed potential of gradual RESET, but further optimization of pulse amplitude and width were required. The effect of Cu and Te composition for the alloy of active electrode was investigated in $\text{Cu}_x\text{Te}_{1-x}/\text{Al}_2\text{O}_3/\text{Si}$ cells [68]. It was found that the device exhibited volatile switching (SET), non-volatile switching with gradual RESET, and non-volatile switching with abrupt RESET as the Te content decreases. This was associated with higher energy barrier to inject Cu into Al_2O_3 for Cu-Te phase compared to pure Cu. This provides a very useful insight on how the cation source characteristic is able to tune the overall memristive cell characteristic for specific applications.

An alternative implementation of the cation devices was proposed through stochastic STDP learning rules. Instead of trying to precisely control the switching operation to produce analog deterministic behavior, the binary probabilistic switching nature of the device is being exploited under these learning rules. This approach provides the equivalent system level functionalities to that of network utilizing the analog synaptic devices under deterministic learning rules [69]. Supervised and unsupervised NNs have been demonstrated using the binary probabilistic synapses [70, 71]. The unsupervised NN was demonstrated using 1T1R and 1R system with Ag/GST memristive structure as the synaptic element. Strong (pulse duration of $\geq 10\mu\text{s}$) and weak (pulse duration of $\leq 1\mu\text{s}$) programming conditions were used to toggle between relatively more deterministic and probabilistic switching operation within the same device. In the strong programming condition, it was observed that high resistance state (HRS) of the device had larger distribution compared to the low resistance state (LRS), which could be attributed to uncontrollable metal filaments dissolution during reset process. With each device carries certain degree of stochasticity, in array level, this characteristic is amplified with the presence of device-to-device variation. The 1T1R and 1R synapse crossbar array was implemented in the core circuit, together with input/output CMOS neuron and pseudo-random number generator (PRNG) circuit under leaky integrate and fire (LIF) neuron model. Specific programming schemes for 1T1R and 1R were implemented to handle asynchronous analog streams of data for unsupervised pattern extraction and recognition. Excellent performance parameters were achieved for auditory pattern sensitivity (> 2.5) and video detection rate (95%), while maintaining extremely low power dissipation of 0.55 μW and 74.2 μW for audio and video demonstrator respectively [71].

Finite Cations Source

Another type of device has also been engineered to improve analog properties of the cation-based devices, in which a fixed amount of metal species within the switching layer is used rather than an active electrode as a source of metal ions. This approach prevents the formation of a localized conductive path during the switching operation, unlike the conventional cation-based devices. This technique was first implemented by sandwiching Ag-doped amorphous Si in between two inert electrodes [72]. The structure was fabricated using co-sputtering technique of Ag and Si to form a gradient mixture of Ag:Si across the switching layer. This results in the presence of rich and poor Ag region that can be modulated under external electric field. The structure successfully achieved gradual conductance change in both potentiation and depression process under identical programming pulse scheme. The device was also integrated with CMOS-based neuron circuits to demonstrate spike timing dependent plasticity (STDP) learning rules. The same approach was successfully adopted in Ag:TiO_x [73]. The device was able to demonstrate the learning and memory functionalities including STDP, PPF, and STP to LTP transition, with an improved timescale of hundreds of nanoseconds as compared to microseconds pulse used in Ag:Si devices.

A slightly different approach was implemented in Ag-doped WO_x [74] and TaO_x [75, 76] with uniform Ag content across the switching layer. The Ag-doped WO_x demonstrated the tunability of the device characteristics with different Ag content. Low Ag concentration within WO_x leads to volatile switching behavior that was able to mimic the forgetting effect of human memory. While, the devices with relatively higher Ag contents enable analog non-volatile switching properties. The Ag:TaO_x device was fabricated via self-doping during the sputtering process. The TaO_x layer was deposited on top of Ag electrode, resulting in the intermixing layer at the interface. The presence of the Ag:TaO_x layer at the interface of Ag and TaO_x layer caused a double switching behavior under different external electric field. The device successfully demonstrated the key synaptic behaviors such as STP, LTP, and spike-rate-dependent plasticity (SRDP).

1.4. RESEARCH OBJECTIVES

Having such promising features to offer as a synaptic element in neuromorphic computing platform, the development of the redox-based memristive devices have been hindered by major research challenges in array level scalability, as well as conductance update linearity and symmetry. The ideal synaptic device properties might not be possible to be

obtained within a single device structure due to the presence of significant trade-off among them. Fortunately, these requirements are highly dependent on the application. Thus, systematic approach in addressing these trade-offs is required to enable tuning of the device properties to meet certain requirements for a specific target application.

In order to fully utilize the scalability of the two-terminal memristive devices, there are two key requirements to be met. The first requirement is the availability of a compatible two-terminal select device to mitigate the inherent sneak-path current issue in a large crossbar (1S1R) array implementation. An alternative approach is having a self-rectifying characteristic within the memristive device itself, thus removing the need of a select device altogether. The second requirement is to avoid the need of an external current limiter during the device switching operation, i.e., compliance free. During the forming and SET process, the current flowing through the device usually must be limited by a series transistor (1T1R) to prevent permanent damage on the device. However, the transistor is known to be the scalability bottleneck in large array integration. The compromise integration approach between 1S1R and 1T1R is 1TnR (1 transistor for n memory elements) [77]. Under 1TnR, the number of parallel memristive elements connected in series with every local line selection transistor (LLST) can be adjusted. Each 1TnR row or column is the equivalent of each row or column in crossbar array in terms of the cells' connectivity. This reduces the number of transistor usage as compared to 1T1R architecture. Furthermore, the presence of the transistor enables compliance current implementation during the device operation, thus controlling the forming and SET process more precisely.

Different learning approaches are available for various artificial neural networks (ANNs), i.e., *ex-situ* and *in-situ* learning approach. Due to the significantly different nature of the weight update required during the training of the network, different synaptic device requirements emerge. For *ex-situ* training, the endurance of the synaptic devices in the array is not as critical as the ones used for *in-situ* training. This is due to much more frequent weight update needed during *in-situ* training. For *in-situ* training, the conductance response of the devices must be as linear and symmetrical as possible with respect to the number of programming pulses during the training. This is because in *ex-situ* training, this inherent non-linearity and asymmetry can be mitigated by multiple iterations of write-verify scheme, while it will be very challenging to implement in *ex-situ* training.

This work aims to address the aforementioned challenges through further understanding of the underlying physics governing the switching dynamics of redox-based memristive devices. The primary aim will be accomplished by achieving the following research objectives:

1. To develop and characterize anion-based memristive devices with analog switching features for synaptic device applications.
2. To develop and characterize cation-based memristive devices with volatile switching capability towards highly scalable select device applications.
3. To develop theoretical framework of the developed devices that can be used to further optimize and improve the synaptic device performance.

1.5. THESIS ORGANIZATION

The thesis consists of 6 chapters, i.e., introduction, research methodology, 3 chapters of experimental findings and analysis, as well as conclusion and future work. **Chapter 2** describes the details of experimental techniques implemented from device fabrication to electrical and material characterizations. **Chapter 3** describes the switching dynamics of anion-based Pt/HfO_x/Ti memristive devices under different electroforming treatments. The multilevel conductance states capability of the devices was also investigated, and its origin is correlated with the modulation of charge trapping levels under external electric field. **Chapter 4** presents the approach in which the developed memristive devices can be implemented as a synaptic element in neuromorphic computing systems. The trade-off among the device properties are required to facilitate different learning approaches, i.e., *ex-situ* and *in-situ* learning. **Chapter 5** presents the findings on cation-based diffusive memristor devices. The volatile and non-volatile behaviour of the device are demonstrated under different electroforming treatments. The temporal response as well as reliability aspect of the device are also discussed. **Chapter 6** summarizes and concludes the key theoretical and experimental findings presented in this thesis. Several suggestions on the future research to further advance the corresponding field are also presented.

1.6. REFERENCES

- [1] G. E. Moore, "Progress in Digital Integrated Electronics," in *Electron Devices Meeting*, 1975, vol. 21, pp. 11-13.
- [2] G. E. Moore, "Cramming More Components onto Integrated Circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff," *IEEE Solid-State Circuits Society Newsletter*, vol. 11, no. 3, pp. 33-35, 2006.
- [3] M. T. Bohr and I. A. Young, "CMOS Scaling Trends and Beyond," *IEEE Micro*, vol. 37, no. 6, pp. 20-29, 2017.
- [4] T. Ghani *et al.*, "A 90nm High Volume Manufacturing Logic Technology Featuring Novel 45nm Gate Length Strained Silicon CMOS Transistors," in *IEEE International Electron Devices Meeting 2003*, 2003, pp. 11.6.1-11.6.3.
- [5] K. Mistry *et al.*, "A 45nm Logic Technology with High-k+Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging," in *2007 IEEE International Electron Devices Meeting*, 2007, pp. 247-250.
- [6] K. Ronse *et al.*, "Lithography Options for the 32nm Half Pitch Node and Beyond," in *2008 IEEE Custom Integrated Circuits Conference*, 2008, pp. 371-378.
- [7] C. Auth *et al.*, "A 22nm High Performance and Low-Power CMOS Technology Featuring Fully-Depleted Tri-Gate Transistors, Self-Aligned Contacts and High Density MIM Capacitors," in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, pp. 131-132.
- [8] N. Waldron *et al.*, "Replacement Fin Processing for III–V on Si: From Finfets to Nanowires," *Solid-State Electronics*, vol. 115, pp. 81-91, 2016.
- [9] N. Loubet *et al.*, "Stacked Nanosheet Gate-All-Around Transistor to Enable Scaling Beyond Finfet," in *2017 Symposium on VLSI Technology*, 2017, pp. T230-T231.
- [10] H. Zhong-Fang, R. Guo-Ping, and R. Gang, "A Simulation Study of Vertical Tunnel Field Effect Transistors," in *2011 9th IEEE International Conference on ASIC*, 2011, pp. 665-668.
- [11] S. Singh and B. Raj, "Vertical Tunnel-FET Analysis for Excessive Low Power Digital Applications," in *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, 2018, pp. 192-197.
- [12] D. Kwon *et al.*, "Negative Capacitance FET With 1.8-nm-Thick Zr-Doped HfO₂ Oxide," *IEEE Electron Device Letters*, vol. 40, no. 6, pp. 993-996, 2019.

- [13] M. H. Moaiyeri, A. Rahi, F. Sharifi, and K. Navi, "Design and Evaluation of Energy-Efficient Carbon Nanotube FET-Based Quaternary Minimum and Maximum Circuits," *Journal of Applied Research and Technology*, vol. 15, no. 3, pp. 233-241, 2017/06/01/ 2017.
- [14] C. Koch, *Biophysics of Computation: Information Processing in Single Neurons (Computational Neuroscience Series)*. Oxford University Press, Inc., 2004.
- [15] G. W. Burr *et al.*, "Neuromorphic Computing Using Non-Volatile Memory," *Advances in Physics: X*, vol. 2, no. 1, pp. 89-124, 2017/01/02 2017.
- [16] S. B. Laughlin and T. J. Sejnowski, "Communication in Neuronal Networks," *Science*, vol. 301, no. 5641, p. 1870, 2003.
- [17] B. Govoreanu *et al.*, "10×10nm² Hf/HfO_x Crossbar Resistive RAM with Excellent Performance, Reliability and Low-Energy Operation," in *2011 International Electron Devices Meeting*, 2011, pp. 31.6.1-31.6.4.
- [18] L. Kai-Shin *et al.*, "Utilizing Sub-5 Nm Sidewall Electrode Technology for Atomic-Scale Resistive Memory Fabrication," in *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers*, 2014, pp. 1-2.
- [19] M. J. Kim *et al.*, "Low Power Operating Bipolar TMO Reram for Sub 10 nm Era," in *2010 International Electron Devices Meeting*, 2010, pp. 19.3.1-19.3.4.
- [20] H. Y. Lee *et al.*, "Low Power and High Speed Bipolar Switching with A Thin Reactive Ti Buffer Layer in Robust HfO₂-Based RRAM," in *2008 IEEE International Electron Devices Meeting*, 2008, pp. 1-4.
- [21] 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," *Nature Materials*, vol. 10, no. 8, pp. 625-630, 2011/08/01 2011.
- [22] D. J. J. Loy, P. A. Dananjaya, X. L. Hong, D. P. Shum, and W. S. Lew, "Conduction Mechanisms on High Retention Annealed MgO-based Resistive Switching Memory Devices," *Scientific Reports*, vol. 8, no. 1, p. 14774, 2018/10/03 2018.
- [23] C.-Y. Lin, C.-Y. Wu, C.-Y. Wu, C. Hu, and T.-Y. Tseng, "Bistable Resistive Switching in Al₂O₃ Memory Thin Films," *Journal of The Electrochemical Society*, vol. 154, no. 9, pp. G189-G192, 2007.
- [24] Y. Wu, S. Yu, B. Lee, and P. Wong, "Low-Power Tin/Al₂O₃/Pt Resistive Switching Device with Sub-20 μA Switching Current and Gradual Resistance Modulation," *Journal of Applied Physics*, vol. 110, no. 9, p. 094104, 2011.

- [25] B. Sarkar, B. Lee, and V. Misra, "Understanding the Gradual Reset in Pt/Al₂O₃/Ni RRAM for Synaptic Applications," *Semiconductor Science and Technology*, vol. 30, no. 10, p. 105014, 2015/08/24 2015.
- [26] Y. Wu, Y. Chai, H.-Y. Chen, S. Yu, and H.-S. P. Wong, "Resistive Switching AlO_x-Based Memory With CNT Electrode for Ultra-Low Switching Current and High Density Memory Application," in *2011 Symposium on VLSI Technology-Digest of Technical Papers*, 2011, pp. 26-27: IEEE.
- [27] Y. Wu *et al.*, "AlO_x-Based Resistive Switching Device With Gradual Resistance Modulation For Neuromorphic Device Application," in *2012 4th IEEE International Memory Workshop*, 2012, pp. 1-4: IEEE.
- [28] K. Park and J.-S. Lee, "Reliable Resistive Switching Memory Based on Oxygen-Vacancy-Controlled Bilayer Structures," *RSC Advances*, 10.1039/C6RA00798H vol. 6, no. 26, pp. 21736-21741, 2016.
- [29] L. Goux *et al.*, "Understanding of the Intrinsic Characteristics and Memory Trade-Offs of Sub-μA Filamentary RRAM Operation," in *2013 Symposium on VLSI Technology*, 2013, pp. T162-T163: IEEE.
- [30] G. C. Adam, B. D. Hoskins, M. Prezioso, F. Merrih-Bayat, B. Chakrabarti, and D. B. Strukov, "3-D Memristor Crossbars for Analog and Neuromorphic Computing Applications," *IEEE Transactions on Electron Devices*, vol. 64, no. 1, pp. 312-318, 2016.
- [31] W. Banerjee, X. Xu, H. Lv, Q. Liu, S. Long, and M. Liu, "Variability Improvement of TiO_x/Al₂O₃ Bilayer Nonvolatile Resistive Switching Devices by Interfacial Band Engineering with an Ultrathin Al₂O₃ Dielectric Material," *ACS Omega*, vol. 2, no. 10, pp. 6888-6895, 2017.
- [32] K.-C. Chuang *et al.*, "Impact of the Stacking Order of HfO_x and AlO_x Dielectric Films on RRAM Switching Mechanisms to Behave Digital Resistive Switching and Synaptic Characteristics," *IEEE Journal of the Electron Devices Society*, vol. 7, pp. 589-595, 2019.
- [33] L. Goux *et al.*, "Asymmetry And Switching Phenomenology In Tin\((Al_2O_3)\)HfO₂\Hf Systems," *ECS Solid State Letters*, vol. 1, no. 4, pp. P63-P65, 2012.
- [34] W. Song *et al.*, "Analog Switching Characteristics In Tiw/Al₂O₃/Ta₂O₅/Ta RRAM Devices," *Applied Physics Letters*, vol. 115, no. 13, p. 133501, 2019.

- [35] J. Woo *et al.*, "Improved Synaptic Behavior under Identical Pulses Using $\text{AlO}_x/\text{HfO}_2$ Bilayer RRAM Array for Neuromorphic Systems," *IEEE Electron Device Letters*, vol. 37, no. 8, pp. 994-997, 2016.
- [36] S. Yu, Y. Wu, Y. Chai, J. Provine, and H.-S. P. Wong, "Characterization of Switching Parameters and Multilevel Capability in $\text{HfO}_x/\text{AlO}_x$ Bi-Layer RRAM Devices," in *Proceedings of 2011 International Symposium on VLSI Technology, Systems and Applications*, 2011, pp. 1-2: IEEE.
- [37] Y. Sun *et al.*, "A $\text{Ti}/\text{AlO}_x/\text{TaO}_x/\text{Pt}$ Analog Synapse for Memristive Neural Network," *IEEE Electron Device Letters*, vol. 39, no. 9, pp. 1298-1301, 2018.
- [38] W. Wu, H. Wu, B. Gao, N. Deng, S. Yu, and H. Qian, "Improving Analog Switching in HfO_x -Based Resistive Memory with a Thermal Enhanced Layer," *IEEE Electron Device Letters*, vol. 38, no. 8, pp. 1019-1022, 2017.
- [39] S. R. Lee *et al.*, "Multi-Level Switching Of Triple-Layered TaO_x RRAM with Excellent Reliability for Storage Class Memory," in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, pp. 71-72: IEEE.
- [40] S. Kim, C. Du, P. Sheridan, W. Ma, S. Choi, and W. D. Lu, "Experimental Demonstration of a Second-Order Memristor and Its Ability to Biorealistically Implement Synaptic Plasticity," *Nano letters*, vol. 15, no. 3, pp. 2203-2211, 2015.
- [41] J. Woo, A. Padovani, K. Moon, M. Kwak, L. Larcher, and H. Hwang, "Linking Conductive Filament Properties and Evolution to Synaptic Behavior of RRAM Devices for Neuromorphic Applications," *IEEE Electron Device Letters*, vol. 38, no. 9, pp. 1220-1223, 2017.
- [42] K. Seo *et al.*, "Analog Memory and Spike-Timing-Dependent Plasticity Characteristics of a Nanoscale Titanium Oxide Bilayer Resistive Switching Device," *Nanotechnology*, vol. 22, no. 25, p. 254023, 2011/05/16 2011.
- [43] B. Govoreanu *et al.*, "Vacancy-Modulated Conductive Oxide Resistive RAM (VMCO-RRAM): An Area-Scalable Switching Current, Self-Compliant, Highly Nonlinear and Wide On/Off-Window Resistive Switching Cell," in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 10.2.1-10.2.4.
- [44] B. Govoreanu *et al.*, "Advanced A-VMCO Resistive Switching Memory through Inner Interface Engineering with Wide ($>10^2$) On/Off Window, Tunable μA -Range Switching Current and Excellent Variability," in *2016 IEEE Symposium on VLSI Technology*, 2016, pp. 1-2.

- [45] B. Govoreanu *et al.*, "A-VMCO: A Novel Forming-Free, Self-Rectifying, Analog Memory Cell with Low-Current Operation, Nonfilamentary Switching and Excellent Variability," in *2015 Symposium on VLSI Technology (VLSI Technology)*, 2015, pp. T132-T133.
- [46] Z. Chai *et al.*, "Impact of RTN on Pattern Recognition Accuracy of RRAM-Based Synaptic Neural Network," *IEEE Electron Device Letters*, vol. 39, no. 11, pp. 1652-1655, 2018.
- [47] S. Stathopoulos *et al.*, "Multibit Memory Operation of Metal-Oxide Bi-Layer Memristors," *Scientific Reports*, vol. 7, no. 1, p. 17532, 2017/12/13 2017.
- [48] J. Park, M. Kwak, K. Moon, J. Woo, D. Lee, and H. Hwang, "TiO_x-Based RRAM Synapse with 64-Levels of Conductance and Symmetric Conductance Change by Adopting a Hybrid Pulse Scheme for Neuromorphic Computing," *IEEE Electron Device Letters*, vol. 37, no. 12, pp. 1559-1562, 2016.
- [49] C. Hsu *et al.*, "Self-Rectifying Bipolar TaO_x/TiO₂ RRAM with Superior Endurance Over 10¹² Cycles for 3D High-Density Storage-Class Memory," in *2013 Symposium on VLSI Technology*, 2013, pp. T166-T167.
- [50] C.-W. Hsu *et al.*, "Homogeneous Barrier Modulation of TaO_x/TiO₂ Bilayers for Ultra-High Endurance Three-Dimensional Storage-Class Memory," *Nanotechnology*, vol. 25, no. 16, p. 165202, 2014/03/25 2014.
- [51] Y.-F. Wang, Y.-C. Lin, I. T. Wang, T.-P. Lin, and T.-H. Hou, "Characterization and Modeling of Nonfilamentary Ta/TaO_x/TiO₂/Ti Analog Synaptic Device," *Scientific Reports*, Article vol. 5, p. 10150, 05/08/online 2015.
- [52] I. T. Wang, C.-C. Chang, L.-W. Chiu, T. Chou, and T.-H. Hou, "3D Ta/TaO_x/TiO₂/Ti Synaptic Array and Linearity Tuning of Weight Update for Hardware Neural Network Applications," *Nanotechnology*, vol. 27, no. 36, p. 365204, 2016/08/02 2016.
- [53] I. Wang, Y. Lin, Y. Wang, C. Hsu, and T. Hou, "3D Synaptic Architecture with Ultralow Sub-10 fJ Energy per Spike for Neuromorphic Computation," in *2014 IEEE International Electron Devices Meeting*, 2014, pp. 28.5.1-28.5.4.
- [54] Y. Lin *et al.*, "Transferable and Flexible Artificial Memristive Synapse Based on WO_x Schottky Junction on Arbitrary Substrates," *Advanced Electronic Materials*, vol. 4, no. 12, p. 1800373, 2018.

- [55] T. Chang, S.-H. Jo, and W. Lu, "Short-Term Memory to Long-Term Memory Transition in a Nanoscale Memristor," *ACS Nano*, vol. 5, no. 9, pp. 7669-7676, 2011/09/27 2011.
- [56] T. Chang, S.-H. Jo, K.-H. Kim, P. Sheridan, S. Gaba, and W. Lu, "Synaptic Behaviors and Modeling of A Metal Oxide Memristive Device," *Applied Physics A*, vol. 102, no. 4, pp. 857-863, 2011/03/01 2011.
- [57] S. Jabeen, M. Ismail, A. M. Rana, and E. Ahmed, "Impact of Work Function on the Resistive Switching Characteristics of M/ZnO/CeO₂/Pt Devices," *Materials Research Express*, vol. 4, no. 5, p. 056401, 2017/05/16 2017.
- [58] M. Kund *et al.*, "Conductive Bridging RAM (CBRAM): An Emerging Non-Volatile Memory Technology Scalable to Sub 20nm," in *IEEE International Electron Devices Meeting, 2005. IEDM Technical Digest.*, 2005, pp. 754-757.
- [59] U. Russo, D. Kamalanathan, D. Ielmini, A. L. Lacaita, and M. N. Kozicki, "Study of Multilevel Programming in Programmable Metallization Cell (PMC) Memory," *IEEE Transactions on Electron Devices*, vol. 56, no. 5, pp. 1040-1047, 2009.
- [60] K. Aratani *et al.*, "A Novel Resistance Memory with High Scalability and Nanosecond Switching," in *2007 IEEE International Electron Devices Meeting*, 2007, pp. 783-786.
- [61] S. Sills *et al.*, "A copper ReRAM cell for Storage Class Memory applications," in *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers*, 2014, pp. 1-2.
- [62] S. Yasuda *et al.*, "A Cross Point Cu-ReRAM with A Novel OTS Selector for Storage Class Memory Applications," in *2017 Symposium on VLSI Technology*, 2017, pp. T30-T31.
- [63] J. Guy *et al.*, "Investigation of the Physical Mechanisms Governing Data-Retention in Down to 10nm Nano-Trench Al₂O₃/CuTeGe Conductive Bridge RAM (CBRAM)," in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 30.2.1-30.2.4.
- [64] S. Fujii *et al.*, "Scaling the CBRAM Switching Layer Diameter to 30 nm Improves Cycling Endurance," *IEEE Electron Device Letters*, vol. 39, no. 1, pp. 23-26, 2018.
- [65] S. Z. Rahaman *et al.*, "Excellent Resistive Memory Characteristics and Switching Mechanism Using a Ti Nanolayer at the Cu/TaO_x Interface," (in eng), *Nanoscale research letters*, vol. 7, no. 1, pp. 345-345, 2012.

- [66] A. Belmonte *et al.*, "90nm W\Al₂O₃\TiW\Cu 1T1R CBRAM Cell Showing Low-Power, Fast and Disturb-Free Operation," in *2013 5th IEEE International Memory Workshop*, 2013, pp. 26-29.
- [67] S. Z. Rahaman *et al.*, "Impact of TaO_x Nanolayer at the GeSe_x/W Interface on Resistive Switching Memory Performance and Investigation of Cu Nanofilament," *Journal of Applied Physics*, vol. 111, no. 6, p. 063710, 2012.
- [68] L. Goux *et al.*, "Influence of the Cu-Te Composition and Microstructure on the Resistive Switching of Cu-Te/Al₂O₃/Si Cells," *Applied Physics Letters*, vol. 99, no. 5, p. 053502, 2011.
- [69] E. O. Neftci, B. U. Pedroni, S. Joshi, M. Al-Shedivat, and G. Cauwenberghs, "Stochastic Synapses Enable Efficient Brain-Inspired Learning Machines," *Frontiers in neuroscience*, vol. 10, p. 241, 2016.
- [70] J. H. Lee and K. K. Likharev, "Defect-Tolerant Nanoelectronic Pattern Classifiers," *International Journal of Circuit Theory and Applications*, vol. 35, no. 3, pp. 239-264, 2007.
- [71] M. Suri *et al.*, "Bio-Inspired Stochastic Computing Using Binary CBRAM Synapses," *IEEE Transactions on Electron Devices*, vol. 60, no. 7, pp. 2402-2409, 2013.
- [72] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale Memristor Device as Synapse in Neuromorphic Systems," *Nano Letters*, vol. 10, no. 4, pp. 1297-1301, 2010/04/14 2010.
- [73] X. Yan *et al.*, "Memristor with Ag-Cluster-Doped TiO₂ Films as Artificial Synapse for Neuroinspired Computing," *Advanced Functional Materials*, vol. 28, no. 1, p. 1705320, 2018.
- [74] T. D. Dongale, S. V. Mohite, A. A. Bagade, R. K. Kamat, and K. Y. Rajpure, "Bio-Mimicking the Synaptic Weights, Analog Memory, and Forgetting Effect Using Spray Deposited WO₃ Memristor Device," *Microelectronic Engineering*, vol. 183-184, pp. 12-18, 2017/11/05/ 2017.
- [75] J. H. Yoon *et al.*, "Truly Electroforming-Free and Low-Energy Memristors with Preconditioned Conductive Tunneling Paths," *Advanced Functional Materials*, vol. 27, no. 35, p. 1702010, 2017.
- [76] Y. Wang *et al.*, "Self-Doping Memristors with Equivalently Synaptic Ion Dynamics for Neuromorphic Computing," *ACS Applied Materials & Interfaces*, vol. 11, no. 27, pp. 24230-24240, 2019/07/10 2019.

- [77] A. C. Wang *et al.*, *Cross-point Resistive Memory: Nonideal Properties and Solutions* (no. 4). Association for Computing Machinery, 2019, p. Article 46.

CHAPTER 2

EXPERIMENTAL TECHNIQUES

This chapter presents the experimental techniques used to characterize all the devices and materials investigated. The entire device fabrication process is described in detail consisting of UV lithography and magnetron sputtering deposition steps. The material characterization methods performed, i.e., atomic force microscopy (AFM), and high-resolution transmission electron microscopy (HR-TEM), are also described. The material analysis was performed to ensure good device yield and obtain an insight into the physical mechanism involved during the device operation. Lastly, the electrical characterization techniques implemented on the corresponding device structures are thoroughly explained.

2.1. DEVICE FABRICATION

The devices investigated in this project were fabricated by using lithography patterning technique. Different device dimensions were achieved with different features designed on soda-lime glass mask. All layers of the thin film in the device structure were deposited by magnetron sputtering system. The details of the steps are given in the following.

Figure 2.1. SUSS MicroTec MJB4 mask aligner used for UV exposure and manual alignment of the bottom and top electrode patterning

2.1.1. Ultra-violet Lithography Patterning

Ultra-violet (UV) lithography technique was used to facilitate the device patterning from $50 \times 50 \mu\text{m}^2$ down to $5 \times 5 \mu\text{m}^2$ cross-point structure. The devices were fabricated on top of commercially available Si/SiO₂ substrate with oxide thickness of 300 nm ($\pm 5\%$).

Substrate Cleaning Treatment

The fabrication process of each sample was started by substrate cleaning treatment. The substrate was cleaned by first acetone and followed by isopropyl alcohol (IPA). The acetone is used to remove organic impurities, i.e., oily and/or greasy contaminants. Due to its rapid rate of evaporation it tends to redeposit the contaminants back onto the wafer surface. In order to prevent this from happening during the cleaning process, subsequent solvent IPA is used as a rinse agent to remove the contaminated acetone. The substrate was soaked and put into ultrasonic cleaner for 15 minutes for each solvent. The substrate was then dried up using Nitrogen (N₂) gun and baked at 130 °C to completely remove the solvents.

UV Lithography Process

For the lithography process, AZ5214E resist was used in the positive tone mode. The resist was spin-coated on the substrate with 6000 RPM for 1 min, resulting in $\sim 1.3 \mu\text{m}$ thick resist. It was then soft baked at 110°C on a hotplate for 2 minutes to get rid of the solvents and solidify the resist film. The UV exposure was done under soda lime glass mask with the bottom electrode (BE) pattern. The exposed sample was then developed using dissolved AZ400K developer (1:4 ratio with DI water) for ~ 40 s, rinsed by DI water, and dried using N₂ gun. The first lithography process was followed by the BE deposition. After the BE deposition, the sample underwent lift-off process. The lift-off step was first done by using acetone followed by IPA, i.e., the same steps used for substrate cleaning. This was done to first remove most of the metals from the sputtering deposition. It was then followed by soaking the sample inside AZ100 remover with heated ultrasonic bath for ~ 60 minutes before the cleaning process using acetone and IPA was repeated. This is to ensure there is no resist residue especially at the edges of the structures. In the second lithography process, the same recipe was also implemented with the additional manual alignment before the UV exposure, i.e., to align the thinnest portion of the cross ensuring the controllable device dimension. The summary of the device fabrication process can be seen on **Figure 2.2**.

Figure 2.2. Summary of the entire fabrication process of the cross-point devices.

2.1.2. Thin Film Deposition Techniques

All the thin film structures investigated in this project were deposited by magnetron sputtering deposition. Two different sputtering techniques were implemented for different materials. All the materials utilized in the structures are metals, i.e., Ti, Pt, and Cu, except HfO_x . The metal thin films were deposited using DC sputtering approach from pure metal targets, while HfO_x was deposited via RF sputtering deposition with the use of stoichiometric HfO_2 target.

Figure 2.3. Ultra-high vacuum of multi-cathode magnetron sputtering deposition system. All of the thin film depositions were performed inside one chamber.

RF sputtering was required for the HfO_x deposition due to the low conductivity of the materials. This is critical to avoid positive charge build-up on the surface of the cathode due to the bombardment of the ions during the deposition. Furthermore, due to the low thermal conductivity of the HfO_x and relatively small sputtering target diameter of 2", the power supply needs to be ramped up slowly to reach the targeted power and then ramped back down slowly to zero. Thus, to initiate the plasma, i.e., usually known as "strike", low power with high pressure is implemented rather than starting at high power directly. This is a precaution required to ensure healthy target condition minimizing the risk of thermal shock due to the thermal gradient between the HfO_x target and the Cu backing plate. High temperature during the deposition process due to too high power might also damage the indium bonding present between the HfO_x and the Cu backing plate. These steps are highly dependent on the size of the sputtering target and the cooling system, i.e., direct or indirect cooling, is implemented in the system.

The sputtering system is equipped with ultra-high vacuum chamber capable of achieving $\sim 10^{-8}$ Torr base pressure with 7 con-focally oriented cathodes and 1 cathode positioned right below the substrate. Because the structures fabricated were not planarized, all the materials were deposited with the con-focally oriented cathodes to ensure similar conformality throughout all the depositions. Each material was pre-calibrated by using a patterned wafer and lift-off process to facilitate step measurement using atomic force microscope (AFM). The materials were deposited with constant 20 sccm Ar concentration under different deposition pressure and power. The metal thin films were deposited with 50 W DC power and 2 mTorr deposition pressure with the plasma usually initiated at 50 W and 20 mTorr. Different metals have different deposition rate under the same power and pressure condition. On the other hand, the HfO_x was deposited at 50 W RF power and 1.5 mTorr deposition pressure. The strike was performed at 30 W and 50 mTorr, and then the power was gradually increased to 50 W with the rate of ~ 67 mW/s and gradually decreased to zero with rate of ~ 167 mW/s.

Figure 2.4. Confocal configuration of 8-gun sputtering deposition system.

Before sputtering of the first thin film layer, the substrate always underwent Ar plasma cleaning to further clean the surface from contamination thus prepare the surface for better adhesion before the deposition. The order in which the thin films were deposited must also be taken into account due to the potential cross contamination from different material depositions within the same chamber. When any of the metals was deposited right after the HfO_x deposition, target surface plasma cleaning was performed for 5 mins before the actual deposition was started, especially for the highly reactive Ti.

2.2. MATERIAL CHARACTERIZATIONS

The material characterization of the devices consists of atomic force microscopy (AFM) and high-resolution transmission electron microscopy (HR-TEM)

2.2.1. Atomic Force Microscopy

AFM was mainly used to check the thickness of the thin film layers and the present of random spikes due to lift-off process issue or sidewalls formed during the confocal sputtering process. The thickness calibration of the sputtering process was done by measuring the step-height of the materials grown on the SiO₂ substrate. All the imaging was done in non-contact AFM mode. This mode utilizes the attractive inter-atomic force between the AFM tip and the sample surface while the tip is oscillating above the surface. The image is generated by the feedback mechanism to maintain the amplitude of the tip oscillation. As the tip gets closer to the sample surface, the amplitude of the oscillation decreases and vice versa. This mode is preferred to prevent sample damage during the scanning and prolong the lifespan of the tip while maintaining high image quality.

Figure 2.5. The cross-point device under optical microscope (a) and AFM topography images of two different devices (b, c). The presence of significantly high random unwanted spikes can be observed in (b).

The presence of the spikes or sidewalls can severely impact the performance of the devices. As an example, the presence of the high spikes or sidewalls after the BE lift-off might lead to thinner oxide effective thickness formed between BE and TE, resulting in the inconsistent switching voltages observed. After the lift-off process performed on the BE patterned sample, the topography of the μm-wide stripes was scanned to ensure no spikes or sidewalls present. The topography check was also performed on the completed cross-point devices, as depicted in **Figure 2.5**.

2.2.2. High Resolution Transmission Electron Microscopy

HR-TEM characterization was performed on two devices of different structures, i.e., Pt/HfO_x/Ti and Pt/HfO_x/Cu. The 10 × 10 μm² cross-point devices were prepared by UV lithography patterning and sputtering deposition process. The IV characteristic of the pristine devices were checked with the DC IV sweep below the forming voltage to ensure all the devices were in the lowest conductance state (pristine insulating state). The devices were then cut using focus ion beam (FIB) with Pt protection layer deposited on top of the TE, as seen on the **Figure 2.6**. The cross-sectional HR-TEM view of the device at the BE edge was used to investigate the coverage of the deposition with respect to the area near the centre of the cross-point device. This was performed on the region of the Pt/HfO_x/Cu cross-point device, as depicted on the **Figure 2.7**. The thickness of the HfO_x in region 1 and 2 are close to 10 nm. Thus, accurate analysis can be done on the device performance associated with the oxide layer thickness.

Figure 2.6. The deposition of Pt protective layer before the FIB process.

Figure 2.7. The schematic of the cross-point device structure and the corresponding cross-sectional HR-TEM images

On the other hand, the TEM analysis was performed on the crystal structure of the thin film layers of each material. The metal thin films, i.e., Pt, Ti, and Cu, used in these structures are found to be in the polycrystalline phase, while the oxide layer, i.e., HfO_x , is in amorphous state. The different phases of the materials were analysed through fast-Fourier transform (FFT) and inverse fast-Fourier transform (IFFT) of the selected region on the HR-TEM images.

Magnified FFT

FFT

IFFT

Figure 2.8. The FFT and IFFT of selected region on the cross-sectional images for the corresponding thin film materials.

2.3. ELECTRICAL CHARACTERIZATIONS

The electrical characterization of the devices can be divided into two categories, i.e., DC IV and pulsed IV measurement. The combination of these two enables the investigation of the different aspect of the devices. These measurements were performed on cascade microtech manual probe station using Keithley 4200A SCS equipped with 1 pulse measure unit (PMU) and 2 source measure units (SMUs) configuration.

Figure 2.9. Cascade microtech manual probe station with integrated Keithley 4200A SCS system

2.3.1. DC IV Measurement

The DC IV measurement is a simple two-terminal current-voltage (IV) sweep in which the external voltage is implemented across the top electrode (TE) and bottom electrode (BE) of the device while measuring the current response of the device. This measurement was performed by using two source-measure units (SMUs) in which one SMU acts as a ground unit applying 0 V to the BE and the other SMU sweep the voltage implemented at the TE while measuring the current response of the device.

In the pristine memristive devices, DC IV sweep is usually used to find the forming voltage of the device. It is especially important when the memristive device is measured as a standalone device or 1R architecture because the SMU is usually equipped with compliance current capability, which is needed to reduce the risk of permanent device degradation during forming and SET operation. While it has a capability to limit the current flowing through the device in the long run, it might not be able to help the current overshoot that occurs in the ultra-fast regime (<10 ns) due to the inherent parasitic capacitance present in the connection. This issue can be seen in the first RESET current of the device, which is usually higher than the CC

implemented during the electroforming process. However, this issue is common in 1R architecture and can only be completely mitigated with the use of a series transistor integrated with the memristive device.

The DC IV characteristic holds substantial information related to the memristive switching dynamics. The commonly observed non-linear response of the memristive devices is critical to understand the physical mechanism involved during the device operation as well as improving the device performance, e.g., finding the optimum value of read voltage to achieve the largest dynamic ratio possible. Thus, despite not mimicking the actual on-chip device operation, DC IV characteristic of the device can provide a critical insight on how the device will perform under the actual pulse programming operation.

2.1.3. Pulsed IV Measurement

Pulsed IV measurement is required to perform characterization of the time-domain response of the devices. It is able to simultaneously generate voltage pulse and measure the current response of the device down to below 80 ns depending on the current range required to measure the current flowing through the device under test (DUT). In order to obtain the full waveform of the pulse generated and the current response of the device, the PMU needs to utilize two timely synchronized channels to separately handle the pulse generation and the current reading. This is to avoid the charging and discharging effect from the channel generating the voltage pulse. However, if the measurement only requires several data point to take an average from, 1 PMU channel can be sufficient to perform an accurate measurement with proper waiting time between the start of the voltage pulse to the end of the current charging/discharging regime.

Unlike SMU, PMU is not equipped with current limiting capability. Thus, when the PMU is utilized to measure the latency of a memristive device, an external resistor is needed to act as a current limiter, unless the DUT exhibits some sort of self-compliance capability. The use of series resistor was implemented during the characterization of cation based diffusive memristor to prevent the excessive migration of metal species into the dielectric layer, while no resistor was connected to the anion-based memristive device. Furthermore, the pulsed IV measurement is usually used to facilitate the endurance measurement, enabling more efficient and faster cycling.

CHAPTER 3

Switching Dynamics of Pt/HfO_x/Ti Anion-Based Memristive Devices

In this chapter, different electroforming treatments, i.e., low compliance forming (LCF) and self-compliance forming (SCF) on Pt/HfO_x/Ti/Pt devices are presented along with the corresponding IV characteristics and multilevel conductance states capability of the devices. The conduction mechanisms involved during the device operation can be associated with the bulk-like interfacial mechanism under the theoretical framework of the space-charged limit conduction (SCL). Three major contributors to the overall device conduction are ohmic, trap-unfill space-charged limited (TU-SCL), and trap-fill space-charged limited (TF-SCL) conduction. Each of this mechanism was found to be more dominant than the others in different voltage regimes. The abrupt SET behaviour of the devices is potentially driven by filamentary gap switching, while the progressive SET is governed by interfacial effect. Based on the extracted parameters of the voltage trap fill limit and trap characteristic temperature, the progressive SET behaviour of the devices was attributed to the trap-controlled SCL switching mechanism in which the switching operation is facilitated by the generation/redistribution of new/existing oxygen vacancy defects at the HfO_x/Ti interface. These oxygen vacancies act as electron traps resulting in transition of among different charge trapping levels during the switching process. Transition between these two potential switching mechanisms were observed on the SCF devices through the experimentally fitted conduction model on the device IV characteristic under different compliance current level.

3.1. DIFFERENT ELECTROFORMING TREATMENTS AND THE CORRESPONDING IV CHARACTERISTICS

HfO_x anion-based memristive devices have been widely investigated since the emergence of memristive devices in the field of non-volatile memory technology. Its simple structure and sub-10 nm scalability attracted many researchers in the field to extensively investigate its characteristics and to further improve its key performance parameters [1, 2]. It has been well-understood that the reactive electrode, e.g., Ti [3-6], TiN [7-9], Hf [1, 10, 11], etc, acts as an oxygen reservoir electrode to facilitate the electric-field induced oxygen vacancy defects generation and recombination during the forming or SET process, i.e., low to high conductance state (LCS to HCS) transition, and the RESET process, i.e., HCS to LCS, respectively.

In HfO_x/Ti system, the presence of TiO_x interfacial layer has been reported due to high oxygen reactivity of the Ti electrode in the pristine device state or after the forming process. The thickness of the TiO_x layer in the pristine device is highly dependent on the overall thickness of the Ti electrode and the stoichiometry as well as morphology of the HfO_x layer [3, 5, 12]. The presence of this TiO_x layer enables the self-compliance characteristics on the structure during the forming and SET operation. Self-compliance characteristic is the intrinsic ability of the device to limit the current flowing through it without external entity such as a series transistor. This property is highly desirable to enable transistor-less memristive array integration.

Electroforming process for the pristine filamentary-based memristive devices has been reported as the critical stage in determining the subsequent memristive device performance [13-18]. It can be used to alter the conductive filament characteristics responsible for the conductance switching behaviour. In general, electroforming and the first RESET operating parameters, i.e., voltage and current, are higher than the operating parameters during subsequent SET and RESET cycles. In cation-based devices, the forming treatment determine the stability and the dimension of the metallic filament involved during the switching. It might lead to either volatile or non-volatile switching characteristics depending on the stability of the filament. In the anion-based devices, the forming step sets the number of oxygen vacancy defects generated within the oxide switching layer, which will determine the conductance levels of the device in the subsequent switching operations. In general, compliance current (CC) is required to limit the current flowing through the device during the electroforming as

well as SET process to prevent excessive formation of defects leading to permanent device degradation, unless the device holds certain degree of self-compliance characteristic.

In order to fully utilize the scalability of memristive devices in the high-density crossbar array implementation, the use of a series transistor for each memristive device must be eliminated. The series transistor plays two critical roles in the 1T1R integration, i.e., as a select device and a current limiting element. Thus, not only a two-terminal select device (selector) is required to replace the transistor, but also a self-compliance feature originated from either the selector or memristive element. Thus, self-compliance feature is highly beneficial for the high-density array application.

Figure 3.1. The IV characteristics of first 70 cycles of LCF devices with forming CC of $200 \mu\text{A}$ (a). The subsequent cycles were performed with programmed voltage of 0 to 1.5 V for SET and 0 to -1.5 for RESET. The forming of SCF devices were done without limiting the current flowing through the devices (b, c). The SCF devices were operated under two different mode, i.e., self-compliant SET (SCF1) and CC- $200 \mu\text{A}$ SET (SCF2).

The $10 \times 10 \mu\text{m}^2$ cross-point devices consist of 10 nm HfO_x layer sandwiched between 10 nm Pt inert electrode and 50 nm Ti oxygen reservoir electrode was fabricated using UV-lithography patterning and sputtering deposition technique. The pristine devices underwent two different electroforming treatments, i.e., low-current forming (LCF) and self-compliance forming (SCF). For simplicity, the devices will be addressed as LCF and SCF devices respectively. The LCF treatment was performed by programming the DC voltage sweep from 0 to 3 V and limiting the current flowing through the device at 200 μA . On the other hand, the SCF treatment was performed without implementing compliance current during the forming, making use of TiO_x interfacial layer formed within the pristine devices as well as during the forming sweep. The IV characteristics of the devices are depicted in **Figure 3.1**. The forming occurred at around 2 V to 2.5 V. The first RESET process started at current level of 2-4 mA for LCF and SCF devices. The subsequent SET cycles were performed with CC of 200 μA for LCF devices, while SCF devices underwent two different DC IV sweep, i.e., self-compliant mode (SCF1) and with SET CC of 200 μA (SCF2). LCF and SCF2 devices exhibit abrupt SET and progressive RESET process under the DC sweep. On the other hand, SCF1 devices showed abrupt SET and a mixed of abrupt and progressive RESET. The three different IV behaviours obtained under different operating modes were used to optimize the operating parameters during the pulse operation. The IV characteristic of SCF2 device suggests that despite having compliance free forming treatment, the subsequent SET and RESET operating parameters as well as conductance levels highly depend on the current flowing through the device during the SET process, which means the SET pulse amplitude and duration can be used to tune the device property in the absence of current limiting element. Furthermore, these IV characteristics were used to analyse the underlying conduction and switching mechanisms involved during the device operation.

The progressive RESET process has been widely used to achieve multilevel conductance states within a single cell. In the DC sweep, the different conductance states can be obtained by implementing different stopping voltages (SVs). While in the pulse programming operation, the different conductance states can be achieved through the implementation of different pulse amplitudes, durations, and/or numbers. The seemingly abrupt SET process during DC IV sweep can also be used to achieve multilevel states via the use of different compliance current values. This approach can be mimicked in the pulse programming operation with the use of a series transistor with the memristive device, i.e., 1T1R architecture, controlling the current through the device via the transistor gate biasing. In 1R

devices, a more rigorous optimization of pulse amplitude and duration during the programming operation can also be utilized to achieve the multilevel conductance states capability of the structure, which was the main approach used throughout the experiment in this work.

Figure 3.2. Comparison of the distribution of the switching voltages (a) and conductance values (b) for LCF and SCF devices. SCF devices successfully achieved higher ON/OFF ratio with lower operating voltages.

For the LCF devices the SET processes occurred at ~ 1.1 V and the progressive RESET started at ~ -0.8 V, while the SCF devices exhibited lower switching voltages of ~ 0.7 V and ~ -0.6 V for SET and RESET respectively. The switching voltage distributions can be seen on the **Figure 3.2(a)**. Due to the progressive RESET behaviour observed on the devices, V_{RESET} is defined as the voltage value in which the progressive RESET process initially started, while V_{SET} is defined as the first voltage value corresponding to the CC level. In the first 70 cycles (after forming and first RESET) of the DC IV sweep, these operating voltages resulted in the conductance values depicted in **Figure 3.2(b)**, at the read voltage of 0.2 V. The LCF devices could achieve dynamic ratio of ~ 5 , while SCF treatment improved the ratio significantly to ~ 20 . The LCF devices exhibited lower conductance values for LCS and HCS as compared to SCF devices. These can be attributed to different nature of conductive filament nucleated during the electroforming process, which will be further discussed later.

3.2. MULTILEVEL CONDUCTANCE STATES

Multilevel conductance capability within a single memristive device has attracted enormous attention in the field of high-density storage and neuromorphic engineering. The progressive RESET behaviour that is commonly observed in memristive devices is usually the main property exploited to store different memory states. Due to the abrupt nature of the SET

process, the device is usually programmed starting from the HCS level and slowly moved towards the lower conductance levels. When the device needs to be written into higher conductance level than the starting conductance, SET process is required to bring the device all the way to HCS before starting the progressive RESET operation. This will require more energy and slow down the programming process significantly, given the inherent cycle-to-cycle and device-to-device variations presence during the memristive device operation. While this property can still be tolerated for high density memory, this has been the major issue in the implementation of large memristive-based synaptic array for neuromorphic computing application. Thus, large amount of efforts has been put into achieving progressive SET and RESET behaviour at the device level.

The extensive characterization of the multilevel conductance states capability of the Pt/HfO_x/Ti/Pt devices is divided into conductance switching behaviours under DC sweep, different pulse amplitudes and durations during progressive RESET operation, and optimized non-identical pulse programming during SET and RESET operation. Based on the previous DC characteristics, showing significantly higher dynamic ratio and lower switching operation, SCF devices were used to investigate the multilevel conductance state capability of the devices and to optimize the pulse parameters. Eventually, the optimized parameters were also implemented on the LCF devices.

DC-IV Sweep

Anion-based memristive devices have been known to exhibit abrupt SET and progressive RESET behaviour. Due to the compliance free forming capability and the higher dynamic ratio observed in **Figure 3.2(b)**, SCF2 devices were used to demonstrate the multilevel conductance states capability of the structure under DC sweep. Different levels of conductance were obtained by first applying different CC values during the SET process to control the conductive filament formation. Starting at LCS, the CC value was increased from 50 μ A to 200 μ A with interval of 25 μ A, obtaining 7 other conductance states in the process. On the other hand, the RESET process was controlled by varying the stopping voltage (SV). Starting from the HCS level, the SV value was varied from -0.6 V to -1.5 V with the interval of 0.1 V achieving 10 different conductance levels in the process, as depicted on **Figure 3.3(a)**. In order to check the non-volatility nature and the stability of the conductance states, each conductance level was measured under 0.2 V read voltage, the reading was done with interval

of 20s for overall 1000s duration. Relatively stable conductance levels were observed, as shown on **Figure 3.3(b), (c), and (d)**.

Figure 3.3. Multilevel conductance states capability of SCF devices. The multilevel conductance values were achieved by applying different CCs during SET and different stopping voltages during RESET. The stability of each state was measured under 0.2 V read voltage for 1000s.

Progressive RESET behaviour under different pulse amplitudes and durations

The conductance modulation behaviour of the SCF devices was further characterized by using different pulse programming amplitudes and durations. Based on the progressive switching behaviour observed on the IV characteristics during the RESET process, short programming pulses with fixed 40 ns rise/fall times and hundreds of ns width were only implemented during the RESET operation, while the SET operation was performed under DC source generated pulse with amplitude of 2 V, duration of 5 ms, and CC of 200 μA . This was performed to minimize the cycle-to-cycle variation of the devices generated during the SET process, thus isolating the device variation during the RESET process. With the well-controlled SET process, the same starting conductance level can also be ensured before the RESET pulses

were sent through the device. Different RESET pulse amplitudes varying from -0.9V to -1.5V with increment of 0.1 V under fixed pulse width of 500 ns were implemented. Two different RESET schemes were investigated in the following.

Scheme 1

Every RESET pulse was followed by SET operation to bring the conductance back to the starting point at HCS. The conductance states under different RESET voltage amplitudes are depicted on the **Figure 3.4**. The set of HCS values can be seen clustered together indicating excellent distribution from the implementation of pulse mode DC source for the SET biasing. Strong correlation between the amplitude of the pulse and the change of conductance values were observed. The average of each conductance value decays exponentially with the linear increase in the RESET voltage amplitude. This can be explained by the increasing gap between the oxygen vacancy defects filament and the inert electrode. As the gap grows, the electric field across the active switching region reduces, thus less oxygen-vacancy recombination can take place.

Figure 3.4. The conductance distribution of the SCF device under fixed RESET pulse width of 500 ns and varying pulse amplitudes (-0.9 V to -1.6 V). The RESET processes were started from HCS value for each cycle.

The pulse width dependence of the conductance modulation was also characterized under the same scheme. Pulse amplitude of -1.5V and -1.3V were implemented with different pulse width from 100 ns to 800 ns. Similar correlation was observed as compared to the pulse amplitude-based modulation, as shown on the **Figure 3.5**. However, significantly broader conductance values distribution was observed indicating more probabilistic nature of the switching with the decrease in pulse duration. Lower set of conductance levels were observed

under -1.5 V as compared to -1.3 V pulse amplitude. This is consistent with the previous observation associated with the change in the value of effective electric field across the active switching region.

Figure 3.5. The conductance distribution of the SCF device under fixed RESET pulse amplitude of -1.5 V and -1.3 V with increasing pulse duration from 100 ns to 800 ns. The conductance value of each amplitude tends to get saturated even with further increase in pulse width. This is potentially due to electric field limited switching nature of the conductive filament.

Scheme 2

The second conductance update scheme investigated on the devices was performed by sending back-to-back RESET pulses without performing SET operation after each RESET cycle. The same set of non-identical RESET pulse amplitudes and durations were used. Lower set of conductance values were obtained because each RESET cycle was done from the lower conductance values rather than starting from HCS. This also resulted in narrower conductance distribution of each corresponding state. This indicates that the conductance update behaviour of the structure depends on not only the pulse amplitude and duration, but also the initial conductance level. The conductance distributions achieved under different amplitudes and pulse widths can be seen on the **Figure 3.6**.

Figure 3.6. The conductance distribution of the SCF devices under back-to-back RESET pulse with 200 ns and 500 ns duration and changing pulse amplitudes from -0.9 V to -1.6 V. The distributions of the conductance were shifted towards the lower conductance level

The exponential relation between the RESET pulse amplitude and the obtained average conductance values agrees with the well-established conductive filament-gap model. Based on this model, during the dissolution of the conductive filament, a tunnelling gap between the oxygen vacancy filament and the electrode is formed. The thickness of this gap can be associated with the tunnelling current flowing through the gap under the Wentzel-Kramers-Brillouin approximation. By only considering the elastic tunnelling process and assuming trapezoidal barrier with linear electrical potential, the transmission probability of the tunnelling electron can be analytically expressed as [9]:

$$T = \exp \left[-\frac{4\sqrt{2m^*}}{3\hbar qF} \left(E_t^{\frac{3}{2}} - (E_t - qFd)^{\frac{3}{2}} \right) \right], \quad (1)$$

with q , F , m^* , d , and E_t are electronic charge quantity, external electric field applied, electron effective mass, distance between the closest traps to the electrode (tunnelling gap), and trap energy below conduction band respectively. Thus, with the exponentially decreasing trend of conductance observed with linearly increasing pulse amplitude, it agrees with the potentially linear correlation between the RESET pulse amplitude and the resulting tunnelling gap during the RESET process [19].

Optimized SET pulse with amplitude of 1.1 V and increasing RESET pulse amplitudes from -0.65 V to -1.5 V with finer interval of 25 mV and fixed duration of 200 ns were implemented to achieve the gradual RESET behaviour. The SET pulse was optimized to attain the SCF2 HCS rather than SCF1. This was done to achieve the progressive RESET rather than the abrupt RESET feature of the devices based-on the obtained IV characteristics. Different conductance levels were achieved on the SCF devices, as shown on the **Figure 3.7(a)**. The average of HCS values obtained was higher than the one achieved by the DC bias with 200 μ A CC. This indicates different conductive filament morphology potentially formed due to the absence of the compliance current. Despite the absence of the CC, the SCF devices were able to maintain the progressive RESET characteristic.

Figure 3.7. The conductance distribution of the SCF devices (a) under single SET voltage of 1.10 V, 200 ns and back-to-back RESET pulse with 200 ns duration and varying pulse amplitudes from -0.65 V to -1.50 V. The device average conductance response with $\pm\sigma$ as a function of pulse number (b)

Progressive SET behaviour under optimized pulse parameters

So far, only the progressive RESET behaviour has been discussed due to the seemingly abrupt switching nature of the devices in the SET regime. Based on the optimized 1.10 V, 200 ns SET pulse, a gradually increasing pulse amplitude from 0.76 V to 1.1 V with 10 mV interval was implemented, while increasing RESET pulse amplitudes from -0.65 V to -1.5 V with interval of 25 mV were used with fixed pulse duration of 200 ns.

Progressive SET and RESET behaviour were achieved in both LCF and SCF devices, as depicted in **Figure 3.8**. In LCF devices, higher conductance levels were achieved as compared to the ones obtained during DC sweep. This can be associated with higher current flowing through the device during pulse SET operation, thus altering the conductive filament dimension formed during the SET process. Despite that, uniform conductance distributions were achieved during SET and RESET process. More deterministic behaviour was observed during the RESET process as compared to the SET process. On the other hand, while achieving progressive RESET operation, the switching behaviour of the SCF devices during the SET process exhibited more stochastic response. This behaviour of SCF devices agreed with the generally reported filamentary anion-based memristive devices in which the SET operation of the devices tend to be more abrupt in nature due to the decreasing gap between the oxygen vacancy defects filament to the electrode that results in the positive feedback mechanism increasing the effective electric field strength within the active switching region.

Based on the characteristics of these devices, the device operated under LCF mode is an excellent candidate for not only analog synaptic device, but also multi-level cell. As an analog synaptic device with 1T1R architecture and properly designed peripheral circuitry, the device can facilitate weight update in both direction while maintaining large dynamic ratio of ~25. The highly uniform conductance distribution achieved during the RESET process demonstrates a very promising multi-level cell (MLC) property. With the help of transistor working in tandem with the device, the uniformity of these conductance states can be further improved.

The excellent progressive SET process demonstrated by the LCF devices were not expected due to the seemingly abrupt switching observed under DC IV sweep. Thus, further study on the conduction and switching mechanisms of the LCF and SCF devices were conducted to understand the physics behind the device characteristics observed.

Figure 3.8. Conductance distribution of LCF devices during SET (a) and RESET (b). The average conductance as a function of pulse number for LCF (c) and SCF (d) devices.

3.3. THE CONDUCTION AND SWITCHING MECHANISMS UNDER LCF AND SCF TREATMENTS

Numerous conduction mechanisms have been reported in metal-insulator-metal (MIM) structures [20]. In general, they can be classified into two major categories, i.e., electrode-limited and bulk-limited mechanism. The electrode-limited conduction is highly dependent on parameters associated with the electrode-insulator interface, e.g, interfacial energy barrier height and dominant charge carriers responsible for the transport. It consists of direct tunnelling, Fowler-Nordheim (FN) tunnelling, Schottky emission, and thermionic-field emission. On the other hand, the bulk-limited mechanism, as its name suggested, relies on the intrinsic properties of the insulating material, such as carrier mobility within the insulator as well as the dielectric trap energy, density, and distribution. Among the bulk-limited mechanisms are Poole-Frenkel (PF) emission, ionic conduction, hopping conduction, grain-

boundary-limited conduction, ohmic conduction, and space-charge limited (SCLC) conduction.

The significant difference in work function of the electrodes, i.e., $\phi_{Pt} = 5.12 - 5.93$ eV and $\phi_{Ti} = 4.33$ eV, and the electron affinity of the HfO_x layer of more than 2.99 eV, should reduce the barrier height at the HfO_x/Ti interface [5]. Theoretically, this will promote noticeably different magnitude of current flow under different bias polarity. However, the IV characteristics of the devices observed showed symmetrical behaviour before the conductance switching occurs. This indicates that the contribution of the electrode-limited conduction current through these devices can be considered negligible. With that being said, the dominant conduction mechanisms involved during device operation are among those of the bulk-limited nature rather than electrode-limited ones.

Linear behaviour of the double logarithmic plot of the cell current and voltage, i.e., $\log(I)-\log(V)$, was observed for both LCF and SCF devices. Up to three different slope regions were obtained within a single sweep. The switching mechanism of HfO_x anion memristive devices have been widely accepted based on the modulation of the oxygen vacancy defects profile. These defect sites can serve as charge carrier traps within the oxide switching layer. Thus, in some reported devices, these properties result in trap-controlled SCLC mechanism [6-8, 13-18, 21, 22]. Similar conduction mechanism has been reported on the $Pt/HfO_x/Ti$ devices with atomic layer deposition (ALD) grown stoichiometric and non-stoichiometric HfO_2 layer [5]. The SCLC mechanism is highly dependent on the distribution of the traps within the band gap region. The widely implemented distribution consists of single discrete [23], exponential [24], and Gaussian traps distribution [25]. In a Gaussian trap distribution, the JV characteristics is modelled based on the presence of shallow and deep Gaussian traps. However, it has been shown that this distribution can be approximated with the equations from exponentially distributed trap density of states and a single discrete trap level. Thus, the JV characteristics of the trap-controlled SCLC can be simply estimated by the single discrete trap level model in low voltages (low occupancies of trap) and exponential trap distribution model in high voltages (high trap occupancies) [25]. It must be noted that these approximations are only implemented in the non-linear IV regime.

In the lower voltage regime, ohmic conduction driven by the thermally activated free electrons within the oxide layer contributes the most to the overall current flow, thus linear correlation of the IV can be modelled by the following equation [20],

$$J_{Ohm} = \frac{q\mu_N N_0}{L} V, \quad (2)$$

in which μ_N and N_0 represent the free electrons mobility and density inside the oxide switching layer, while L is the thickness of the oxide layer. As the voltage increases, the injected electrons from the ohmic contacts at the electrode-oxide interfaces started to exceed the limit, as the duration of the carrier transit time gets shorter and eventually surpasses the Ohmic relaxation time at a specific threshold voltage. Beyond this threshold voltage, the current is driven by electrode injected charge carriers that transit in the oxide traps rather than the thermally activated intrinsic electrons, entering the quadratic portion known as trap-unfilled SCLC (TUSCLC) regime. The threshold voltage in which the transition between Ohmic and TUSCLC conduction occurs is usually labelled as V_{ON} . The current density flowing through the device in this voltage range can be expressed as the following [23, 26],

$$J_{TUSCLC} = \frac{9\varepsilon_0\varepsilon_R\mu_N\theta}{8L^3} V^2, \quad (3)$$

ε_R represents the dielectric constant of the oxide, while θ depends on the ratio of the free charge carriers to total carriers density, i.e., free (n_F) and trapped (n_T) carriers, $\theta = \frac{n_F}{n_T + n_F}$.

This expression is essentially similar to the one that has been used to model the dielectric conduction in the absence of traps, i.e., Mott-Gurney law of conduction, scaled by a factor of θ . The physical significance of this scaling factor is often described as the decrease in either the effective carriers' mobility or dielectric constant.

From the IV characteristics of the devices, it was observed that further increase in the voltage resulted in steeper $\log(I)$ - $\log(V)$ slope, corresponding to higher voltage exponent. This is potentially caused by the traps filling below the quasi-Fermi energy level, which leads to the overall conduction highly depend on the trap density and energy distribution. If exponential traps distribution is assumed in the band gap, the voltage exponent dependence of higher than 2 can be attributed to trap-filled SCLC (TFSCCLC). Its J-V characteristics can be expressed as,

$$J_{TFSLC} = \frac{q^{1-\tau} \mu_N n_c}{L^{2\tau+1}} \left(\frac{2\tau+1}{\tau+1} \right)^{\tau+1} \left(\frac{\tau}{\tau+1} \frac{\epsilon_0 \epsilon_R}{n_t} \right)^{\tau} V^{\tau+1}. \quad (4)$$

τ is the ratio between the characteristic temperature, T_C , which is related to the trap distribution function, to the actual device temperature, T . n_c and n_t are the conduction band density of states and total density of states respectively. The transition between TUSCL and TFSLC conduction is indicated by another threshold voltage known as the trap-fill limit voltage, V_{TFL} .

The LCS and the HCS of the devices were fitted with these three conduction models. The V_{ON} and V_{TFL} values of the states were extracted by extrapolating the model fitting line of each region and intersecting the extrapolated lines between the adjacent voltage regimes. Fairly symmetrical voltage exponents and threshold voltages extracted from the experimental data for SET and RESET process emphasizes the bulk-limited nature of the conduction processes. For LCF devices, all three conduction regimes were found in both LCS and HCS. The extracted voltage exponent of each region and the threshold voltages, i.e., V_{ON} and V_{TFL} , are depicted on the **Figure 3.9**.

Figure 3.9. The conduction mechanisms exhibited by LCF devices and LCS and HCS were studied. The typical IV characteristics of the devices fit well with the three bulk-limited conduction mechanisms, i.e., Ohmic (red), SCL (green), and TF-SCL (blue).

The first key observation from LCF devices is the read voltage, i.e., $V_{READ} = 0.2 \text{ V}$, that was significantly higher than the V_{ON} in LCS and positioned in the transition between ohmic and TUSCLC region in HCS. Thus, the reading operation was done within the nonlinear IV region. It can be qualitatively observed from the graph that the dynamic ratio increases as the voltage reduced into the ohmic regime. By extracting the value of the conductance at 0.1 V, the dynamic ratio can be improved from ~ 5 to ~ 7.5 , which was an improvement by 50% from the original value, as depicted on **Figure 3.10**. This is critical in achieving sufficient reading margin during the device operation. However, during the pulse operation, 0.2 V was maintained throughout the experiment to enable higher read current range used for the analyser, thus the settling time during the measurement can be minimized. The dynamic ratio of the LCF devices were also improved due to higher current flowing through the devices during SET process, resulting in higher HCS values.

Figure 3.10. The conductance values distribution obtained different reading performed under different read voltage of 0.2V (green) and 0.1V (orange). Read voltage of 0.1V falls within ohmic region resulting in improved dynamic ratio.

The second key observation is the value of the V_{TFL} of $\sim 0.97 \text{ V}$ during the SET and V_{TFL} of ~ 0.91 during RESET process. The difference in the extracted V_{TFL} values potentially indicate the influence of the different electronic state at the Pt/HfO_x and HfO_x/Ti interfaces that results in asymmetrical trap potential property in the presence or the absence of the trapped carriers. As discussed earlier, with the significantly different electrode work functions and the electron affinity of the HfO_x, it is highly likely that the Schottky-like barrier with the depletion layer is formed at the interface, but it does not influence the bulk-limited SCLC behaviour.

Moreover, the trap-fill SCLC region could barely be observed before the switching occurs in the SET operation. This suggests the high possibility of trap-controlled SCL switching mechanism in which the conductance modulation is driven by the oxygen vacancies modulation at the HfO_x/Ti interface, resulting in transition between trap-unfilled to trap-filled SCLC regime, instead of conductive filament gap. Similar IV characteristic and switching dynamics have been reported in $\text{Ag}/\text{La}_{0.7}\text{Ca}_{0.3}\text{MnO}_3/\text{Pt}$ and $\text{Ag}/\text{Pr}_{0.7}\text{Ca}_{0.3}\text{MnO}_3/\text{Pt}$ heterostructures [22, 26]. However, with the implementation of the CC during DC sweep, the full loop of the IV characteristics could not be observed. These underlying mechanisms agree with the excellent progressive SET process achieved by LCF devices in the previous section.

The trap-controlled switching mechanisms is driven by the modulation of the charge carrier trapping levels causing the shift of quasi-Fermi level towards the valence band under the external electric field [26]. The non-volatile change in conductance only starts to occur when the quasi-Fermi energy level meets the exponential trap energy distribution, $V > V_{TFL}$. The change in the slope of the TFSCLC region under different conductance states can be attributed to the different traps distribution characteristic, T_C . During the DC sweep of the SET process, as the voltage increased in the positive direction, the conduction profile changed from ohmic to TUSCL and followed by TFSCLC right before the switching occurs. The modulation of charge trapping level was indicated by the change in the slope of the curve, from ~ 7.09 to ~ 3.20 right before and after the SET process. Assuming the process occurred at constant room temperature, based on equation (4), a decrease in characteristic temperature can be extracted. During the SET process the characteristic temperature switched from $T_C \sim 1827$ K to $T_C \sim 660$ K. On the other hand, during the RESET process, the opposite shifts from $T_C \sim 501$ K to $T_C \sim 1824$ K was observed. This indicates that the switching from LCS to HCS and vice versa, consistently occurs between the two charge-trapping levels, reflected by the unique value of T_C . Lower T_C value corresponds to trap distributions varying more rapidly, while higher T_C value represents a slower varying trap distribution with respect to the energy. The change in T_C value is used as a parameter indicator to determine the role of the charge trapping levels modulation in conductance switching process. The increase in conductance value is accompanied with the decrease in T_C for trap-controlled SCL dominant switching mechanism.

The retention of the trapped charges after the removal of external electric field can be explained by structural reordering near the trap sites. The trapped carriers escape frequency can be expressed by the following [22, 26]:

$$\nu = \nu_0 \exp \left[- \left(\frac{T_c}{T} - \frac{\Delta S}{k} \right) \right]. \quad (5)$$

where ν_0 is the fundamental escape frequency and ΔS is the entropy change required to cause an escape from a kT_c trap depth. The retention time of the trapped charges is the inverse of the escape frequency. Thus, in order for the charges to remain trapped as long as possible after the removal of the SET voltage, the escape frequency value must be as low as possible, which implies the requirement of large decrease in the entropy of the system $\Delta S < 0$, i.e., increase in the symmetry or ordering of the material system. During the RESET process, beyond certain threshold voltage (V_{TFL} of ~ -0.91), the ordering is damaged with the increase of entropy, i.e., $\Delta S > 0$, causes the trapped charges to be released and the conductance value decreases.

Figure 3.11. The conduction mechanisms involved under SCF treatment. Similar to LCF devices, the conduction fits well under Ohmic (red), SCL (green), and TF-SCL (blue) conduction at different voltage regions.

On the other hand, SCF devices exhibited the three conduction regimes for the LCS while solely ohmic conduction mechanism was observed in HCS, as seen on **Figure 3.11**. Different from LCF devices, the V_{ON} is higher than the V_{READ} , thus the read operation of the devices was performed in the optimized linear region of the conduction. The V_{TFL} extracted at LCS is ~ 0.63 V, which is significantly smaller than the V_{TFL} of LCF devices and more importantly quite far from the median of the SET voltage from **Figure 3.2(a)**. Thus, the

switching mechanism of the device is potentially dominated by the modulation of oxygen vacancy defects profile at the filamentary gap during the SET process rather than the transition from unfilled to filled traps conduction, as reported in numerous HfO_x/Ti devices. However, the transition of the conduction mechanism from trap-controlled SCLC at LCS to ohmic conduction at HCS prompted the investigation upon the possible transition in dominant switching mechanisms.

Figure 3.12. The conduction mechanisms of SCF devices were investigated under different CCs and SVs. The trap-controlled SCL dominant switching mechanism is indicated by the change in trapping level parameter T_c .

The conduction models fitting was performed on the DC sweep with different CC values during SET process and different SV values during RESET process, as seen on **Figure 3.12**. In the trap-controlled SCL switching mechanism, the change in conductance must be accompanied with the change in T_c in the opposite manner. In the SET direction of the DC sweep, the change in conductance from LCS to the higher conductance state after the SET process at $50 \mu\text{A}$ CC was not accompanied by the change in T_c , which stayed around the same value of $\sim 843 \text{ K}$. Unfortunately, the TFSCLC region was no longer observable from $50 \mu\text{A}$ CC and above. On the other hand, the TFSCLC region was started to be observable after SV of -0.9 V , with first extracted T_c of $\sim 810 \text{ K}$. As the absolute value of SV increased and the conductance value decreased, instead of exhibiting an increasing trend, the value of T_c fluctuates between $\sim 710 \text{ K}$ to $\sim 850 \text{ K}$. These findings suggest that in the lower device conductance regime, the switching mechanism was dominated by the oxygen vacancy defects modulation at the filamentary gap, instead of the change in charge carrier trapping level. Based on the experimental observation

obtained for LCF and SCF devices, the switching mechanism proposed can be illustrated in the **Figure 3.13**.

Figure 3.13. The proposed switching mechanism under different forming treatment. Trap-controlled SCL switching dominate LCF device and SCF device near HCS regime. On the other hand, filamentary oxygen vacancies gap modulation mainly governs the SCF device near LCS regime.

3.4. SUMMARY

In this chapter, the characterization of Pt/HfO_x/Ti/Pt devices under different electroforming treatments was extensively discussed. Multilevel conductance capability of the devices was investigated under DC sweep and pulse programming. Excellent progressive RESET characteristics was achieved through optimized pulse programming amplitude and duration for the devices. On the other hand, progressive SET behaviour could only be achieved in LCF devices, while the SET process under increasing pulse amplitude scheme for SCF devices exhibited relatively more stochastic nature. This was attributed to different dominant switching region involved during the device operation. From the experimentally fit model, the correlation between the key parameter associated with the trap distribution (T_c) and the direction of the conductance switching was observed, which can be attributed to oxygen vacancies modulation at the HfO_x/Ti interface. The trap-controlled SCL switching mechanism

was observed in LCF devices, which can be used to explain the origin of the excellent progressive SET behaviour under the non-identical voltage pulse programming. On the other hand, the analysis on the DC sweep of SCF devices under different CC and SV values suggested that the transition from lateral oxygen vacancy defects modulation near inert electrode (filamentary gap model) in the lower conductance regime to trap-controlled SCL dominated switching mechanism driven by the switching at the interface. This insight is critical to unlock more ideal device properties, not only for high density storage class memory, but also analog synaptic device for neuromorphic computing applications. This understanding was further implemented in Chapter 4, where the key synaptic properties of the Pt/HfO_x/Ti memristive devices are analysed.

3.5. REFERENCES

- [1] B. Govoreanu *et al.*, "10×10nm² Hf/HfO_x Crossbar Resistive RAM with Excellent Performance, Reliability and Low-Energy Operation," in *2011 International Electron Devices Meeting*, 2011, pp. 31.6.1-31.6.4.
- [2] Y. Hou *et al.*, "Sub-10 nm Low Current Resistive Switching Behavior in Hafnium Oxide Stack," *Applied Physics Letters*, vol. 108, no. 12, p. 123106, 2016.
- [3] Z. Fang *et al.*, "The Role of Ti Capping Layer in HfO_x-Based RRAM Devices," *IEEE Electron Device Letters*, vol. 35, no. 9, pp. 912-914, 2014.
- [4] S. Z. Rahaman *et al.*, "The Role of Ti Buffer Layer Thickness on the Resistive Switching Properties of Hafnium Oxide-Based Resistive Switching Memories," *Langmuir*, vol. 33, no. 19, pp. 4654-4665, 2017/05/16 2017.
- [5] A. S. Sokolov *et al.*, "Influence of Oxygen Vacancies in ALD HfO_{2-x} Thin Films on Non-Volatile Resistive Switching Phenomena with a Ti/HfO_{2-x}/Pt Structure," *Applied Surface Science*, vol. 434, pp. 822-830, 2018/03/15/ 2018.
- [6] D. Walczyk *et al.*, "Resistive Switching Behavior in TiN/HfO₂/Ti/TiN Devices," in *2012 International Semiconductor Conference Dresden-Grenoble (ISCDG)*, 2012, pp. 143-146.
- [7] L. Goux *et al.*, "On the Gradual Unipolar and Bipolar Resistive Switching of TiN\HfO₂\Pt Memory Systems," *Electrochemical and Solid-State Letters*, vol. 13, no. 6, p. G54, 2010.

- [8] C. Sun, S. M. Lu, F. Jin, W. Q. Mo, J. L. Song, and K. F. Dong, "Control The Switching Mode of Pt/HfO₂/TiN RRAM Devices By Tuning The Crystalline State Of TiN Electrode," *Journal of Alloys and Compounds*, vol. 749, pp. 481-486, 2018/06/15/ 2018.
- [9] S. Yu, X. Guan, and H.-S. P. Wong, "Conduction Mechanism Of TiN/HfO_x/Pt Resistive Switching Memory: A Trap-Assisted-Tunneling Model," *Applied Physics Letters*, vol. 99, no. 6, p. 063507, 2011.
- [10] J. Bi and Z. Han, "Characteristics of HfO₂/Hf-Based Bipolar Resistive Memories," *Journal of Semiconductors*, vol. 36, no. 6, p. 064010, 2015/06 2015.
- [11] R. Nakajima, A. Azuma, H. Yoshida, T. Shimizu, T. Ito, and S. Shingubara, "Hf Layer Thickness Dependence of Resistive Switching Characteristics of Ti/Hf/HfO₂/Au Resistive Random Access Memory Device," *Japanese Journal of Applied Physics*, vol. 57, no. 6S1, p. 06HD06, 2018/05/17 2018.
- [12] P. Calka *et al.*, "Engineering of the Chemical Reactivity of the Ti/HfO₂ Interface for RRAM: Experiment and Theory," *ACS Applied Materials & Interfaces*, vol. 6, no. 7, pp. 5056-5060, 2014/04/09 2014.
- [13] W. Hu *et al.*, "Insights to the Influences of Electroforming Process on Resistive Switching Types In Pt/InGaZnO/W Memory Device," *Ceramics International*, vol. 44, pp. S88-S92, 2018/11/01/ 2018.
- [14] A. Marchewka, R. Waser, and S. Menzel, "Physical Modeling of the Electroforming Process in Resistive-Switching Devices," in *2017 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 2017, pp. 133-136.
- [15] M. Noman, A. A. Sharma, Y. M. Lu, M. Skowronski, P. A. Salvador, and J. A. Bain, "Transient Characterization of the Electroforming Process in TiO₂ Based Resistive Switching Devices," *Applied Physics Letters*, vol. 102, no. 2, p. 023507, 2013.
- [16] G.-H. Buh, I. Hwang, and B. H. Park, "Time-Dependent Electroforming in NiO Resistive Switching Devices," *Applied Physics Letters*, vol. 95, no. 14, p. 142101, 2009.
- [17] B. Butcher *et al.*, "Hot Forming to Improve Memory Window and Uniformity of Low-Power HfO_x-Based RRAMs," in *2012 4th IEEE International Memory Workshop*, 2012, pp. 1-4.

- [18] A. Kalantarian *et al.*, "Controlling Uniformity of RRAM Characteristics through the Forming Process," in *2012 IEEE International Reliability Physics Symposium (IRPS)*, 2012, pp. 6C.4.1-6C.4.5.
- [19] L. Zhao *et al.*, "Multi-Level Control of Conductive Nano-Filament Evolution in HfO₂ ReRAM by Pulse-Train Operations," *Nanoscale*, 10.1039/C4NR00500G vol. 6, no. 11, pp. 5698-5702, 2014.
- [20] F.-C. Chiu, "A Review on Conduction Mechanisms in Dielectric Films," *Advances in Materials Science and Engineering*, vol. 2014, p. 18, 2014, Art. no. 578168.
- [21] E. W. Lim and R. Ismail, "Conduction Mechanism of Valence Change Resistive Switching Memory: A Survey," *Electronics*, vol. 4, no. 3, pp. 586-613, 2015.
- [22] A. Odagawa *et al.*, "Colossal Electroresistance of a Pr_{0.7}Ca_{0.3}MnO₃ Thin Film at Room Temperature," *Physical Review B*, vol. 70, no. 22, p. 224403, 12/03/ 2004.
- [23] L. Brehmer, "Electrical Transport in Solids with Particular Reference to Organic Semiconductors. Von K. C. KAO und W. HWANG. 1. Auflage. Oxford/New York/Toronto/Sidney/Paris/Frankfurt: Pergamon Press 1981. 663 S., US \$ 120.-, £ 50.-," *Acta Polymerica*, vol. 33, no. 1, pp. 91-91, 1982.
- [24] P. Mark and W. Helfrich, "Space-Charge-Limited Currents in Organic Crystals," *Journal of Applied Physics*, vol. 33, no. 1, pp. 205-215, 1962.
- [25] H. T. Nicolai, M. M. Mandoc, and P. W. M. Blom, "Electron Traps in Semiconducting Polymers: Exponential Versus Gaussian Trap Distribution," *Physical Review B*, vol. 83, no. 19, p. 195204, 05/04/ 2011.
- [26] D. S. Shang, Q. Wang, L. D. Chen, R. Dong, X. M. Li, and W. Q. Zhang, "Effect Of Carrier Trapping On The Hysteretic Current-Voltage Characteristics in Ag/La_{0.7}Ca_{0.3}MnO₃/Pt Heterostructures," *Physical Review B*, vol. 73, no. 24, p. 245427, 06/22/ 2006.

CHAPTER 4

Synaptic Behaviour of Pt/HfO_x/Ti Anion-based Memristive Devices

In this chapter, the synaptic characteristics, i.e., LTP and LTD, of Pt/HfO_x/Ti anion-based memristive devices are extensively investigated. Firstly, the trade-off in device dynamic ratio is presented to improve the bidirectionality of the SCF devices. The performance of LCF and SCF devices under IPP scheme was compared under the same dynamic ratio. With lower asymmetric non-linearity (ANL) factor obtained by SCF devices, further characterization of the SCF devices was performed under different pulse amplitude and pulse number during potentiation and depression process. Seemingly linear correlation between the device dynamic ratio and ANL was observed. The ANL increases as the dynamic ratio increases. On the other hand, the increase in pulse number resulting in increase in dynamic ratio seems to achieve the opposite. The potential origin of these observations is discussed. The chapter ends with the reliability test on the SCF devices under NPP and IPP scheme.

4.1. DIFFERENT LEARNING APPROACHES AND CORRESPONDING SYNAPTIC DEVICE REQUIREMENTS

The implementation of various redox-based memristive devices as a synaptic element in artificial neural network (ANN) has been widely investigated. This is mainly due to the promising non-volatile multilevel conductance feature of the devices with highly scalable array architecture, i.e., two-terminal nature of the device enables the pure crossbar array implementation under different array configurations. The memristive devices have also been reported to have excellent device latency, endurance, and retention capability. The non-volatility of these devices allows the on-chip storage of the weight values that can further accelerate the processing speed and reduce energy consumption on the system level. The multilevel conductance states property is required to achieve analog synaptic performance in facilitating the programming intensive weight update activities during the ANN training. This is not only beneficial to the network learning capability but also the overall physical footprint of the network on the chip. The more bits can be stored in each synaptic cell, the smaller number of synapses required to deliver certain task and thus the smaller area required for the synaptic array on the chip.

The learning process of an ANN can be performed by two different approaches, i.e., *ex-situ* and *in-situ* learning [1]. In *ex-situ* learning, the training of the weights is done externally in a software-based system, followed by the conversion of the trained weights into conductance values in the synaptic array. The transferring of the weight values into each device conductance level is done once before the chip can be deployed to carry out any inference or classification task. On the other hand, the *in-situ* training performs the weight updates on the hardware directly rather than just transferring pre-trained weights from the software. These two learning approaches will lead to different synaptic requirements. As an example, the endurance and retention aspect of the devices. In *ex-situ* learning, the weight values are not refreshed as frequent as in *in-situ* learning, thus the endurance requirement can be more relaxed. However, the retention capability of the device is critical to increase the number of inference or classification tasks can be executed before the weights need to be re-programmed. For *in-situ* training, the weight values are updated frequently, leading to programming-intensive treatment that requires relatively higher endurance capability. After the learning process is completed, the synaptic elements need to maintain the weight values until the next training data is being input to the network.

The analog property of the memristive devices has been used to mimic different synaptic behaviours such as long-term potentiation (LTP) and long-term depression (LTD). However, the inherent nonidealities of memristive devices, such as device variability, stochasticity, and yield, has been shown to have a negative impact towards the network learning accuracy. For *ex-situ* learning, the stochasticity of the conductance values can be mitigated by the implementation of write-verify scheme. Relatively more complex yet efficient programming schemes can be implemented with the proper design of peripheral circuitry. The main challenge for the *ex-situ* learning will be the presence of Random Telegraph Noise (RTN) in the memristive devices, which might be amplified when the network needs to perform current sum operation. For *in-situ* learning approach, it has been shown that the ANNs are most sensitive towards the behaviour of the weight update as a function of the programming pulse rather than those random effects present in the array of synaptic devices, especially during the learning activities of the network. In order to accommodate *in-situ* training, an ideal synaptic device should be able to maintain linear and symmetrical response throughout the synaptic weight update activities, i.e., conductance increase (SET, LTP) and conductance decrease (RESET, LTD). The linearity aspect of the conductance update refers to the relationship between the change in conductance value per programming cycle, while the symmetry characteristic is associated with the change in conductance value per programming cycle in the opposite directions. These are critical features to enable direct mapping between the weight values and device conductance. The deviation from these two properties have been shown to result in severe degradation of the network learning accuracy.

The ideal linearity and symmetry of the synaptic behaviour further extend their importance in the process of designing the suitable programming and required circuitry during the network training. The non-linear and asymmetrical behaviour of the conductance update will result in history-dependent conductance response because the change in conductance varies with the conductance value. The impact of this history-dependent behaviour on the network accuracy has been simulated with the help of a “jump-tables” [2-5]. Different types of conductance modulation behaviour in memristive devices have been recorded and their corresponding jump-tables were constructed. Different effects from the different jump-tables on the network learning accuracy were captured. Those behaviours can be classified into bounded or unbounded, linear or non-linear, and unidirectional or bidirectional. In all reported devices, the conductance values are modulated within a finite range, thus only bounded conductance behaviour is considered practical. This unbounded nature of the conductance

caused a slight degradation on the network learning accuracy. The uni- and bi-directionality of the conductance update are used to describe symmetry aspect of the weight update, i.e., the feasibility of the gradual change in conductance of the device. As an example, the memristive devices with abrupt SET and progressive RESET behaviour are classified as unidirectional devices, while those devices exhibit progressive SET and RESET are considered bidirectional.

In most of the reported two-terminal memristive-based synaptic devices, non-linear and asymmetrical behaviour of the conductance response is commonly observed. Different approaches have been investigated to mitigate the issue from the programming and device engineering viewpoint. From programming perspective, linearity and symmetry of the conductance response can be improved with the use of non-identical pulse programming (NPP) scheme, i.e., pulse train with increasing amplitudes and/or durations, during LTP and LTD. However, in large synaptic array, state-independent programming using identical pulse programming scheme (IPP) is more desirable and practical as compared to NPP because of the complex peripheral circuitry required for non-identical pulse generation. The pulse parameters need to be adjusted based on the current conductance state value, i.e., current state reading is required before each programming cycle, to ensure accurate weight jump. This is a major challenge for the implementation of in-situ learning functionalities, while the non-linearity and asymmetry can be addressed with write-verify approach in ANN with ex-situ learning. Thus, the main objective of this section of the work is to investigate the device conductance response under NPP and IPP scheme.

4.2. TRADE-OFF BETWEEN DYNAMIC RATIO AND PROGRESSIVE CONDUCTANCE SWITCHING BEHAVIOUR OF THE DEVICES

Large dynamic ratio of conductance can be considered as one of the important features for synaptic device implementation. It is required to provide sufficient read margin between the conductance states and reducing the effect of soft errors during the reading of the weights. However, among all the properties required to reach the objective of an ideal synaptic device, it has been shown that the bidirectionality of the conductance update has bigger impact on the learning accuracy of the neural network as compared to degradation induced by the inherent random effects present on the device and array level [5]. The drop in learning accuracy has been shown in the abrupt SET behaviour observed in Ti/HfO₂ devices and a significant improvement was observed through material engineering to achieve the gradual SET process [6].

As discussed in Chapter 3, Pt/HfO_x/Ti-based memristive devices have been demonstrated to have progressive RESET and abrupt SET for SCF devices, while both progressive SET and RESET behaviour was obtained under LCF treatment. The origin of the progressive SET behaviour on the LCF devices was investigated through the fitting of the SCLC conduction model. It was found that the trap-controlled SCL switching mechanism lead to the gradual change in conductance, not only during the RESET, but also the SET operation. The possible transition between oxygen vacancy modulation to trap-controlled SCL dominated switching was suggested in SCF devices based on the experimentally fitted model of the DC sweep under different CC and SV values. Based on the model, the trap-controlled SCL switching mechanism potentially dominated the switching process in the higher conductance range rather than the lower ones. Thus, the progressive SET and RESET behaviour of the SCF devices are further improved through pulse parameters optimization in the higher conductance regime.

The optimized programming pulse parameters under NPP scheme with SET amplitude from 0.71 V to 0.90 V and RESET amplitude from -0.65 V to -1.10 V under fixed pulse duration of 200 ns were implemented. The SCF devices successfully achieved both progressive SET and RESET with excellent conductance distribution, as depicted on the **Figure 4.1**. This supports the idea in which the trap-controlled SCL switching mechanism is more dominant at higher conductance range for the SCF devices. This is a significant improvement as compared to the wider range of conductance utilized previously, where more stochastic SET process was obtained under the non-identical SET pulse amplitudes.

Figure 4.1. (a) The conductance switching behaviour of the SCF devices utilizing the full range of conductance under SET 0.76 V to 1.1 V with 10 mV interval and RESET of -0.65 V to -1.5 V with interval of 25 mV. (b) More linear region of conductance behaviour was achieved under 0.71 V to 0.90 V and RESET amplitude from -0.65 V to -1.10 V, with the associated conductance distribution (c, d)

Furthermore, not only more deterministic progressive SET operation was achieved but also a significantly improved conductance distribution during RESET process, can be seen on the **Figure 4.2**. However, the dynamic ratio was reduced from ~ 60 to ~ 10 . This range of conductance and programming parameters were used as a reference to optimize the identical pulse programming scheme, enabling state-independent programming operation of the devices.

Figure 4.2. More uniform conductance distribution was achieved under reduced conductance range.

The same set of pulses were implemented on LCF devices to make a direct comparison in terms of operating voltages. The conductance response of the device obtained was improved in terms of linearity and symmetry. However, the dynamic ratio of the device was significantly reduced from ~ 25 to ~ 1.6 . This suggests higher voltage pulse is required to achieve the equivalent dynamic ratio observed in SCF devices. This was also taken into consideration in optimizing the identical pulse train parameters.

Figure 4.3. The conductance switching behaviour of the LCF device utilizing the full range of conductance (a) as compared to the more linear region (grey area in (a)) of the conductance response (b). Improved linearity and symmetry were attained, however the dynamic ratio reduced significantly to ~ 1.6 . The distribution of the conductance values under the progressive SET (c) and progressive RESET (d) within smaller conductance

4.3. IDENTICAL PULSE PROGRAMMING OPERATION OF LCF AND SCF DEVICES

In IPP scheme, fixed duration of 200 ns pulse was implemented throughout the experiment. Difference in amplitude dependence between the potentiation and depression were investigated. This is a critical parameter to achieve potentiation and depression with stable conductance range. If tuning of the amplitude is not properly optimized, the devices tend to drift towards one side of the conductance and either get stuck in the extreme conductance values or having a very small dynamic ratio of less than 1. As discussed in Chapter 3, the conductance update behaviour of the devices depends not only on the programming parameters, but also the starting point of the conductance. Thus, based on the conduction value distribution

of the devices, i.e., tighter distribution near HCS, as well as the bidirectionality of the device conductance modulation, the IPP operation was implemented starting from HCS of the corresponding devices. Thus, the switching operation was always started with depression rather than potentiation. At the start of the potentiation and depression cycles, the devices exhibited decrease in maximum and minimum conductance value before reaching an equilibrium range, as depicted on the **Figure 4.4**. This method was maintained throughout the IPP scheme implementation.

Figure 4.4. The decreasing trend of maximum and minimum conductance is considered as non-equilibrium range of the conductance response. The investigated conductance responses were characterized within equilibrium region (Potentiation: 0.75V, 200ns; Depression: -0.80V, 200ns)

Both LCF and SCF devices achieved non-linear, bounded, and bidirectional conductance update, as displayed in **Figure 4.5**. The devices exhibited different conductance range during the IPP operation. The range of conductance obtained in LCF devices were lower compared to the SCF devices with generally higher operating voltages were implemented for the LCF devices. This can be attributed to different oxygen vacancy defects profile initiated during the forming process.

Figure 4.5. Conductance modulation behaviour of LCF (a) and SCF (b) devices under IPP scheme. The potentiation and depression pulse parameters are highlighted in green and red respectively. Each potentiation and depression cycle consist of 256 programming pulses.

The non-linear and asymmetrical conductance modulation as a function of the cumulative number of programming pulses can be mathematically modelled by the following set of equations [7]:

$$G_P(P) = G_{\min} + \frac{\alpha P}{P + e^\beta}, \quad (1)$$

for conductance modulation during potentiation,

$$G_D(D) = G_{\max} - \frac{\alpha(N-D)}{(N-D) + e^\beta}, \quad (2)$$

for conductance modulation during depression, and α can be described as:

$$\alpha = \frac{(G_{\max} - G_{\min})(N + e^\beta)}{N}, \quad (3)$$

N is the maximum number of pulses for each of the complete potentiation and depression. P and D represent n^{th} programming pulse for potentiation and depression respectively, while β is the adjustable fitting parameter. The equations fit well with the conductance behaviour of both type of devices.

Figure 4.6. The conductance modulation fits well with the mathematical model for both SCF (a) and LCF (b) devices. The normalized conductance plot (c) from the fitted mathematical model shows that the SCF device exhibits better linearity during potentiation, thus achieving overall lower ANL factor. These can be ascribed to the significantly sharper drop in conductance change at the start of the potentiation cycle (d) for the LCF devices.

In order to make a fair comparison between the two different forming treatments, optimized pulse parameters to achieve the same dynamic ratio for both devices were implemented. Dynamic ratio of ~ 2.5 were achieved for LCF and SCF devices, as depicted on the **Figure 4.6(a)** and **(b)**. The normalized values of conductance based on the experimentally fitted mathematical model for both devices are plotted in **Figure 4.6(c)**. This provides a complete picture in terms of linearity and symmetry property of the devices. The SCF devices achieved a better overall linearity characteristic as compared to the LCF devices. Both devices exhibited similar linearity during the depression process, but significantly less linear behaviour can be observed from LCF devices during the potentiation. Furthermore, the rate of change of the conductance, ΔG , against the starting conductance G were also plotted in **Figure 4.6(d)**. The rate of change in conductance reduces more sharply after the first few programming pulses of the LCF devices during the potentiation as compared to the SCF devices. This can be ascribed to the observation of different charge-trapping level involved during the switching, indicated by different T_C values extracted for LCF and SCF devices in Chapter 3. Thus, further optimization on the pulse parameters implemented during the programming of LCF devices are required.

Trade-off between the device dynamic ratio and the conductance update linearity and symmetry has been observed in several memristive-based synaptic devices. However, the direct correlation between these two parameters has not been investigated. Thus, thorough characterization on the synaptic behaviour of the SCF devices under different pulse amplitudes and numbers were performed to investigate the correlation between the dynamic ratio and the non-linearity as well as asymmetry property of the devices. To simplify the analysis, the asymmetric non-linearity factor (ANL) is defined as [7]:

$$ANL = \frac{G_P(N/2) - G_D(N/2)}{G_{\max} - G_{\min}}, \quad (4)$$

For an ideal synaptic device with linear and symmetrical conductance update ANL is equal to zero. Thus, the further this value from zero, the further the device properties from an ideal synaptic device performance. The ANL values for the corresponding potentiation-depression curve on **Figure 4.6** are ~ 0.90 and ~ 0.83 for LCF and SCF devices respectively.

4.4. CORRELATION BETWEEN DYNAMIC RATIO AND ASYMMETRIC NON-LINEARITY FACTOR

In order to achieve different dynamic ratio on the SCF devices, the pulse amplitudes and the number of pulses during potentiation and depression cycles were varied under constant 200 ns pulse duration. The first approach was performed by keeping constant potentiation amplitude at +0.80 V, while varying the depression pulse amplitude from -0.80 V to -0.95 V, relatively constant maximum conductance values were achieved with decreasing minimum values, as displayed on the **Figure 4.7**. The ANL increased from ~0.55 to ~0.99 with the increase in dynamic ratio ~1.1 to ~13.8 followed.

Figure 4.7. The conductance response with constant potentiation amplitude and increasing depression amplitude. Increasing dynamic ratio was achieved with relatively constant maximum conductance.

The second approach was conducted by increasing both potentiation and depression amplitudes linearly, keeping constant difference of 0.10 V. Increasing maximum conductance level and decreasing minimum conductance level were obtained. The same trend can be observed with ANL increased from ~ 0.66 to ~ 1.38 and dynamic ratio increased from ~ 1.30 to ~ 6.07 .

Figure 4.8. The conductance response with linearly increasing potentiation and depression amplitudes. The dynamic ratio increases as the maximum conductance increases and minimum conductance decreases.

Under different pulse amplitudes, both starting and end point of the conductance during the potentiation and depression were modulated, which resulted in the change of dynamic ratio of the device conductance, as can be seen on **Figure 4.9**. It can be observed in both the amplitude modulation schemes that the ANL increases as the dynamic ratio increases. However, there is a significance difference in the rate of change of ANL with respect to the dynamic ratio. When the potentiation pulse was kept constant with increasing amplitude of depression pulses, higher dynamic ratio could be achieved while maintaining relatively lower ANL as compared to the scheme in which both amplitudes were increased. This could be

associated with the higher degree of non-linearity induced during the potentiation process as compared to the depression. When higher amplitude of potentiation pulse was implemented, sharper drop in conductance change occurred at the beginning of the potentiation cycle. This resulted in the more aggressive increase in ANL factor of the device. Thus, further optimization especially in the potentiation pulse duration is required to improve the linearity of the devices while maintaining sufficiently high dynamic ratio.

Figure 4.9. The change in ANL factor as a function of device dynamic ratio. Both schemes show nearly-linear increase in ANL with the increase in dynamic ratio.

The third approach was performed by keeping constant pulse parameters with potentiation pulse amplitude of +0.75 V and depression amplitude of -0.85 V, while varying the number of pulses from 64 pulses to 512 pulses for each potentiation and depression cycles. The maximum conductance increased, and the minimum conductance decreased as more programming pulses were implemented during potentiation and depression cycle. The ANL values were relatively constant for 64, 128, and 256 with the increasing dynamic ratio. This is due to the same conductance response, i.e., conductance rate of change as a function of conductance level, was achieved under the same pulse parameters. However, a significant drop was observed with 512 programming pulses, as depicted on the **Figure 4.10**. This is due to the overall potentiation and depression behaviour was dominated by the more linear region of the conductance response. This behaviour can be beneficial in improving the dynamic ratio of the

conductance. For example, if each programming cycle consists of 8 pulses instead of a single pulse, significant improvement in both dynamic ratio and ANL can be achieved.

Figure 4.10. The conductance modulation under different number of programming pulses (a). The ANL factor trend as a function of dynamic ratio (b).

4.5. RELIABILITY ASPECT OF THE DEVICES

During the training process of the ANN, the conductance of the devices needs to be updated frequently according to the weight update required by the algorithm. Thus, sufficient endurance capability is required to ensure the device is able to sustain the learning process, i.e., maintaining the conductance range, symmetry, and linearity. The requirement of the device endurance highly depends on the number of weight update required, which is related to the learning algorithm and the target applications.

The SCF devices underwent two different endurance tests. The first endurance test was performed under NPP scheme. The previously optimized programming pulse parameters with potentiation amplitudes from 0.71 V to 0.90 V and depression amplitudes from -0.65 V to -1.10 V under fixed pulse duration of 200 ns followed by read pulse of 0.2 V amplitude and 100 μ s duration were implemented. This was conducted to investigate the response of the device under higher programming voltages and larger range of conductance. The average of first, 10th, 20th, 30th, and 40th 20 cycles consisting of full potentiation and depression range with total of 40 conductance updates in each cycle are plotted in **Figure 4.11**. The SCF device successfully maintained the average LCS of ~ 45 μ S with dynamic ratio of ~ 11 for more than 1000 full potentiation and depression cycles comprising of over 40,000 conductance updates.

Figure 4.11. The endurance test of the SCF device under NPP scheme. No significant degradation can be observed

The second endurance test was performed under IPP scheme with pulse amplitude of +0.75 V during potentiation and -0.85 V during depression. Each potentiation and depression cycle contain 256 programming pulses. The endurance test was performed for 1500 potentiation-depression cycles, equivalent of 768,000 weight updates. It successfully maintained average dynamic ratio of ~ 2.67 with ANL ~ 0.92 . The average of first, 10th, 20th, 30th, and 40th 20 cycles consisting of full potentiation and depression range with total of 40 conductance updates in each cycle are depicted on the **Figure 4.12**.

Figure 4.12. The endurance test of the SCF device under IPP scheme. No appreciable degradation can be observed

4.6. SUMMARY

The synaptic behaviour of Pt/HfO_x/Ti anion devices was investigated under different pulse scheme, i.e., NPP and IPP. Under NPP, the trade-off in device dynamic ratio successfully improved device progressive switching behaviour and the uniformity of the conductance distribution, especially in SCF devices. This enables the implementation of compliance-free anion synaptic device for highly scalable crossbar array architecture. Under IPP scheme, it was found that the SCF device could achieve better symmetry and linearity as compared to the LCF device while maintaining the same dynamic ratio. This was associated with the charge trapping level involved during the switching process. Excellent endurance capability was maintained with more than 40,000 weight updates for NPP and more than 700,000 weight updates for IPP scheme.

4.7. REFERENCES

- [1] S. Yu, "Neuro-Inspired Computing with Emerging Nonvolatile Memorys," *Proceedings of the IEEE*, vol. 106, no. 2, pp. 260-285, 2018.
- [2] A. Fumarola *et al.*, "Bidirectional Non-Filamentary RRAM as an Analog Neuromorphic Synapse, Part II: Impact of Al/Mo/Pr_{0.7}Ca_{0.3}MnO₃ Device Characteristics on Neural Network Training Accuracy," *IEEE Journal of the Electron Devices Society*, vol. 6, pp. 169-178, 2018.
- [3] K. Moon *et al.*, "Bidirectional Non-Filamentary RRAM As An Analog Neuromorphic Synapse, Part I: Al/Mo/Pr_{0.7}Ca_{0.3}MnO₃ Material Improvements and Device Measurements," *IEEE Journal of the Electron Devices Society*, vol. 6, pp. 146-155, 2017.
- [4] G. W. Burr *et al.*, "Experimental Demonstration and Tolerancing of a Large-Scale Neural Network (165 000 Synapses) Using Phase-Change Memory as the Synaptic Weight Element," *IEEE Transactions on Electron Devices*, vol. 62, no. 11, pp. 3498-3507, 2015.
- [5] S. Sidler *et al.*, "Large-scale Neural Networks Implemented with Non-Volatile Memory as the Synaptic Weight Element: Impact of Conductance Response," in *2016 46th European Solid-State Device Research Conference (ESSDERC)*, 2016, pp. 440-443.
- [6] J. Woo *et al.*, "Improved Synaptic Behavior under Identical Pulses Using AlO_x/HfO₂ Bilayer RRAM Array for Neuromorphic Systems," *IEEE Electron Device Letters*, vol. 37, no. 8, pp. 994-997, 2016.
- [7] C. Chang *et al.*, "Mitigating Asymmetric Nonlinear Weight Update Effects in Hardware Neural Network Based on Analog Resistive Synapse," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 8, no. 1, pp. 116-124, 2018.

4.8. APPENDIX

T_c is determined by extracting the gradient of the log(I)-log(V) plot in the TFSCLC region. The derivation is as follows:

$$I_{TFSCLC} = \frac{Aq^{1-\tau}\mu_N n_c}{L^{2\tau+1}} \left(\frac{2\tau+1}{\tau+1} \right)^{\tau+1} \left(\frac{\tau}{\tau+1} \frac{\varepsilon_0 \varepsilon_R}{n_t} \right)^{\tau} V^{\tau+1}$$

$$\log(I_{TFSCLC}) = \log \left[\frac{Aq^{1-\tau}\mu_N n_c}{L^{2\tau+1}} \left(\frac{2\tau+1}{\tau+1} \right)^{\tau+1} \left(\frac{\tau}{\tau+1} \frac{\varepsilon_0 \varepsilon_R}{n_t} \right)^{\tau} V^{\tau+1} \right]$$

$$\log(I_{TFSCLC}) = \log \left[\frac{Aq^{1-\tau}\mu_N n_c}{L^{2\tau+1}} \left(\frac{2\tau+1}{\tau+1} \right)^{\tau+1} \left(\frac{\tau}{\tau+1} \frac{\varepsilon_0 \varepsilon_R}{n_t} \right)^{\tau} \right] + \log(V^{\tau+1})$$

$$\log(I_{TFSCLC}) = \log \left[\frac{Aq^{1-\tau}\mu_N n_c}{L^{2\tau+1}} \left(\frac{2\tau+1}{\tau+1} \right)^{\tau+1} \left(\frac{\tau}{\tau+1} \frac{\varepsilon_0 \varepsilon_R}{n_t} \right)^{\tau} \right] + (\tau+1)\log V$$

$$y = C + mx$$

The IV sweep was performed at room temperature (T = 300K) and the value was extracted before the conductance switching occurs at each sweep. Thus, T_c can be calculated by the following:

$$T_c = T(m-1)$$

CHAPTER 5

Cation-Based Diffusive Memristor

In this chapter, an experimental and theoretical framework of a cation-based diffusive memristor is presented ¹. The ability of the Pt/HfO_x/Cu/Pt structure to toggle between a memory element, i.e., non-volatile property, and a threshold switch, i.e., volatile property, is investigated. The key approach relies on the initial electroforming treatment given to the pristine Pt/HfO_x/Cu/Pt structure. A gradual electroforming process was implemented on the pristine devices to realize volatile threshold switching characteristics of a diffusive memristor. The devices exhibit stable unidirectional threshold switching properties with high selectivity of $>10^7$ and ultralow OFF current of ~ 100 fA for over 10^4 endurance cycles. Nucleation theory on spheroidal-shaped metallic filament growth is used to extensively discuss the structural changes of the device after gradual forming treatments by analysing the applied bias amplitude dependency of the finite delay time required by the device to turn ON under external electric field. On the other hand, the Rayleigh instability model was implemented for the aforementioned spheroidal metallic nucleus to explain the relaxation dynamics of the device. It was shown that the relaxation time of the device depends on the initial profile of the nucleus within the insulating layer. The reliability aspects of the device in terms of endurance and noise observation are also discussed.

¹Reproduced with permission from "Unidirectional Threshold Switching Induced by Cu Migration with High Selectivity and Ultralow OFF Current under Gradual Electroforming Treatment," ACS Applied Electronic Materials, vol. 1, no. 10, pp. 2076-2085, 2019/10/22 2019. Copyright 2019 American Chemical Society.

5.1. DIFFUSIVE MEMRISTOR HIGH DENSITY CROSSBAR ARRAY AND OTHER EMERGING SYSTEMS

Memristive devices are promising non-volatile memory technologies due to its high scalability, high-speed operation, low power consumption, and multibit per cell capability. These desirable characteristics are required for both embedded and standalone mass storage devices. Recently, the artificial intelligence technology, which utilises new high-level computing approach, has been on a quest in search of ideal synaptic devices to realize a highly dense computing platform, which requires high synaptic weight precision. Thus, the high scalability nature and multibit per cell capability demonstrated in many memristive devices have attracted a great deal of attention from the neuromorphic computing community to fulfil their ideal synaptic device requirements.

In order to fully utilise the scalability of a memristive device through highly connected crossbar array, a select device or selector is required to work in tandem with the memory element to mitigate the inherent sneak path current issue of the architecture. In the recent years, different types of selectors have been widely investigated. These selectors can be classified into two major categories, *i.e.* non-linear IV and threshold IV selector. Non-linear IV selector are those devices implementing metal-oxide Schottky barrier modulation [1-7], crested and variable oxide tunnel barrier [8-13], and mixed of ionic electronic conductors (MIEC) [14-17]. On the other hand, the threshold IV selector consists of ovonic threshold switch (OTS), [18-20] metal-insulator transition (MIT) [21-24], and diffusive memristor (DM) [7, 25-45]. Among these select devices, DM has been found to exhibit the most promising performance in terms of selectivity and OFF current. In addition to the crossbar array integration, DM devices, due to their distinctive device dynamics, are implemented in emerging systems, *e.g.* true random number generators [27], steep subthreshold slope transistors [30, 33], and artificial synaptic devices [28, 29, 37].

The reported DM devices can be classified into Ag [7, 25-40] and Cu [40-45] based devices. DM devices work based on the electric field induced migration of active metal atoms, *e.g.* Cu and Ag, with high mobility inside dielectric layer. The device is switched ON when the applied electric field exceeds the threshold field required to form unstable metallic filament within the dielectric and is switched OFF due to its dissolution after the removal of electric field. This mechanism is analogous to electrochemical metallization (ECM) cells. The same structure might exhibit both non-volatile and volatile switching characteristics, which can be

used for memory and threshold switch applications respectively [35, 44]. Thus, complete control over these two opposite characteristics must be thoroughly investigated.

Gradual electroforming process from pristine Pt/HfO_x/Cu/Pt structure is introduced to achieve volatile threshold switching rather than non-volatile memory characteristics. Crosspoint device structures of 10 $\mu\text{m} \times 10 \mu\text{m}$ cell area, shown in **Figure 5.1**, were fabricated by a two-step UV-lithography patterning followed by lift-off processes. All layers of the materials were deposited via magnetron sputtering deposition, *i.e.* DC sputtering for metals and RF sputtering from ceramic target (HfO₂) for the oxide layer. The thickness of each layer in the structure is 10nm. The electrical characterizations were done using Keithley 4200A-SCS Parameter Analyzer, *i.e.* Source Measure Unit (SMU) and preamplifier for DC IV measurement and Pulse Measure Unit (PMU) for fast IV measurement.

Figure 5.1. (a) Top view of the crosspoint device under optical microscope. (b) 3D view of the structure taken by atomic force microscopy

5.2. GRADUAL FORMING PROCESS AND THRESHOLD SWITCHING CHARACTERISTICS OF DM

In order to specifically achieve threshold switching properties, two common fabrication techniques have been implemented. The first technique involves co-sputtering and/or reactive sputtering deposition to form metal (Cu or Ag) and insulating host composites. It requires high control over the number of metal atoms within the insulating host. The second technique utilises a post-annealing treatment to allow metal atoms diffusion into insulating layer. Temperature and duration of the annealing process are the two critical parameters to achieve

the desired properties. Due to the bulk nature of the atomic diffusion, it usually requires an insertion of a thin tunnelling layer to maintain the high selectivity of the threshold switch. HfO_x/Cu-based structures have been reported as to exhibit both memristive and threshold switching properties within one structure. This technique was implemented to alter the memristive property of the device into volatile threshold switching characteristics [44]. However, the temporal response of the device has not been discussed.

It has been reported that the compliance current (CC) during the operation of DM plays an important role in controlling the nature of metallic filament induced under external electric field [26, 35, 45]. Relatively higher CCs have been observed to result in non-volatile memory switching behaviour, while the lower CCs lead to unstable metallic filament resulting in volatile threshold switching characteristics. For the structure investigated in this work, the ON state current target was set to be 10 μ A or higher, which was based on the common switching current range reported for memristive devices. The devices underwent direct forming using compliance current of 10 μ A or higher tend to form strong conductive filament, which resulted in non-volatile switching behaviour, depicted in **Figure 5.2**.

Figure 5.2. The behaviour of the devices underwent direct electroforming under 10 μ A CC showed a mix of volatile and non-volatile switching behaviour in between cycles (a), while the devices underwent gradual forming treatment exhibited pure volatile threshold switching characteristics (b).

The electroforming process was performed by gradually increasing the CC from 100 nA to 10 μ A with certain subsequent CC ratio. Two critical parameters in the gradual forming treatment are the starting CC value and the subsequent CC ratio. It was found that the optimized forming condition to achieve stable volatile threshold switching characteristics were with the starting CC of 100 nA and subsequent CC ratio from 2 to 40.

Figure 5.3. Gradual forming process of the device from the lowest compliance current (a) to the highest one (c).

More detailed discussion on the implementation of the gradual forming treatment is discussed in the following.

Implementation of Gradual Forming Treatment and Constraints

Different starting CC of the forming treatment was used, i.e. 100 nA, 1 μ A, and 10 μ A, to investigate the switching characteristics of the devices. Without gradual forming treatment, the devices could only achieve stable volatile switching behaviour under 100 nA and 1 μ A CC, as shown in **Figure 5.4**. However, a mix of volatile and non-volatile switching behaviour was observed for 10 μ A CC.

Under the starting CC of 100 nA, different increments of CC value were investigated, ratio of 2 and 40. The maximum ON current of 102.4 μ A and 160 μ A were achieved respectively, as depicted in **Figure 5.5** and **Figure 5.6**. From the smallest increment of CC ratio of 2, the devices started to have non-volatile switching behaviour from 204.8 μ A CC value. By comparing **Figure 5.5** and **Figure 5.6**, it can be observed that despite being able to successfully achieve volatile switching characteristics without gradual forming treatment under

1 μA CC, the properties of the device achieved are less desirable because of the lower threshold voltage and higher variation of OFF current value as compared to the devices underwent gradual forming treatment.

Despite successfully achieving ON current of 160 μA , the subsequent investigations were done on lower current regime. This is due to noticeably long relaxation time observed in the devices with higher CC used. Thus, only the ON current regime of 10 to 20 μA was investigated for the rest of the measurements.

Figure 5.4. The volatile threshold switching characteristics achieved under 100 nA and 1 μA without gradual forming treatment. Significant difference in average threshold voltage of the devices was observed, ~ 0.4 V and ~ 0.2 V for 100 nA and 1 μA CC respectively. Higher variation of OFF current can be observed for the device underwent direct forming at 1 μA CC.

Figure 5.5. Switching behaviour of the devices under starting CC of 100 nA and subsequent CC ratio of 2. Volatile switching behaviour was successfully maintained up to 102.4 μ A (Figure 3a to 3c). At 204.8 μ A CC, the devices exhibited non-volatile switching behaviour after several cycles.

Figure 5.6. Switching behaviour of the devices under starting CC of 100 nA and subsequent CC ratio of 40. Volatile switching behaviour was successfully maintained up to 160 μA (**Figure 5.6(a) to 5.6(c)**).

During the gradual forming treatment, the devices had to undergo long duration of voltage stress under DC sweep. Any structural changes exhibited by the device during this forming process is critical for the subsequent device operation. Thus, single run and multiple runs (50 cycles) of DC gradual forming treatment for each CC value were conducted to investigate the device structural changes during the forming process. The multiple forming treatments were performed by repeating this forming steps 50 times at each CC value. The multiple gradual forming cycles is not a necessity to obtain functioning DM devices in 10 μA or higher regime ON current, but the gradual forming treatment is. Under single gradual forming cycle, a more controllable formation conductive filament was achieved, as compared to those in direct single forming treatment. With the use of multiple cycles of forming runs for each CC, the decrease in threshold voltage during DC sweep and the decrease in device delay time during pulse measurements were experimentally observed as compared to the single gradual forming treatment. The delay times variation was also found to be lower under the

multiple gradual forming treatment. This phenomenon can be explained by nucleation theory in which there was an increase in amount Cu atoms residing within dielectric, which potentially resulted in the decrease of effective oxide thickness at the conductive filament vicinity, leading to more efficient switching behaviour.

Figure 5.7. (a) DC IV sweep of gradual forming treatment. (b) Decreasing trend of threshold voltage with the increase of compliance current

Threshold voltage and unidirectional switching

The gradual forming process allows more controllable and localized metallic filament of ionizable Cu electrode into HfO_x dielectric, shown in **Figure 5.3**. A decreasing trend in the threshold voltage (V_{th}) of the device was observed with increasing CC value under DC IV sweep, as shown in **Figure 5.7(a)** and **5.7(b)**. This can be attributed to the amount of Cu atoms residing within dielectric and the effective thickness of the oxide involve during the switching process, which is discussed further in the following segment of the work. After the gradual forming treatment was performed, the device under test was connected to series resistor of 47 k Ω to facilitate subsequent DC cycling and pulse measurements. The forming process was able to successfully achieve stable threshold switching behaviour with more than 100 DC cycles (few selected cycles are shown in **Figure 5.8**).

Figure 5.8. DC IV sweep with series resistor of 47 k Ω connected to the DM device shows the unidirectionality of the DM devices. Ultra-low OFF state current about 100fA and high selectivity of 10^7 was observed under V/2 operating scheme.

The devices tested also preserved unidirectional characteristics after forming and the subsequent cycles. The unidirectionality of the switching is due to the asymmetrical structure of the devices and the nature of the Cu filament growth within the HfO_x. There are two possible growth mechanisms in Cu-based devices, i.e. filamentary growth of Cu atoms from the active copper electrode towards the inert electrode [46], and filamentary growth in opposite direction [47]. This suggests the localized filamentary growth of Cu atoms from the active copper electrode towards the inert electrode [46], rather than the localized filamentary growth in opposite direction during gradual forming of the device [47]. Asymmetrical structures exhibiting bidirectional volatile switching behaviour have only been reported in Ag-based devices [25, 48, 49]. This is consistent with reported literatures on the observation of Ag filament growth [35, 50]. The Ag conductive filaments tend to grow from the inert electrode towards the active electrode. Thus, when the unstable filament formed during volatile switching process, the rupture of the filament does not occur at the furthest Ag atoms position from active electrode, which will leave Ag atoms residue on the inert electrode. After the first few switching cycles of the Ag-based DM devices, there will be sufficient amount of Ag atoms on both side of the dielectric to achieve bidirectional volatile threshold switching behaviour.

This asymmetrical structure will not be able to achieve bidirectional selector behaviour in as fabricated devices.

5.3. TEMPORAL RESPONSE OF DM DEVICES

The underlying mechanism of DM utilises the ionic movement of the active electrode inside dielectric. While it results in an ultra-low OFF current and high selectivity, it also leads to the device having a finite delay and relaxation time during ON and OFF operations respectively. The reported values for delay time vary from about 70 ns up to 70 ms, while the relaxation time varies from 70 ns to 25 s [48]. This temporal response of the device is determined by the specific structure used, applied electric field, and device operating temperature. The focus of this study is on the influence of applied electric field amplitude on the temporal response of the devices.

The time domain measurements were performed by sending higher amplitude voltage pulse to turn ON the DM followed by a “read” pulse of 0.1 V with 20 ms pulse width after 1 μ s interval, while the current flowing through the device was simultaneously measured. **Figure 5.9(b)** shows a complete ON-OFF cycle of the device. The delay time of the device was captured through the time domain measurement from the moment “write” pulse implemented until the observation of abrupt increase in cell current, as depicted in **Figure 5.9(a)**. On the other hand, the finite relaxation time was measured from the timestamp of the removal of the electric field until the sudden drop in cell current was observed, as shown in **Figure 5.9(c)**.

Figure 5.9. Complete ON-OFF cycle of the time domain measurement of the device (b) to capture device delay time (a) and device relaxation time (c).

than R_0 , while resulting in non-volatile switching behaviour if it increases past R_0 . The delay time of the device can be correlated to the nucleation energy barrier height with the following equation [7].

$$\tau_d = \tau_0 \exp \left[\frac{U_N(E)}{kT} \right]. \quad (1)$$

The temperature of the devices was kept constant at room temperature throughout the experiment. τ_0 is defined as the characteristic delay time, which is unique to the active material and dielectric used in the system. $U_N(E)$ is defined as nucleation energy barrier height, that can be expressed as the following,

$$U_N(E) = U_0 E_0 \alpha^2 \frac{3}{2} \frac{d}{V}. \quad (2)$$

V and d are defined as external voltage applied and effective thickness of the dielectric involved during the volatile switching operation. The effective thickness refers to smallest distance between the Cu atoms to the Pt inert electrode across the dielectric. U_0 is defined as the nucleation barrier energy at zero-field, which is unique to the dielectric material in which the metal atoms move within, e.g. HfO_x (0.47 eV) and TiO_x (0.71 eV) [7]. In this work, the dielectric material used was the same, i.e. HfO_x . Thus, under the two forming treatments, the value of parameter U_0 did not change. E_0 is defined in literatures as voltage acceleration factor or characteristic field. This value is independent of the external field or temperature. The acceptable value used in calculation is 1 MV/cm [7], regardless of the type of structures or systems. α is geometric factor of spheroid nucleus. Its value ranges from 0.1 to 0.5 and it was assumed to be 0.5 in the literature because it corresponds to the highest nucleation barrier energy. This value is dependent on the shape of the nucleus. Across different dielectric materials, this value might be different, but within the same dielectric material, this geometric factor can be safely assumed to be identical.

Figure 5.11. Voltage amplitude dependence of device delay time under single and multiple forming treatments. The extracted characteristics time, τ , values are very close in value that indicates both relations come from the same structure of Pt/HfO_x/Cu/Pt.

The focus of this segment of the work is to investigate the effect of external electric field (voltage) on the temporal response of the device. Hence, to simplify the expression in (1), another critical parameter ζ is defined to replace $U_N(E)$.

$$\zeta = U_0 E_0 \alpha^{\frac{3}{2}} \frac{d}{kT} . \quad (3)$$

Thus, equation (1) can be modified into the following,

$$\tau_d = \tau_0 \exp \left[\frac{\zeta}{V} \right] . \quad (4)$$

Equation (4) was used to fit the experimental data obtained for device delay time under different pulse amplitude, shown in **Figure 5.11**. The experimental data fits well under field-induced nucleation theory with a spheroidal-shape metallic filament. The device delay time varies from approximately 5 ms to under 300 μ s for the single run forming treatment, while it ranges from 250 μ s to as fast as 44 μ s for multiple runs forming treatment under different voltage amplitudes. Two parameters could be extracted from the fitting curves, i.e. τ_0 and ζ . τ_0 values extracted from the fitting are ~19 μ s and ~21 μ s for devices underwent single and multiple forming treatments respectively. This is consistent with the two forming conditions

being strictly comparing the same Pt/HfO_x/Cu/Pt structures. On the other hand, the extracted ζ values under two different forming treatments exhibit a significant difference. Based on the field nucleation theory, it is attributed to the change in the effective thickness of the oxide layer involved throughout the nucleation process. It was found that the value of ζ_1 is about 4 times higher than ζ_2 . The devices underwent single forming treatment was assumed to have very small change in the effective thickness from the pristine devices, i.e. $d \sim 10$ nm. The thickness of the HfO_x layer in the pristine devices was experimentally confirmed by TEM images of the multilayer structure, as shown in **Figure 2.7**. This indicates the effective thickness of the oxide was reduced significantly from ~ 10 nm originally to ~ 2.5 nm.

Figure 5.12. Increasing trend of average relaxation time of the device observed under increasing voltage pulse amplitude.

The relaxation time of the device was also investigated under different voltage amplitude in the range of 0.8 V-1.2 V with 10 ms pulse width. From **Figure 5.12**, an increasing trend of relaxation time was observed from about 1 ms to 7 ms with an increase of voltage amplitudes. This trend is potentially due to the size of the spheroidal nucleus generated under different voltage pulses. The higher the amplitude of the pulse, the stronger or larger metallic nucleus dimension formed. Thus, larger metallic nucleus will take a relatively longer time to rupture after the removal of electric field.

Figure 5.13. Metallic filament dissolution process based on Rayleigh instability theorem

This observation is supported by the Rayleigh instability model of the metallic filament rupture inside dielectric [26, 36, 37, 40]. Rayleigh instability suggests that after a removal of an electric field, the metallic filament ruptures into series of spherical nanoclusters through the minimization of total free energy of the system. With our earlier confirmation on spheroidal metallic nucleus growth by looking at the device delay time dependency on voltage pulse amplitude, the initial filament shape before the removal of electric field is taken as spheroid, as shown in **Figure 5.13**. Thus, an analytical model was proposed better understand the evolution of filament morphology. After the removal of electric field, the spheroidal nucleus is divided into N number of nanospheres. This process is assumed to be due to the atomistic re-arrangement of Cu atoms, causing the total volume of the shape to be constant throughout the process. The number of Cu spherical nanoclusters (N) formed can be expressed as,

$$N = \rho \frac{R^3}{D^3}, \quad (5)$$

where ρ is defined as aspect ratio of L/R , in which L and R the initial major and minor axis dimension of the nucleus respectively. D is defined as the average diameter of nanoclusters

formed in the OFF state of the device. This phase transition is driven by the change in total free energy consists of both surface and volume terms. However, with the assumption of constant volume throughout the process, the change in surface free energy will be the only terms contributing into the system. By approximation of very thin metallic filament ($\rho \gg 1$), the change in free energy was estimated in the following:

$$\Delta G \approx \pi \rho R^2 \left(\frac{R}{D} - \frac{\pi}{4} \right). \quad (6)$$

Spontaneous evolution of the filament after the removal of the electric field has been experimentally confirmed in the rupture of Ag filament as per this literature [26]. Thus, the change in free energy should be less than zero and the dimension of the nanoclusters should follow this relation:

$$D > \frac{4}{\pi} R. \quad (7)$$

After applying longitudinal damped sinusoidal perturbation along the major axis of the spheroid, the surface perturbation equation is given by:

$$r(z, t) = R(z) + \delta_2(t) \delta_1(z) \sin kz, \quad (8)$$

where k is the perturbation wavenumber along the z -axis, $\delta_1(z)$ and $\delta_2(t)$ are the spatial and temporal dependence of the perturbation amplitude terms respectively. While the temporal dependence of the perturbation amplitude has been frequently reported [26, 36, 37, 40], the spatial dependence of the amplitude was added due to the spheroidal nature of the filament. If $z = z_N$ was taken as the centre point of one of the nanospheres, the relaxation time of the device is determined by the following equation:

$$\delta_2(\tau_d) = \frac{D - R(z_N)}{\delta_1(z_N) \sin kz_N}. \quad (9)$$

From this equation, it was concluded that the relaxation time is mainly driven by initial filament profile, which is driven by the current or voltage amplitude during the “write” operation.

5.4. RELIABILITY ASPECTS OF DM: DEVICE ENDURANCE AND OBSERVATION OF RANDOM TELEGRAPH SIGNAL (RTS)

The endurance test was performed on the device which underwent multiple runs of forming treatment. By using the correlation of pulse amplitude to the delay time of the device in **Figure 5.11**, the endurance test was performed by sending 200 μs of 0.4 V pulse to turn ON the device and 200 μs of 0.2 V pulse to measure the OFF state of the device. The OFF state of the device was beyond the measurement limit of the equipment. No selectivity degradation was observed after more than 10^4 cycles, shown in **Figure 5.14**. The current distribution of the ON state was observed to be lower than the current flow of 1-Resistor under the same voltage, which means the DM device reached a comparable resistance value to that of the resistor as expected. However, the distribution of the current of DM device was noticeably asymmetrical and has broader variation compared to 1-Resistor current.

Figure 5.14. (a) Actual endurance cycle reading of the device with over 104 ON-OFF cycles. (b) Comparison of current distribution between 1-DM-1Resistor and 1-Resistor under 0.4V voltage amplitude

The origin of the ON-state current variation is explained by the presence of random telegraph signal (RTS) in DM devices. RTS is defined as a random bimodal or multilevel fluctuation of current or voltage during device operation. RTS has been widely reported in memristive devices, especially oxide-based devices utilising the movement of oxygen ions and/or oxygen vacancy-type defects modification for its operations. For these types of devices, the origin of the RTS is believed to be from the reversible movement of electrons between metal electrodes and defects within dielectric layer which is also known as charge trapping and de-trapping activities [54]. However, DM device structure and underlying switching

mechanism mimics those of electrochemical metallization cells (ECM). Thus, the origin of RTS in DM devices is expected to incline towards those similar type of devices in which very few studies have been conducted. In Cu doped $\text{Ge}_{0.3}\text{Se}_{0.7}$ devices, it is suggested that the RTS presence in the device is potentially due to thermally activated movement of Cu species within ionic or redox “double-site traps” [55]. Another report on $\text{Ge}_x\text{Se}_{1-x}$ -based devices showing the observation of RTS in ovonic threshold switching mechanism after the forming process [56]. **Figure 5.15** shows different RTS observed in the DM structure during the ON state of the device under 0.4 V voltage amplitude. Both bimodal and multilevel RTS was observed within the same structure but under different writing cycles. This demonstrated the random and complex nature of the signal. The current readings of the device ON and OFF states during the endurance test were done by taking the average of the data points at the same 10 μs window. Thus, with the time domain variation of RTS, the cell current distribution would be larger than the current distribution of 1-R configuration.

Figure 5.15. RTS captured from different cycles operation of the device during the ON state under 0.4V voltage amplitude. Unstable complex RTS were observed exhibiting multilevel RTS nature.

Further investigation of the RTS behaviour was performed under different applied voltage amplitudes. Two different types of RTS were observed in different voltage regimes, depicted in **Figure 5.15**. The first RTS was more dominantly observed in the low voltage regime (≤ 0.6 V), resulting in random fluctuations throughout the ON state of the device. It was the same type of RTS causing the current distribution broadening during endurance test at 0.4 V. This RTS became less apparent with the increase of the external applied voltage. The second RTS was observed in the high voltage regime of more than 0.6 V. This RTS was more likely to be observed in the early stage of the ON state rather than throughout the ON state. This could be attributed to the competing nature of the two parameters, *i.e.* external electric field and Joule heating. Higher electric field tends to result in stronger filament but at the same time it will generate larger Joule heating effect that might rupture the metal filament. The current reading was observed to stabilize after several current jumps. This could be an indication of a stronger and larger filament had been formed after several ruptures. This property must be taken into consideration in the device implementation as a select device in crossbar array, especially during the “write” operation of the memory cell that requires higher voltage amplitude. The average time constant of this behaviour was above 10 μ s. Thus, it will not affect the memory operation of the devices operating in the faster regime in which most of the reported memristive devices stand.

5.5. SUMMARY

This chapter has discussed a new technique, from device operation point of view, to achieve volatile threshold switching behaviour from pristine Pt/HfO_x/Cu/Pt structure via gradual electroforming treatment. Excellent selectivity of more than 10^7 and extremely low OFF current of 100 fA were maintained with more than 10^4 ON-OFF cycles. An insight into the structural evolution of the DM was discussed by analysing the temporal response of the device under different external electric fields in the frame of the field-induced nucleation theory and Rayleigh instability model. This analysis can be potentially used to extrapolate the device degradation mechanism in which excessive amount of copper species eventually reside within the insulating layer. The reliability aspect of the device considering the current distribution broadening observed during endurance test was also discussed. It was caused by the presence of RTS potentially originated from thermally induced Cu ionic or redox “double-site traps” activities in the structure. The RTS dependence on external applied voltages was also

investigated, which gives an important insight on the time domain behaviour of the device under different voltage regime.

Figure 5.16. The RTS behaviour was observed under different voltage amplitudes. Red-dashed lines are the current peak due to measurement settling time and black-dashed line indicates the current flow of the 1-R configuration (taking current reading of 1.0V as an example).

5.6. REFERENCES

- [1] J. Huang, Y. Tseng, C. Hsu, and T. Hou, "Bipolar Nonlinear Ni/TiO₂/Ni Selector for 1S1R Crossbar Array Applications," *IEEE Electron Device Letters*, vol. 32, no. 10, pp. 1427-1429, 2011.
- [2] J. Huang, T. Yi-Ming, L. Wun-Cheng, H. Chung-Wei, and T. Hou, "One Selector-One Resistor (1S1R) Crossbar Array For High-Density Flexible Memory Applications," in *2011 International Electron Devices Meeting*, 2011, pp. 31.7.1-31.7.4.
- [3] J.-J. Huang, C.-W. Kuo, W.-C. Chang, and T.-H. Hou, "Transition of Stable Rectification to Resistive-Switching in Ti/TiO₂/Pt Oxide Diode," *Applied Physics Letters*, vol. 96, no. 26, p. 262901, 2010.
- [4] G. H. Kim *et al.*, "Schottky Diode with Excellent Performance for Large Integration Density of Crossbar Resistive Memory," *Applied Physics Letters*, vol. 100, no. 21, p. 213508, 2012.
- [5] W. Y. Park *et al.*, "A Pt/TiO₂/Ti Schottky-Type Selection Diode for Alleviating The Sneak Current in Resistance Switching Memory Arrays," *Nanotechnology*, vol. 21, no. 19, p. 195201, 2010/04/19 2010.
- [6] G. Tallarida *et al.*, "Low Temperature Rectifying Junctions for Crossbar Non-Volatile Memory Devices," in *2009 IEEE International Memory Workshop*, 2009, pp. 1-3.
- [7] J. Yoo, J. Park, J. Song, S. Lim, and H. Hwang, "Field-Induced Nucleation in Threshold Switching Characteristics of Electrochemical Metallization Devices," *Applied Physics Letters*, vol. 111, no. 6, p. 063109, 2017.
- [8] A. Kawahara *et al.*, "An 8 Mb Multi-Layered Cross-Point ReRAM Macro with 443 MB/s Write Throughput," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 1, pp. 178-185, 2013.
- [9] W. Lee *et al.*, "High Current Density and Nonlinearity Combination of Selection Device Based on TaO_x/TiO₂/TaO_x Structure for One Selector–One Resistor Arrays," *ACS Nano*, vol. 6, no. 9, pp. 8166-8172, 2012/09/25 2012.
- [10] W. Lee *et al.*, "Varistor-Type Bidirectional Switch ($J_{\max} > 10^7$ A/Cm², Selectivity $\sim 10^4$) for 3D Bipolar Resistive Memory Arrays," in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, pp. 37-38.

- [11] J. Shin *et al.*, "Tio₂-Based Metal-Insulator-Metal Selection Device for Bipolar Resistive Random Access Memory Cross-Point Application," *Journal of Applied Physics*, vol. 109, no. 3, p. 033712, 2011.
- [12] J. Woo, D. Lee, E. Cha, S. Lee, S. Park, and H. Hwang, "Multilayer-Oxide-Based Bidirectional Cell Selector Device for Cross-Point Resistive Memory Applications," *Applied Physics Letters*, vol. 103, no. 20, p. 202113, 2013.
- [13] J. Woo *et al.*, "Multi-Layer Tunnel Barrier (Ta₂O₅/TaO_x/Tio₂) Engineering for Bipolar RRAM Selector Applications," in *2013 Symposium on VLSI Technology*, 2013, pp. T168-T169.
- [14] G. W. Burr *et al.*, "Large-Scale (512kbit) Integration of Multilayer-Ready Access-Devices Based on Mixed-Ionic-Electronic-Conduction (Miec) at 100% Yield," in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, pp. 41-42.
- [15] K. Gopalakrishnan *et al.*, "Highly-Scalable Novel Access Device Based on Mixed Ionic Electronic Conduction (Miec) Materials for High Density Phase Change Memory (PCM) Arrays," in *2010 Symposium on VLSI Technology*, 2010, pp. 205-206.
- [16] R. S. Shenoy *et al.*, "Endurance and Scaling Trends of Novel Access-Devices for Multi-Layer Crosspoint-Memory Based on Mixed-Ionic-Electronic-Conduction (Miec) Materials," in *2011 Symposium on VLSI Technology - Digest of Technical Papers*, 2011, pp. 94-95.
- [17] K. Virwani *et al.*, "Sub-30nm Scaling and High-Speed Operation of Fully-Confined Access-Devices for 3d Crosspoint Memory Based on Mixed-Ionic-Electronic-Conduction (Miec) Materials," in *2012 International Electron Devices Meeting*, 2012, pp. 2.7.1-2.7.4.
- [18] M. Anbarasu, M. Wimmer, G. Bruns, M. Salinga, and M. Wuttig, "Nanosecond Threshold Switching of GeTe₆ Cells and Their Potential as Selector Devices," *Applied Physics Letters*, vol. 100, no. 14, p. 143505, 2012.
- [19] K. DerChang *et al.*, "A Stackable Cross Point Phase Change Memory," in *2009 IEEE International Electron Devices Meeting (IEDM)*, 2009, pp. 1-4.

[20] M. Lee *et al.*, "Highly-Scalable Threshold Switching Select Device Based on Chalcogenide Glasses for 3D Nanoscaled Memory Arrays," in *2012 International Electron Devices Meeting*, 2012, pp. 2.6.1-2.6.3.

[21] S. Kim *et al.*, "Ultrathin (<10nm) Nb₂O₅/NbO₂ Hybrid Memory with Both Memory and Selector Characteristics for High Density 3D Vertically Stackable RRAM Applications," in *2012 Symposium on VLSI Technology (VLSIT)*, 2012, pp. 155-156.

[22] X. Liu *et al.*, "Diode-Less Bilayer Oxide (WO_x-NbO_x) Device for Cross-Point Resistive Memory Applications," *Nanotechnology*, vol. 22, no. 47, p. 475702, 2011/11/04 2011.

[23] M. Son *et al.*, "Excellent Selector Characteristics of Nanoscale VO₂ for High-Density Bipolar ReRAM Applications," *IEEE Electron Device Letters*, vol. 32, no. 11, pp. 1579-1581, 2011.

[24] M. Son *et al.*, "Self-Selective Characteristics of Nanoscale VO Devices for High-Density ReRAM Applications," *IEEE Electron Device Letters*, vol. 33, no. 5, pp. 718-720, 2012.

[25] A. Bricalli, E. Ambrosi, M. Laudato, M. Maestro, R. Rodriguez, and D. Ielmini, "SiO_x-Based Resistive Switching Memory (RRAM) for Crossbar Storage/Select Elements with High On/Off Ratio," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 4.3.1-4.3.4.

[26] C.-P. Hsiung *et al.*, "Formation and Instability of Silver Nanofilament in Ag-Based Programmable Metallization Cells," *ACS Nano*, vol. 4, no. 9, pp. 5414-5420, 2010/09/28 2010.

[27] H. Jiang *et al.*, "A Novel True Random Number Generator Based on A Stochastic Diffusive Memristor," *Nature Communications*, vol. 8, no. 1, p. 882, 2017/10/12 2017.

[28] S. La Barbera, D. Vuillaume, and F. Alibart, "Filamentary Switching: Synaptic Plasticity through Device Volatility," *ACS Nano*, vol. 9, no. 1, pp. 941-949, 2015/01/27 2015.

[29] T. Ohno, T. Hasegawa, T. Tsuruoka, K. Terabe, J. K. Gimzewski, and M. Aono, "Short-Term Plasticity and Long-Term Potentiation Mimicked in Single Inorganic Synapses," *Nature Materials*, vol. 10, p. 591, 06/26/online 2011.

- [30] N. Shukla *et al.*, "Ag/HfO₂ Based Threshold Switch with Extreme Non-Linearity for Unipolar Cross-Point Memory and Steep-Slope Phase-FETs," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 34.6.1-34.6.4.
- [31] J. Song *et al.*, "Monolithic Integration of AgTe/TiO₂ Based Threshold Switching Device with TiN Liner for Steep Slope Field-Effect Transistors," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 25.3.1-25.3.4.
- [32] J. Song *et al.*, "Bidirectional Threshold Switching in Engineered Multilayer (Cu₂O/Ag:Cu₂O/Cu₂O) Stack for Cross-Point Selector Application," *Applied Physics Letters*, vol. 107, no. 11, p. 113504, 2015.
- [33] J. Song *et al.*, "Steep Slope Field-Effect Transistors With Ag/TiO₂-Based Threshold Switching Device," *IEEE Electron Device Letters*, vol. 37, no. 7, pp. 932-934, 2016.
- [34] J. Song, J. Woo, A. Prakash, D. Lee, and H. Hwang, "Threshold Selector with High Selectivity and Steep Slope for Cross-Point Memory Array," *IEEE Electron Device Letters*, vol. 36, no. 7, pp. 681-683, 2015.
- [35] H. Sun *et al.*, "Direct Observation of Conversion Between Threshold Switching and Memory Switching Induced by Conductive Filament Morphology," *Advanced Functional Materials*, vol. 24, no. 36, pp. 5679-5686, 2014.
- [36] J. van den Hurk, E. Linn, H. Zhang, R. Waser, and I. Valov, "Volatile Resistance States in Electrochemical Metallization Cells Enabling Non-Destructive Readout of Complementary Resistive Switches," *Nanotechnology*, vol. 25, no. 42, p. 425202, Oct 24 2014.
- [37] Z. Wang *et al.*, "Memristors with Diffusive Dynamics as Synaptic Emulators for Neuromorphic Computing," *Nature Materials*, Article vol. 16, p. 101, 09/26/online 2016.
- [38] J. Yoo, J. Woo, J. Song, and H. Hwang, "Threshold Switching Behavior of Ag-Si Based Selector Device and Hydrogen Doping Effect on Its Characteristics," *AIP Advances*, vol. 5, no. 12, p. 127221, 2015.
- [39] M. Yu *et al.*, "Self-Selection Effects and Modulation of TaO_x Resistive Switching Random Access Memory with Bottom Electrode of Highly Doped Si," *Journal of Applied Physics*, vol. 119, no. 19, p. 195302, 2016.

- [40] X. Zhao, H. Xu, Z. Wang, L. Zhang, J. Ma, and Y. Liu, "Nonvolatile/Volatile Behaviors and Quantized Conductance Observed in Resistive Switching Memory Based on Amorphous Carbon," *Carbon*, vol. 91, pp. 38-44, 2015.
- [41] W. Chen, H. J. Barnaby, and M. N. Kozicki, "Volatile and Non-Volatile Switching in Cu-SiO₂ Programmable Metallization Cells," *IEEE Electron Device Letters*, vol. 37, no. 5, pp. 580-583, 2016.
- [42] S. Lim, J. Yoo, J. Song, J. Woo, J. Park, and H. Hwang, "Excellent Threshold Switching Device ($I_{\text{off}} \sim 1$ pA) with Atom-Scale Metal Filament for Steep Slope (< 5 mV/Dec), Ultra Low Voltage ($V_{\text{dd}} = 0.25$ V) FET Applications," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 34.7.1-37.7.4.
- [43] T. Liu, M. Verma, Y. Kang, and M. Orlowski, "Volatile Resistive Switching in Cu/TaO_x/δ-Cu/Pt Devices," *Applied Physics Letters*, vol. 101, no. 7, p. 073510, 2012.
- [44] Q. Luo *et al.*, "Cu BEOL Compatible Selector with High Selectivity ($>10^7$), Extremely Low Off-Current ($\sim$ pA) and High Endurance ($>10^{10}$)," in *2015 IEEE International Electron Devices Meeting (IEDM)*, 2015, pp. 10.4.1-10.4.4.
- [45] J. Woo, D. Lee, E. Cha, S. Lee, S. Park, and H. Hwang, "Control of Cu Conductive Filament in Complementary Atom Switch for Cross-Point Selector Device Application," *IEEE Electron Device Letters*, vol. 35, no. 1, pp. 60-62, 2014.
- [46] Q. Liu *et al.*, "Real-Time Observation on Dynamic Growth/Dissolution of Conductive Filaments in Oxide-Electrolyte-Based ReRAM," *Advanced Materials*, vol. 24, no. 14, pp. 1844-1849, 2012.
- [47] W. A. Hubbard *et al.*, "Nanofilament Formation and Regeneration During Cu/Al₂O₃ Resistive Memory Switching," *Nano Letters*, vol. 15, no. 6, pp. 3983-3987, 2015/06/10 2015.
- [48] Z. Wang *et al.*, "Threshold Switching of Ag or Cu in Dielectrics: Materials, Mechanism, and Applications," *Advanced Functional Materials*, vol. 28, no. 6, p. 1704862, 2018.
- [49] G. Du, C. Wang, H. Li, Q. Mao, and Z. Ji, "Bidirectional threshold switching characteristics in Ag/ZrO₂/Pt electrochemical metallization cells," *AIP Advances*, vol. 6, no. 8, p. 085316, 2016.

- [50] Y. Yang, P. Gao, S. Gaba, T. Chang, X. Pan, and W. Lu, "Observation of Conducting Filament Growth in Nanoscale Resistive Memories," *Nature Communications*, Article vol. 3, p. 732, 03/13/online 2012.
- [51] V. G. Karpov, Y. A. Kryukov, S. D. Savransky, and I. V. Karpov, "Nucleation Switching in Phase Change Memory," *Applied Physics Letters*, vol. 90, no. 12, p. 123504, 2007.
- [52] M. Nardone, V. G. Karpov, D. C. S. Jackson, and I. V. Karpov, "A Unified Model of Nucleation Switching," *Applied Physics Letters*, vol. 94, no. 10, p. 103509, 2009.
- [53] A. B. Pevtsov, A. V. Medvedev, D. A. Kurdyukov, N. D. Il'inskaya, V. G. Golubev, and V. G. Karpov, "Evidence of Field-Induced Nucleation Switching in Opal: VO₂ Composites And VO₂ Films," *Physical Review B*, vol. 85, no. 2, p. 024110, 2012.
- [54] Z. Chai *et al.*, "RTN-Based Defect Tracking Technique: Experimentally Probing The Spatial and Energy Profile of the Critical Filament Region and Its Correlation with HfO₂ RRAM Switching Operation and Failure Mechanism," in *2016 IEEE Symposium on VLSI Technology*, 2016, pp. 1-2.
- [55] R. Soni *et al.*, "Probing Cu Doped Ge_{0.3}Se_{0.7} Based Resistance Switching Memory Devices with Random Telegraph Noise," *Journal of Applied Physics*, vol. 107, no. 2, p. 024517, 2010.
- [56] Z. Chai *et al.*, "RTN in Ge_xSe_{1-x} OTS Selector Devices," *Microelectronic Engineering*, vol. 215, p. 110990, 2019.

CHAPTER 6

Conclusion and Future Work

This chapter concludes the work has been discussed in this thesis. Some of the research contributions have been made are highlighted. In addition, some future research work and investigations are suggested with some preliminary results. It includes the device design considerations in achieving bidirectional threshold switching behaviour, ISIR analog synaptic device integration, and Random Telegraph Signal characterization.

6.1. CONCLUSION

Anion-based memristive devices based on Pt/HfO_x/Ti system have been developed using industry-friendly process of UV-lithography patterning and magnetron sputtering deposition. The basic device characteristics, e.g., operating voltages, conductance range, and switching behaviour (abrupt or progressive), can be tuned under different electroforming treatments, i.e., low current forming of 200 μ A (LCF) and self-compliance forming (SCF). The LCF devices exhibited excellent progressive SET and RESET process that is desirable, especially for analog synaptic device applications. The origin of the excellent progressive SET and RESET behaviour of the LCF devices were thoroughly analysed through the dominant conduction mechanisms involved during different conductance states of the devices. Based on the experimentally fitted model, the LCF devices operate under trap-controlled space-charge-limit switching mechanism driven by the oxygen vacancies modulation at HfO_x/Ti interface. Due to the compliance current required during the electroforming process, the implementation of LCF devices on the array level would require a series transistor for each of the memristive element to precisely control the current flowing through the device. Hence, the LCF devices are only suitable for relatively smaller synaptic array density due to the scalability bottleneck of the integrated CMOS transistor.

In order to fully exploit the two-terminal nature of the memristive devices, the requirement of CC to limit the current flowing through the devices during the switching processes, i.e., forming and SET, must be removed altogether. With the presence of TiO_x layer at the interface of the HfO_x and Ti layer induced during the thin film deposition and/or the forming step of the device, Pt/HfO_x/Ti devices can be operated under compliance free forming mode (SCF). The SCF resulted in a mixed of abrupt and progressive switching behaviour during SET and RESET operation as well as higher operating conductance values with higher voltage and current required to switch the devices, arising a challenge in the implementation of the multilevel conductance state capability of the device. However, with optimized programming pulse parameters and conductance range, the excellent progressive switching characteristic of the devices could be retained. The switching mechanism involved in the switching of the SCF devices is potentially originated from the combination of oxygen vacancy defects modulation near the Pt inert electrode and HfO_x/Ti interface. Each mechanism is more dominant than the other under different conductance range. Thus, the multilevel switching behaviour of the devices could be further improved under optimized conductance range. The optimized SCF device operating parameters were implemented to characterize the synaptic

behaviour of the devices towards realizing in-situ on-chip learning capability of ANNs. The trade-off between the dynamic ratio and the asymmetric nonlinearity factor of the devices as well as their endurance capability were thoroughly investigated. The two device operating modes enable wide range of applications for the same device structure under different potential array architectures, i.e., 1T1R, 1TnR, and 1S1R. With the well control device CC with the series transistor element, the LCF mode of the device can be utilized to promote the analog behaviour of the devices with 1T1R and 1TnR architecture. On the other hand, compliance free mode of the SCF devices can be potentially implemented in transistor-less 1S1R integration enabling the real crossbar array architecture.

In Pt/HfO_x/Ti memristive devices, Ti electrode plays a crucial role in determining the physical mechanisms responsible to the conductance switching behaviour of the devices. Thus, changing this oxygen reservoir electrode to an electrochemically active material such as Cu and Ag can significantly alter the switching property of the devices. Cation-based diffusive memristor (DM) was developed using Pt/HfO_x/Cu structure. Volatile and non-volatile switching behaviour were observed in the pristine devices under different CC regimes. A direct implementation of high CC level tended to result in non-volatile switching behaviour on the devices, while gradual electroforming treatment enabled the volatile switching behaviour of the devices even at high CC values. An excellent selectivity of more than 10^7 and extremely low OFF current of 100 fA were successfully achieved, which are desirable for its implementation as a select device in 1S1R crossbar array integration. An insight into the structural evolution of the DM was gained through the analysis of the temporal response of the DM under external electric field. The field-induced nucleation theory and Rayleigh instability theoretical framework were used to explain the finite switching delay and relaxation time observed during the device operation. This analysis can be potentially used to extrapolate the device degradation mechanism in which excessive amount of copper species eventually reside within the insulating layer. The reliability aspect of the device in terms of endurance capability and random telegraph noise present in the structure were also characterized. The excellent performance of the devices was maintained for more than 10^4 ON-OFF cycles.

6.2. FUTURE WORK

6.2.1. Bidirectional Threshold Switching

The Pt/HfO_x/Cu diffusive memristor discussed in chapter 5 is based on infinite cations source of Cu electrode. The growth of the Cu filament always begins from the Cu electrode towards the inert electrode, resulting in preserved unidirectional threshold switching behaviour throughout the device operation, unlike Ag-based devices in which the Ag cations tend to get reduced at the inert electrode resulting filament growth in the opposite direction. In order to promote bidirectional threshold switching behaviour on the structure, two different approaches can be implemented, i.e., symmetrical infinite cations and finite cations source design.

Symmetrical infinite cations source can be developed by sandwiching the HfO_x layer between two Cu electrodes, as depicted on **Figure 6.1(a)**. Theoretically, with this mirrored structure, the electric field induced nucleation model can still be valid. The performance of this mirrored structure selector can be tuned by the insertion of denser oxide layer right in the centre of the oxide switching layer. It has been demonstrated that the presence of this oxide insertion layer can significantly alter the ON voltage of the structure [1]. This approach can be immediately implemented by adding oxygen into the sputtering environment to increase the oxygen content at the centre of HfO_x.

On the other hand, the finite cations source, depicted in **Figure 6.1(b)**, is realized by doping the dielectric material with the neutral cation atom. This finite cation devices have also been used to achieve analog synaptic device behaviour. Thus, it requires a rigorous optimization of the doping percentage to map out the regime in which the devices exhibit volatile and non-volatile switching behaviour [2-4].

Figure 6.1. The schematic of the proposed symmetrical infinite cations source device (a) and finite cation source device (b) to achieve bidirectional threshold switching behaviour

6.2.2. 1S1R Analog Synaptic Device

The ultimate goal of the development of the two-terminal analog synaptic device and its compatible select device is to achieve analog 1S1R synapse implementation. This will enable the development of the highly scalable analog synaptic array. 1S1R devices have been widely demonstrated for digital memristive devices in its application as high density storage class memory. However, the impact of the selector integration with an analog memristive element have rarely been discussed.

Several simulation works have been performed in order to predict how the select device will influence the conductance response of the synaptic device for the inference and online training applications [5, 6]. The simulation was performed on two different selectors, i.e., exponential-IV selector (ES) and threshold switching selector (TS). The read-out current value was found to be vulnerable to the IR drop due to the high wire resistance. This is especially critical in the inference stage of the network. 1ES-1R system is affected more as compared to the 1TS-1R. This is due to the non-linear IV response of the ES causing the slightest deviation of read voltage across the 1ES-1R cell will lead to significant drop in the conductance. On the other hand, the 1TS-1R system is not significantly affected by the IR drop. This is due to the abrupt threshold behaviour of the selector allowing the memristive element to take over the overall IV response when the TS is switched ON. Thus, the presence of the TS will not affect the read-out current value. For online training application, the conductance response of 1ES-1R and 1TS-1R have also been simulated. Based on the simulation, 1ES-1R performed better than 1TS-1R due to the more gradual voltage shift between the ES and the memristive cell. When the voltage applied across the 1ES-1R cell, part of this voltage drops across the ES and the remaining drops across the memristive element. Under identical pulse train programming, as the conductance of the memristive element increases, the voltage drop across the memristive device will decrease. This allow the system to mimic the non-identical pulse programming scheme. On the other hand, for 1TS-1R system, the non-linearity and symmetry of the conductance response highly depends on the memristive device element. When the TS is switched ON, there will be a sudden increase in the voltage across the memristive device, resulting in higher conductance change for the first few programming pulses in the beginning.

Based, on these simulation results, different type of select device might be required to meet certain target applications. However, further experimental verifications are required to support the claim.

The 1TS-1R analog synaptic device can be investigated with the developed analog memristive devices (in Chapter 3 and 4) and the cation-based diffusive memristor (in Chapter 5). Gradual electroforming process needs to be carefully implemented during the forming of the 1TS-1R cell. The proposed device configuration is depicted in **Figure 6.2**.

Figure 6.2. Proposed 1S1R configuration based on the developed anion and cation-based devices.

6.2.3. Random Telegraph Signal

Random Telegraph Signal (RTS) is defined as random discrete fluctuations in voltage or current signal caused by intrinsic defects presence within dielectric. This type of noise signal is commonly observed in dielectric-based devices in addition to the thermal and flicker noise. The probabilistic nature of RTS makes it one of the most difficult type of noises to deal with. In the context of memristive devices, RTS can be classified into two major classes, i.e., defect perturbation (d-RTS) and electron transport (e-RTS)-induced RTS [7]. The defect perturbation induced RTS tend to occur in relatively higher read voltage amplitude with longer time scale and larger noise amplitude as compared to e-RTS. A simple e-RTS consists of two discrete levels of signal, i.e., bimodal e-RTS, that corresponds to the movement of a single electron

from the electrodes into the trap site within the switching layer and vice versa. However, due to the device underlying switching mechanism, i.e., the intrinsic defects of the dielectric layer are being modulated, many defect sites will be established from the application of external electric field. This can potentially lead to the observation of more complex RTS waveforms, where multi-level RTS originated from superposition of e-RTS from different defect sites as well as a mixed of e-RTS and d-RTS. This complex signal requires more sophisticated tools to analyse, i.e., factorial hidden Markov model (FHMM). FHMM is used to decompose a single multilevel RTS into several bimodal RTS spectrums [8].

Figure 6.3. Schematic origin of e-RTN [9].

The presence of this noise signal during the read operation of the memristive device is an obvious disadvantage from the application viewpoint. This noise will contribute to the broadening of the states' variation as well as introduce “soft errors” during reading operation if the dynamic ratio of between the two nearest states are not sufficiently large, i.e., which is the case for most of the reported analog synaptic devices. Thus, it is of an extremely important reliability aspect of memristor to understand and control. Besides the issues might arise from it, the RTS has been reported to carry a crucial information related to the defects profile within the structure. e-RTS can be used to map out the defects spatial profile within the structure,

while d-RTS with the help of quantum point contact (QPC) model can be used to describe the number of defects at the filament constriction where most of the switching activities occur [7].

For e-RTS alone, there are three different origins of the fluctuation depending on different direction of the electron transport, as depicted in **Figure 6.3**. The first type of electron transport is when the electron jumps into the defect site and jumps back to the same electrode. The second one occurs when the electron originated from first electrode ends up in the second electrode by jumping through one or more defect sites within dielectric. The third one is due to the transfer of the electron from one electrode to the defect site followed by thermionic emission. The different nature of the electron transport results in different average emission and capture dependency of the RTS on the read voltage amplitude, as shown in the figure. The first type of e-RTN governs by classical theory, comprises of the following three equations [9]:

$$\frac{\tau_c}{\tau_e} = e^{(E_T - E_F)/K_B T} \quad (1)$$

$$K_B T \ln \left(\frac{\tau_c}{\tau_e} \right) = \varphi_0 - \left[(E_{C,ox} - E_T) + \left| q \cdot \frac{X_T}{T_{ox}} \cdot V_{ox} \right| \right] \quad (2)$$

$$\frac{X_T}{T_{ox}} = \frac{K_B T}{q} \cdot \frac{\partial}{\partial V} \left[\ln \left(\frac{\tau_c}{\tau_e} \right) \right] \quad (3)$$

where E_T and E_F is the oxide trap and fermi energy level respectively. With the help of energy band diagram, (1) can be modified into (2), having the difference of electrode work function and the oxide electron affinity (φ_0), conduction band edge of the HfO_{2-x} ($E_{C,ox}$), ratio of trap location (X_T), and the thickness of tunneling current region in the oxide (T_{ox}) with voltage drop across it as V_{ox} . Equation (3) and the linear fit in Fig.7c was used to extract relative defects spatial location. X_T from the reverse and forward bias, is calculated from two opposite electrodes.

Preliminary results of RTS observation in Pt/HfO_x/Ti anion memristive devices is depicted in **Figure 6.4**. RTS is known to be observed at various timescale, ranging from μs to hours average switching time. It might be observed in the form of a simple bimodal discrete levels or relatively more complex waveform, which is highly dependent on external electric field, relative defect position within dielectric, and defect energy profile. Thus, the sampling rate of the setup and the amplitude of the read voltage used were carefully optimized. Most of the RTS in this structure can be captured under 8kHz sampling rate within 1s period for each

bias value. The reading process was designed in such a way that sufficient number of fluctuations recorded for the analysis while not exposing the device with long voltage stress that potentially alter the defects spatial configuration or form new defects. The RTS measurement was done after each writing cycle under different voltage read polarity and amplitude. The presence of bimodal and multilevel RTS can be observed in **Figure 6.4**.

Figure 6.4. Time-domain multilevel RTS observed in the high voltage regime (inset: time-lag plot)

6.3. REFERENCES

- [1] Y. Sun *et al.*, "Performance-Enhancing Selector via Symmetrical Multilayer Design," *Advanced Functional Materials*, vol. 29, no. 13, 2019.
- [2] T. D. Dongale, S. V. Mohite, A. A. Bagade, R. K. Kamat, and K. Y. Rajpure, "Bio-Mimicking the Synaptic Weights, Analog Memory, and Forgetting Effect Using Spray Deposited WO₃ Memristor Device," *Microelectronic Engineering*, vol. 183-184, pp. 12-18, 2017/11/05/ 2017.

- [3] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, "Nanoscale Memristor Device as Synapse in Neuromorphic Systems," *Nano Letters*, vol. 10, no. 4, pp. 1297-1301, 2010/04/14 2010.
- [4] X. Yan *et al.*, "Memristor with Ag-Cluster-Doped TiO₂ Films as Artificial Synapse for Neuroinspired Computing," *Advanced Functional Materials*, vol. 28, no. 1, p. 1705320, 2018.
- [5] J. Woo, X. Peng, and S. Yu, "Design Considerations of Selector Device in Cross-Point RRAM Array for Neuromorphic Computing," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, pp. 1-4.
- [6] J. Woo and S. Yu, "Impact of Selector Devices in Analog RRAM-Based Crossbar Arrays for Inference and Training of Neuromorphic System," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 27, no. 9, pp. 2205-2212, 2019.
- [7] N. Raghavan *et al.*, "Microscopic origin of random telegraph noise fluctuations in aggressively scaled RRAM and its impact on read disturb variability," in *2013 IEEE International Reliability Physics Symposium (IRPS)*, 2013, pp. 5E.3.1-5E.3.7.
- [8] F. M. Puglisi and P. Pavan, "RTN Analysis with FHMM as a Tool For Multi-Trap Characterization in HfO_x RRAM," in *2013 IEEE International Conference of Electron Devices and Solid-state Circuits*, 2013, pp. 1-2.
- [9] Z. Chai *et al.*, "RTN-Based Defect Tracking Technique: Experimentally Probing the Spatial and Energy Profile of the Critical Filament Region and Its Correlation with HfO₂ RRAM Switching Operation and Failure Mechanism," in *2016 IEEE Symposium on VLSI Technology*, 2016, pp. 1-2.