8-2019

Device Modeling and Circuit Design of Neuromorphic Memory Structures

Sherif Amer
University of Tennessee, samer1@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Sherif Amer entitled "Device Modeling and Circuit Design of Neuromorphic Memory Structures." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Computer Engineering.

Garrett Rose, Major Professor

We have read this dissertation and recommend its acceptance:

Garrett Rose, Benjamin J. Blalock, Nicole McFarlane, Andy Sarles

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Device Modeling and Circuit Design of Neuromorphic Memory Structures

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Sherif Hassanein Amer

August 2019

This Dissertation is dedicated to my family

Acknowledgments

I would like to express my sincere gratitude to my advisor Dr. Garrett Rose for his support and guidance throughout my years working for his group. His continuous support and advice has helped me grow significantly as a researcher.

Special thanks to Dr. Benjamin Blalock, Dr. Nicole McFarlane and Dr. Andy Sarles for their insights, specially during my proposal, which helped me improve my dissertation.

I would also like to thank my lab mates in SENECA group: Sagar Varma, Musabir Adnan, Md Majumder, Gangotree Chakma, Mesbah Uddin, Ryan Weiss and Sam Brown for the good time we spent together as colleagues and friends and making my journey here at the University of Tennessee gratifying and enjoyable.

Special Thanks for Dr. Sakib Hasan for the valuable discussions we had on device modeling and the insights he provided

Special thanks to Dr. Karsten Beckmann, Maximillian Liehr and Dr. Natheniel Cady at SUNY Polytechnic for the valuable insights and discussions about the Hafnium Oxide memristors and the device data they provided to us.

I would also like to express my appreciation to my mother Mona for her continuous emotional support, my uncle Ahmed for his advice and career insights and my father Hassanein for introducing me to academic research and educating me on how to write solid academic publications.

Finally, I would like to express my earnest and sincere gratitude to my grandfather Hamdi ElBanbi. His advice throughout the years has helped me grow in all respects and has shaped my personality to become who I am today.

Abstract

The downscaling of CMOS technology and the benefits gleaned thereof have made it the cornerstone of the semiconductor industry for many years. As the technology reaches its fundamental physical limits, however, CMOS is expected to run out of steam instigating the exploration of new nanoelectronic devices. Memristors have emerged as promising candidates for future computing paradigms, specifically, memory arrays and neuromorphic circuits. Towards this end, this dissertation will explore the use of two memristive devices, namely, Transition Metal Oxide (TMO) devices and Insulator Metal Transition (IMT) devices in constructing neuromorphic circuits.

A compact model for TMO devices is first proposed and verified against experimental data. The proposed model, unlike most of the other models present in the literature, leverages the instantaneous resistance of the device as the state variable which facilitates parameter extraction. In addition, a model for the forming voltage of TMO devices is developed and verified against experimental data and Monte Carlo simulations. Impact of the device geometry and material characteristics of the TMO device on the forming voltage is investigated and techniques for reducing the forming voltage are proposed. The use of TMOs in syanptic arrays is then explored and a multi-driver write scheme is proposed that improves their performance. The proposed technique enhances voltage delivery across the selected cells via suppressing the effective line resistance and leakage current paths, thus, improving the performance of the crossbar array.

An IMT compact model is also developed and verified against experiemntal data and electro-thermal device simulations. The proposed model describes the device as a memristive system with the temperature being the state variable, thus, capturing the temperature dependent resistive switching of the IMT device in a compact form suitable for SPICE

implementation. An IMT based Integrate-And-Fire neuron is then proposed. The IMT neuron leverages the temperature dynamics of the device to deliver the functionality of the neuron. The proposed IMT neuron is more compact than its CMOS counterparts as it alleviates the need for complex CMOS circuitry. Impact of the IMT device parameters on the neuron's performance is then studied and design considerations are provided.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

With Moore's law plateauing, new nanoelectronic devices are explored that are capable of perpetuating the gains previously gleaned by CMOS downscaling by either replacing CMOS in some applications or complementing them in others. One particular application where memristive devices offer a significant potential is neuromorphic circuits due to their miniature feature size and incremental resistance programming. To this end, the work considered here explores memristive devices and their application in building neuromorphic memory architectures. Device models and circuit techniques are proposed in this work to improve the robustness and performance of memristive neuromorphic circuits.

1.1 Memristive Devices

Memristive devices are amongst the novel devices that have been heavily explored over the past few years. Their existence was theoretically predicted by Leon Chua in 1971 [10] but hadn't been physically realized until 2008 when HP announced the first manufactured memristor based on Titanium dioxide TiO_2 [2]. Leon Chua argues, however, that any two terminal device exhibiting a pinched hysteresis loop in the I-V plane that passes through the origin is a memristor and that memristive dynamics are not exclusive to a particular material or process but rather it is a system theory [11, 12]. A memristor can be viewed as an electrically programmable resistor where the resistance is modulated based on the applied voltage.

Memristive devices can be broadly classified into two categories: volatile memristors and nonvolatile memristors. Volatile memristors, as their name suggests, do not possess a long term memory and lose their state (memory) as the applied voltage is removed unlike nonvolatile memristors which can, ideally, maintain their state indefinitely.

Nonvolatile memistors are typically employed in two applications, namely: memory and neuromorphic systems. In memory architectures, memristors are used as resistive binary switches where the logic value of the cell is encoded in the devices resistance. For example, a High Resistance State (HRS) may represent a logic 0 while a Low Resistance State (LRS) may represent a logic 1. On the other hand, the full resistance range of the device (i.e. analog programming) is leveraged in neuromorphic applications to represent synaptic weights.

Similar to nonvolatile memristors, volatile memristors are also employed in memory and neuromorphic systems but deliver different functionality. In memory arrays, volatile memristors are typically used as selector devices to provide high cell non-linearity and suppress sneak path currents. In neuromorphic systems, volatile memristors can be used as neurons.

1.2 Memristive Circuits

Memristors are often organized in a nanoelectronic structure known as the crossbar array shown in Fig 1.1. They are integrated at the intersection of two orthogonal wires and, typically, assume the size of a via as shown in Fig 1.2. This enables high level of integration density which is critical in all of the aforementioned applications.

Accessing memristive devices in the crossbar array requires robust and energy efficient read and write techniques. Ideally, accessing a memristor cell for either read or write operation involves biasing the row and column connected to this cell while floating all other lines. This, however, may lead to unwanted current flow through the unselected cells, since the crossbar array is a passive structure, which may result in erroneous operation. To address this limitation, intelligent write and read schemes have been proposed. In the write operation, the V/2 and the V/3 bias schemes [13, 14, 15] have been proposed where the unselected cells are biased with half or third the write voltage, respectively.

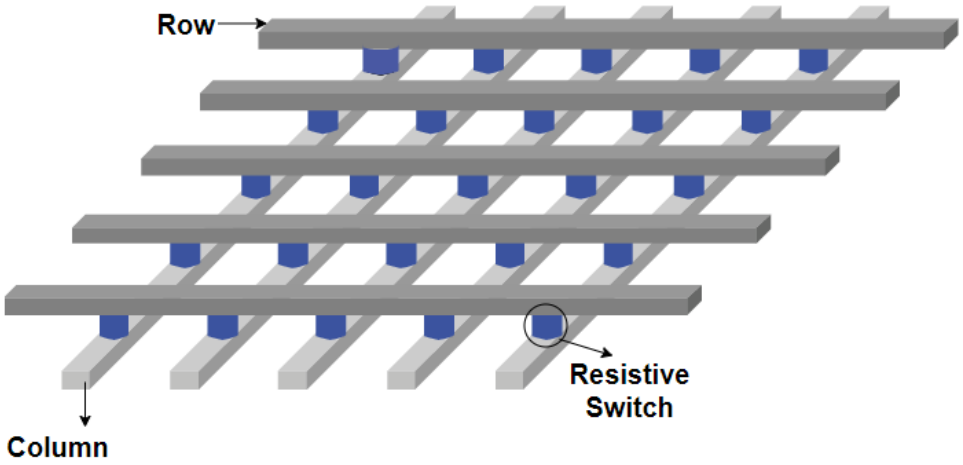

Figure 1.1: Crossbar array structure adapted from [1]. Two sets of orthogonal wires with the memristor device integrated at the intersection of each row and column.

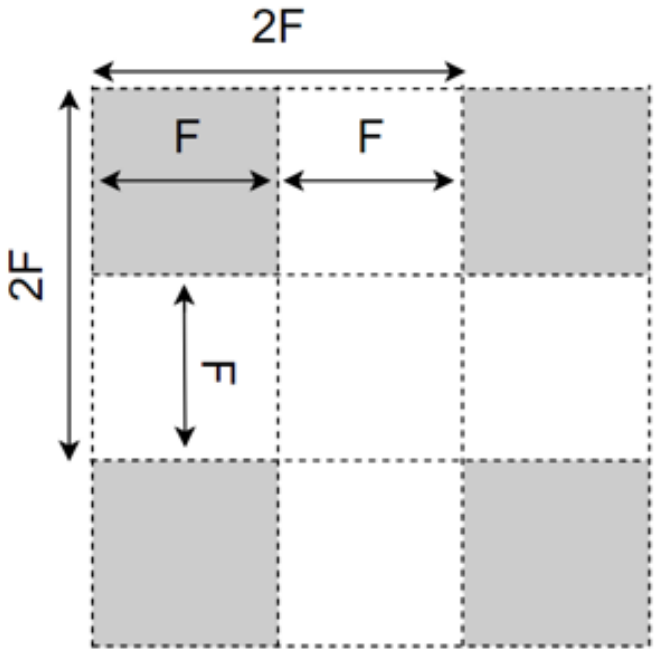

Figure 1.2: Demonstration of the $4F^2$ area occupancy of the TMO device.

In the read operation, techniques such as the ones proposed in [16] have been adopted to limit sneak path currents flowing in the crossbar array, thus, boosting the read margin.

1.3 Neuromporhic Circuits from Memristive Devices

The small feature size and analog resistance programming of memristive devices make them promising candidates for neuromorphic applications. A typical neuromorphic system consists of two components: synapses and neurons. Memristive synapses can be integrated in a crossbar array structure and carry synaptic weights. Neurons are processing elements that fire should the synaptic input cross (inputs multiplied by their respective synaptic weights) a certain threshold. In analogy to conventional Von Neumann architecures, the synaptic array represents the memory while the neurons represent the processors. However, unlike in Von Neumann architectures, the processors in neuromorphic systems are distributed which alleviates the memory wall bottle neck found in modern micro-architectures [17] as shown in Fig 1.3.

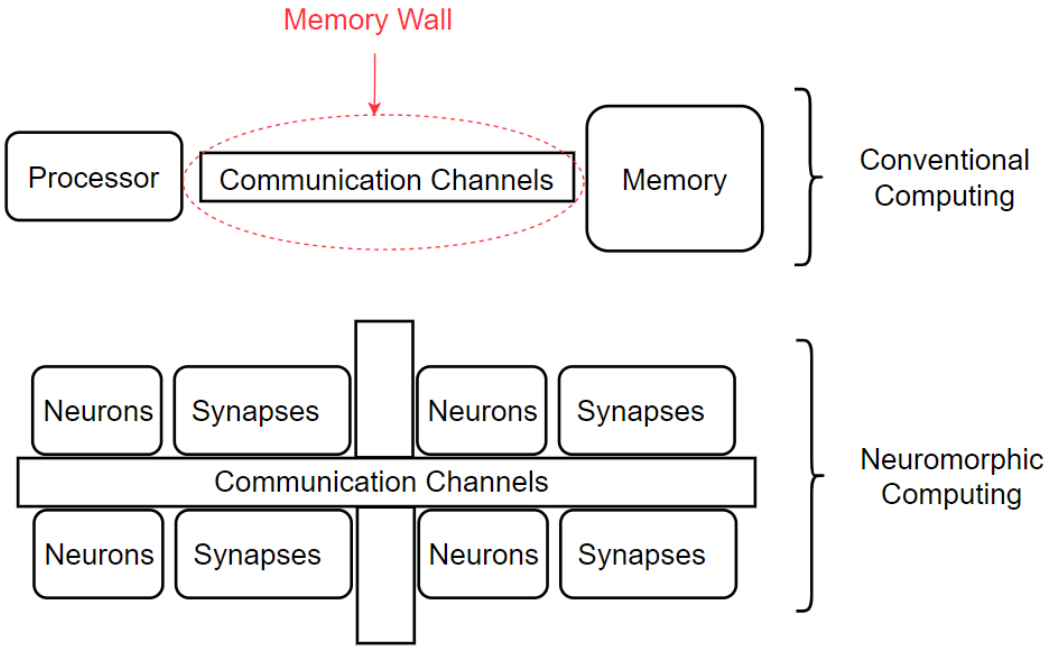

Figure 1.3: Comparison between conventional computing and neuromorphic computing illustrating the memory wall problem.

Currently, state-of-the-art CMOS-based neuromorphic systems such as IBM's TrueNorth [18], MIT's Eyeriss [19] and other Convolutional Neural Network (CNN) accelerators [20, 21, 22] use SRAM cells as synaptic elements. This, however, leads to increased power consumption and area occupancy since each analog synaptic weight is represented by multiple bits. This limitation gave rise to memristive neuromorphic systems where the memristive device can be leveraged as the synaptic element[17]. The synaptic weight can be, thus, programmed in the device's resistance reducing the dot product operation to simple Ohm's law. However, designing robust memristive neuromorphic circuits still require accurate modeling of memristive synaptic elements for circuit design and simulation as well as robust write schemes for programming these devices. On the neuron front, more compact neurons that the currently used CMOS neurons may be required to further improve the performance of the neuromorphic circuit which entice the exploration of new devices that can deliver the neuron's functionality with less hardware [23].

1.4 Research Goal and Scope

The goal of this work is to explore the application of non-volatile and volatile memristors in the construction of neuromorphic circuits. Two devices are explored, namey: Transition Metal Oxide (TMO) devices (non-volatile memristor) and Insulator Metal Transition (IMT) devices (volatile memristor).

This work follows a bottom-up approach and is mainly concerned with the device and circuit abstractions. On the device front, this work focuses on SPICE level modeling and does not delve deeper into the device physics except when needed such as in the case of electroforming where atomistic level simulations were performed. Even at this level, many of the device details were abstracted and only the characteristics of interest were studied. On the circuit front, all simulations were performed using Spectre circuit simulator from Cadence and the device models used are all fitted to real devices weather BSIM models describing the CMOS part of the system or the memristor models developed by the author.

1.5 Research Contributions

This work develops techniques and methodologies to improve the performance of the memristive neuromorphic circuits. Specifically, device models of the two memristive devices studied in this dissertation are first developed to provide better understanding of their behavior when integrated in neuromorphic circuits. The models are SPICE compatible to enable efficient circuit design and simulation. Circuit techniques to effienctly access/write these devices are also developed.

First, a compact model for TMO memristors is developed. Unlike the previously proposed models that are peculiar to some specific switching mechanism, the proposed model is generic and is based on physically accessible parameters which makes it readily amenable for parameter extraction. The model ensures smoothness across all regions of operation to facilitate convergence during circuit simulation. The Verilog-A model of the TMO device is provided in Appendix A.

TMO memristors often require a one time process known as electroforming. Electroforming requires high forming voltages, higher than the nominal voltages used in state-of-the-art CMOS technology nodes, which severely hampers the compatibility of TMO memristors with the CMOS process. Efforts have been undertaken to reduce the forming voltages to levels that are compatible with CMOS at the device level. Those efforts, however, are mostly experimental studies and do not provide a model that help understand how the different device parameters can affect the forming voltage. Thus, this work investigates the physical mechanisms involved in electroforming at the atomistic level. A Monte Carlo simulation framework and a physical closed form model are developed that identify the key physical and structural parameters that can be varied to lower the forming voltage.

The application of TMO memristors in synaptic arrays is then explored and circuit level issues are analyzed. It is shown that the line resistance and the leakage current paths result in voltage degradation across the selected cell which can severely hamper the performance of synaptic crossbar arrays. A multi-driver write scheme is proposed that improves voltage delivery to the selected cells via reducing the effective line resistance and leakage current paths.

A novel volatile memristor known as Insulator Metal Transition (IMT) device is then studied. IMT devices exhibit temperature controlled resistive switching. Several efforts have been made in modeling IMT devices. However, most compact models are behavioral and do not describe the temperature dynamics of the IMT device. On the other hand, the models that describe the role of temperature in IMT switching are TCAD models and cannot be integrated in SPICE-like simulators. This work proposes an IMT SPICE model verified against experimental data and electrothermal device simulations. The Verilog-A model of the IMT device is provided in Appendix B.

An Integrate And Fire (IAF) neuron is then proposed that leverages the switching dynamics of the IMT device and is simulated using the proposed model. The operating principle is explained and design expressions are derived. Impact of the IMT device parameters on the performance of the neuron is also investigated.

Chapter 2

Background

2.1 Memristors

The memristor concept was first proposed by Leon Chua [10] as the fourth fundamental circuit element along with resistors, capacitors and inductors. Voltage and current, voltage and charge and current and flux are connected via resistance, capacitance and inductance, respectively. Voltage is the time derivative of the flux and, similarly, current is the time derivative of the charge. A missing link between the flux and charge existed, as shown in Fig 2.1, which Chua postulated as the memristor element and can be mathematically described as follows:

$$I = G(w, V)V, \tag{2.1}$$

$$\frac{dw}{dt} = f(w, V), \tag{2.2}$$

where equation (2.2) is the state equation, (2.1) is the output equation and w is the state variable. Chua also argues that any device that exhibits hysteresis in the V-I plane is a memristor as shown in Fig 2.2. He later generalized his theory into memristive systems [11] to encompass a myriad set of other elements that had not been considered memristors before.

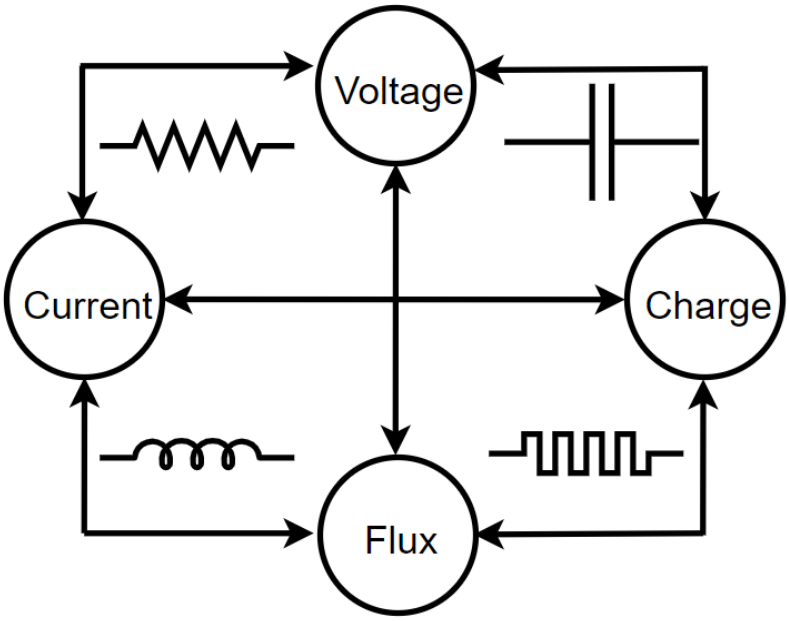

Figure 2.1: Layout of the four fundamental electric quantities

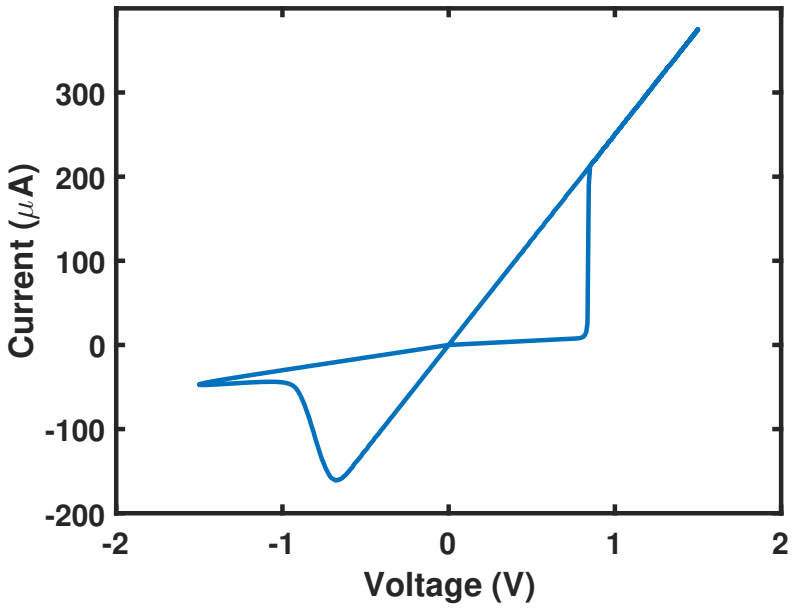

Figure 2.2: Hysteresis in the V-I plane

2.2 Physical Realizations of Memristors

The first memristor to be physically realized was developed by HP in 2008 when they announced the first working memristor based on Titanium dioxide process [2, 24] which exhibits resistive switching based on oxygen vacancy creation and annihilation. Other memristive devices, however, were later proposed based on different properties and/or switching mechanisms. Broadly speaking, memristors can be classified into nonvolatile and volatile memrsitors depending on whether they possess or lack memory.

The three main nonvolatile memristors are: Transition Metal Oxides (TMOs) (also referred to as RRAMs), Phase Change Memory (PCM) and Magnetic RAM (MRAM). The switching mechanism in TMOs is based on the dynamics of oxygen vacancies. Resistive switching in PCM relies on switching between amorphous and crystalline phases [25] while in MRAMs switching relies on device magnetization [26]. Table 2.1 compares the three different nonvolatile memristor devices.

One can readily observe that PCM devices might be best suited for neuromorphic applications for they provide intermediate resistance states. MRAM devices, on the other

Table 2.1: Comparison between nonvolatile memristive devices

Property	PCM	TMO	MRAM
Switching mechanism	crystalline to amorphous transition	Oxygen vacancy creation and annihilation	Spin Transfer Torque (STT)
Resistance range	$10k\Omega$ to $200k\Omega$	$10k\Omega$ to $100k\Omega$	$2k\Omega$ to $10k\Omega$
Resistance ratio	10^3 to 10^6	1 to 1000	1.5 to 3
Write latency	$150ns$	$10ns$ to $50ns$	$2ns$ to $20ns$
Tunability	Intermediate states	Stochastic	Bistable
Endurance	$< 10^8$	10^8 to 10^{12}	$> 10^{12}$

hand, are bistable and are more suitable for memory applications. TMO devices may provide intermediate states yet their switching dynamics are highly stochastic which introduces challenges in circuit applications. Despite these challenges, TMO devices are widely used for they possess low operating voltages [27] which facilitates their integration in standard CMOS processes.

Volatile memristors encompass devices such as Mott devices [28], Insulator Metal Transition (IMT) devices [29], Dipole Induced Bilayer (DIB) devices [30, 31] and Thermistors [11]. The state dynamics of these devices are very similar to nonvolatile ones except that a leakage mechanism is often present which results in the device losing its state (memory) over time. For example, in Thermistors and IMTs, the state dynamics are controlled by the temperature evolution of the device resulting from Joule heating. A leakage mechanism exists, however, induced by conduction and/or convection [32]. This class of devices are also sometimes referred to as diffusive memristors and have been shown to exhibit interesting characteristics that can be leveraged in neuromorphic computing [33, 34]. The IMT in particular has recently shown a significant potential for use in neuromorphic structures [23, 32]. SPICE models, however, are still needed to enable the full exploration of these devices in circuit environment.

2.3 Memristor models

Transition Metal Oxides (TMO) and Insulator Metal Transition (IMT) devices, as mentioned earlier, warrant attention by neuromorphic circuit designers for the interesting dynamics they can deliver. This, however, requires accurate yet computationally efficient SPICE compatible models to facilitate circuit design and simulation. This section will review some of the commonly used models for both devices and address some of the existing challenges.

2.3.1 Tansition Metal Oxide (TMO) models

Transition Metal Oxides in their pristine state possess few or no Oxygen vacancies to enable cycling (i.e. switching or regular operation). An electroforming step [35] is first required

before the device can be used for cycling. This said, this discussion will be divided into two parts: switching models and electroforming models.

TMO Switching Models

Several models have been proposed in the literature since 2008. In [2], the linear ion drift model was proposed in which oxygen vacancies were assumed to undergo a linear drift with the applied electric field. The linear ion drift model was modified later in [36, 37, 38] where an empirical window function was added to account for the non-linear vacancy dynamics near the device boundaries. In [39], the non-linear ion drift model was proposed where an exponential dependence of the vacancy drift velocity on the electric field was assumed. Other works such as [40, 41, 42, 43] proposed models based on the Simmon's barrier tunneling model [44]. Models in [45, 46] are also based on [44] but followed simpler formulation to enable faster and simpler execution in SPICE simulators. In [3, 47], models based on reaction rate equations were proposed where the dynamics of oxygen vacancies were described by Arrhenius law. To the best of the author's knowledge, these models are the closest to the actual physical dynamics taking place within TMOs as agreed by the community.

The stochastic dynamics of Oxygen vacancies in TMOs have posed significant challenges on the device modeling front. Unlike conventional semiconductor devices, memristors are dynamic devices that possess an internal state variable controlling their state evolution. For example, in expressions (2.1) and (2.2) the state variable w controls the conductance (reciprocal of memristance) of the device. Thus, the validity of the device model is dependent on the correct physical interpretation of the state variable controlling the switching and its relationship to the external excitation as well as the relationship between the state variable and the device's resistance. For example, the first TMO model proposed in [2] was described as in expressions (2.3) and (2.4):

$$V = (R_{on}\frac{w(t)}{D} + R_{off}(1 - \frac{w(t)}{D}))I, \tag{2.3}$$

$$\frac{dw}{dt} = \mu_v \frac{R_{on}}{D} I, \tag{2.4}$$

Fig 2.3 depicts a visualization of the physical picture posed by the HP model. This model, as alluded to before, assumed a linear drift of the Oxygen vacancies under the applied electric field. The state variable w is chosen to be the length of the oxygen vacancy rich region and its rate of change (i.e. velocity of the vacancies) is linearly dependent on the applied electric field such that $v = \mu E$. The relationship between the the state variable and the device's resistance is a weighted average of the oxygen rich and oxygen deficient regions as shown in expression (2.3). This physical interpretation was later proven to be inaccurate which, accordingly, limited the validity of the model.

The models in [3, 4] presented the most accurate physical interpretation known to date for TMO memristors based on Arrhenius law. These models use the reaction rate equation to model oxygen vacancy dynamics. The challenge, however, rests in choosing a state variable that accurately represents resistive switching!

Resistive switching in TMOs occur due to filament formation. A filament is a chain of oxygen vacancies. The reaction rate equation only describes the dynamics of Oxygen vacancies, and not the filaments. Fig 2.4 depicts the structure of Oxygen vacancy filaments in TMOs. In order to model filament formation accurately given the reaction rate equation, a Monte Carlo simulation framework needs to be employed [48, 49]. SPICE models, however, need to be more compact and closed form. Hence, an approximation to the geometric structure of the filament is usually employed. For example, the model in [3] assumed that resistive switching occurs due to creation and annihilation of oxygen vacancies near the electrode and, accordingly, the state variable was chosen to be the gap between the tip of the filament and the electrode. On the other hand, the model in [4] assumed that the filament already shunts both electrodes and that resistive switching occurs due to change in the filament size and, therefore, the cross sectional area of the filament was chosen as the state variable. Fig 2.5 depicts a visual representation of both models.

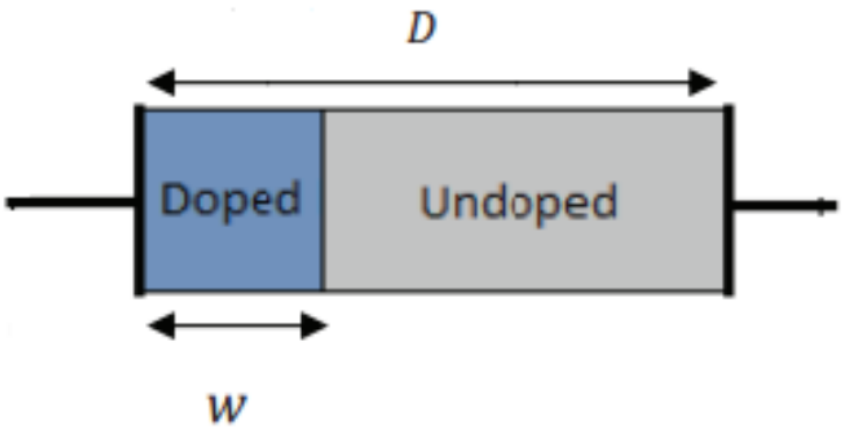

Figure 2.3: Visualization of the Oxygen vacancy dynamics in the linear ion drift model. Figure is adapted from [2].

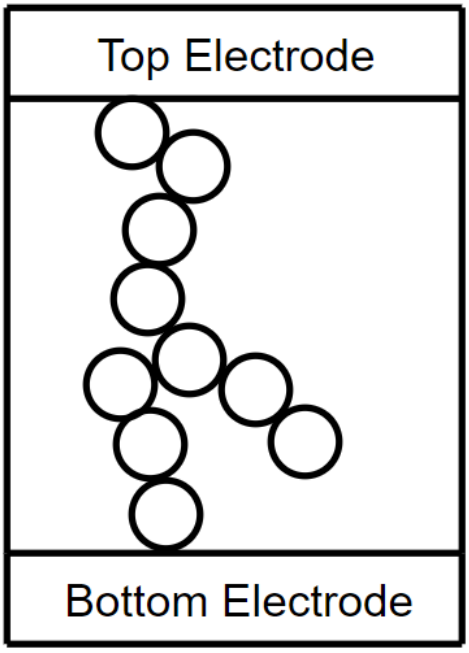

Figure 2.4: Filament formation in TMOs.

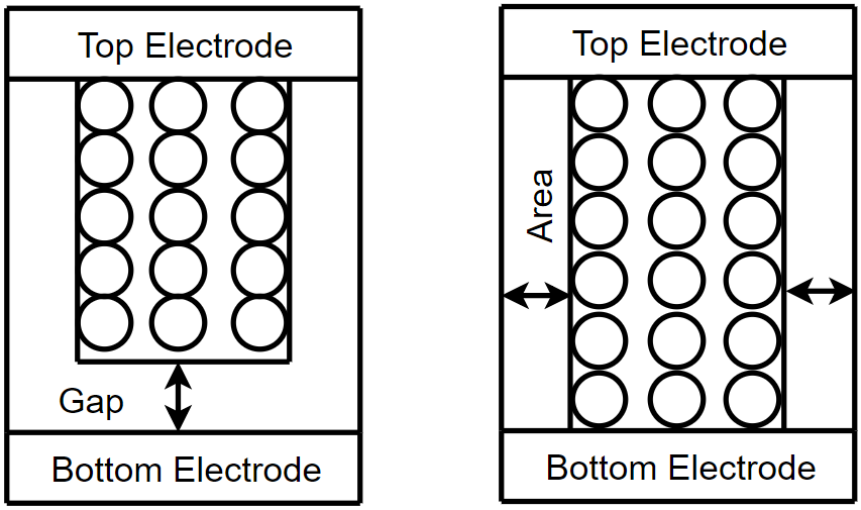

Figure 2.5: Visual representation of the approximate model of filament formation in [3](left) and [4] (right).

Given the stochastic nature of filament formation, it is usually challenging to ensure that either picture is the correct one. In fact, both might occur in the same device [50, 51].

This dilemma gave rise to resistance based models which leverage the instantaneous resistance of the device as the state variable [52, 6, 50, 51]. These models are motivated by the fact that experimental data is often reported in the form of I-V sweeps as a function of the external excitation from which a direct relation between the applied stimulus and the resistance can be drawn. These models are often empirical and although may not be predictive nor scalable, they are often preferred by circuit designers since they are simple, intuitive and, most importantly, based on measurable parameters which facilitates parameter extraction.

Electroforming of Transition Metal Oxide memrsitors

Transition Metal Oxides, despite their merits such as low operation voltage, fast switching times and high data retention, suffer a major drawback which is the requirement of electroforming. Electroforming typically requires voltages that are higher than the voltages used in most of the advanced CMOS processes. Several experimental studies have been presented trying to lower the forming voltages of TMO devices to enable their integration in standard CMOS processes. Those studies focused on varying devices geometry and/or

process parameters and study their impact on the forming voltage. In [8, 53], impact of scaling the device's thickness and area on the forming voltage was studied. It was shown that decreasing the device's thickness and/or increasing the device's area results in reducing the forming voltage. Other factors such as local field enhancement may also impact the forming voltage [27].

On the modeling front, Kinetic Monte Carlo (KMC) simulation studies were conducted in [54, 55, 56] to study the characteristics of electoforming in TMO devices. To the best of the author's knowledge, little has been done on the analytic modeling of electroforming. While KMC simulation is a useful numerical vehicle that helps provide better understanding of the forming process, analytic models may be preferable for they provide more insight into the parameters affecting the forming voltage which will be addressed in this dissertation.

2.3.2 Insulator Metal Transition devices

Insulator Metal Transition (IMT) devices are thermally driven resistive switches [28]. Their switching behavior can be described as a volatile memristor device [12]. Several experimental studies have been reported on IMTs. Works in [57, 58] argue that temperature due to Joule heating is the main source of resistive transition in those devices while [59, 60, 61] suggest that Joule heating is insufficient and that electric field is the main cause of phase transition.

Few models have been also proposed such as the models in [32, 62] which are TCAD-like models that rely on solving coupled differential equations. The work in [23] presented a behavioral model that is compatible with SPICE. Yet, the model did not consider the temperature dynamics of IMT devices. A SPICE compatible model that captures the temperature dynamics of IMT switching is still lacking which will be addressed in this work.

2.4 Crossbar arrays of memristive devices

The crossbar array is the basic nanoelectronic structure in which memristors are integrated. Crossbars consist of two sets of perpendicular wires with memristors integrated at the intersection between each horizontal and vertical wire as shown in Fig 2.6

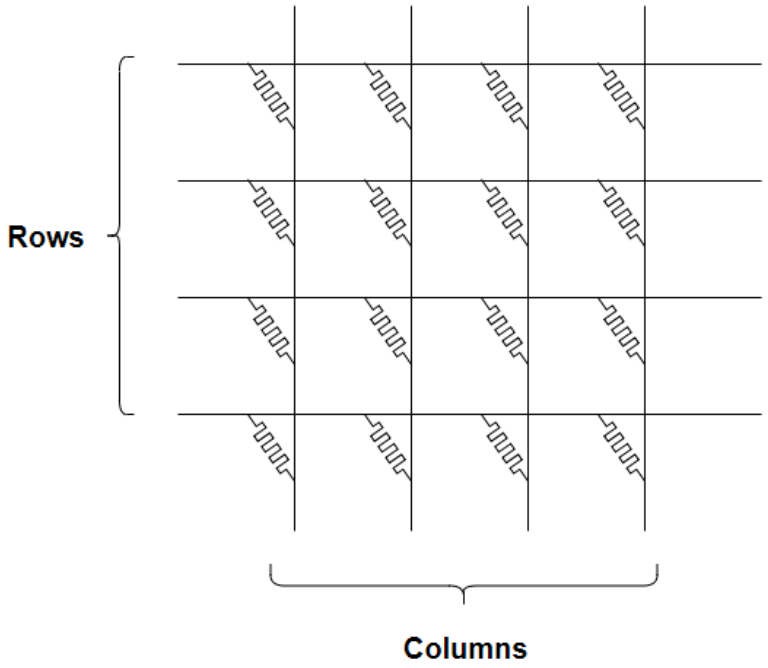

Figure 2.6: The crossbar array with memristors integrated at each junction.

In memory arrays, each cell represents a bit which can take either logic high or logic low depending on whether the device is at low resistance state (LRS) or high resistance state (HRS), respectively. In neuromorphic arrays, on the other hand, each cell represents a synaptic weight and can, ideally, take any intermediate resistance value.

2.4.1 Crossbar Memory Arrays

Crossbar memory arrays based on memristive devices are often referred to as Resistive RAMs or simply: ReRAMs. Fig 2.7 depicts the crossbar array under read and write operations. The write operation can be either SET or RESET:

- SET: V_{wr}/ground are applied across the row/column of the selected cell.
- RESET: ground/V_{wr} are applied across the row/column of the selected cell.

where V_{wr} is the write voltage. In the read operation, V_{read} is applied to the row of the selected cell while a load resistance R_{load} is connected to the column of the selected cell. The voltage drop across R_{load}, V_{sense}, is then fed to a sense amplifier to interpret the logic value

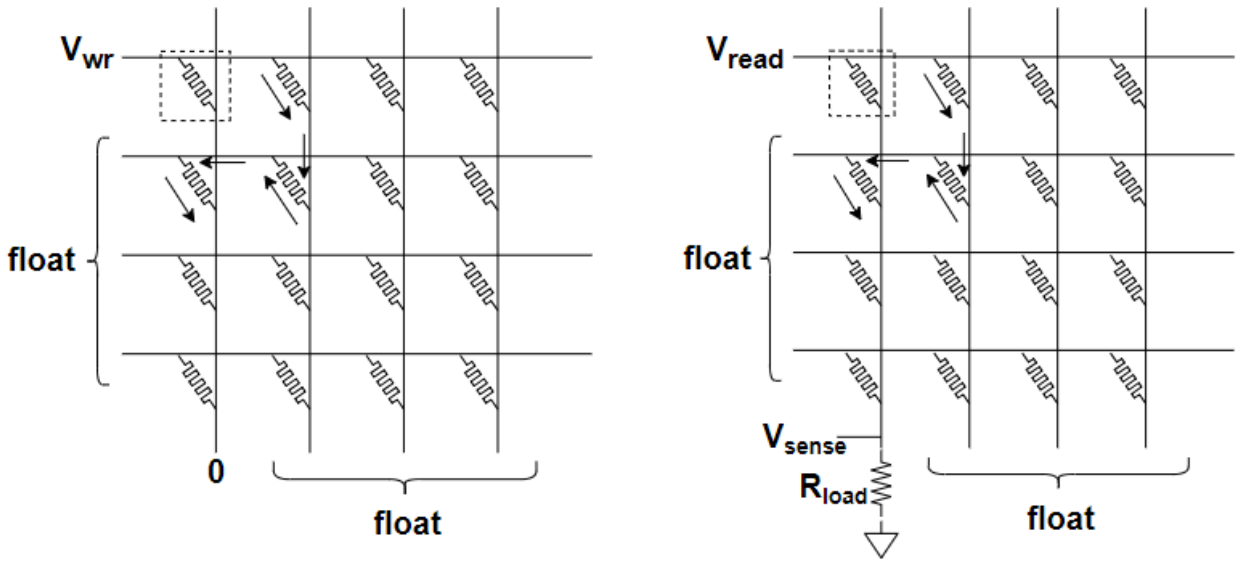

Figure 2.7: Write (left) and Read (right) operation in ReRAM arrays.

of the cell. In both scenarios, all unselected cells are left floating which gives rise to sneak path current. Sneak paths are unintended paths for current flow which are inherent to the crossbar structure. The arrows in Fig 2.7 depict the sneak path. Sneak paths may result in erroneous read/write operations.

1T1R ReRAM arrays

One way to eliminate sneak paths is adding a series FET device to the memristor at each cell as shown in Fig 2.8, thus, known as: One-Transistor One-Resistor (1T1R). The gate terminal of the FET device is used as a selector terminal such that only the FET connected to the selected device is activated while all other FETs are deactivated. Given the high OFF resistance of the FET devices (usually Giga Ohm range), sneak path currents are significantly suppressed. The drawback, however, is that the introduction of a FET device limits the Back-End-Of-Line (BEOL) of ReRAM arrays and increases the footprint of each memory cell. These challenges lead to the development of 1S1R arrays which are considered in this work.

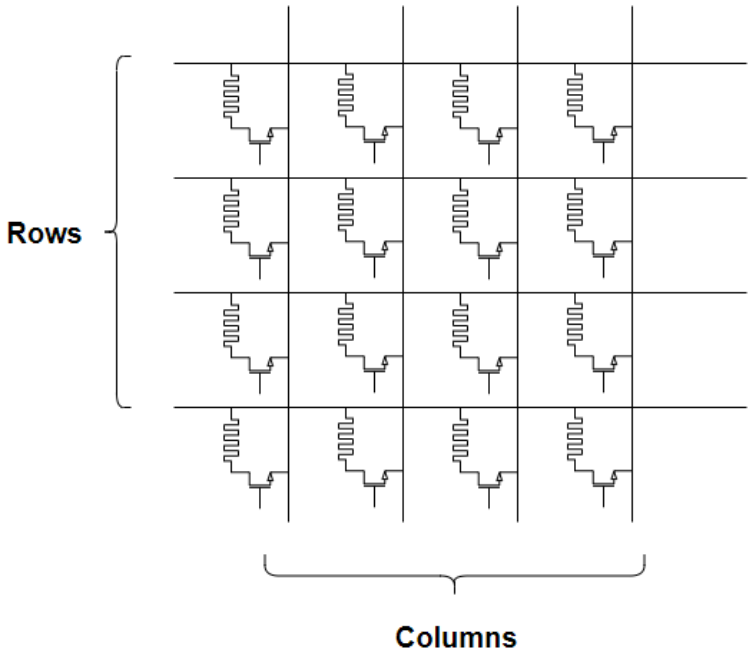

Figure 2.8: 1T1R ReRAM arrays.

1S1R ReRAM arrays

One-Selector One-Resistor (1S1R) arrays combine advantages from both 1R arrays and 1T1R arrays. A two terminal selector element is added in series with the memory device to circumvent sneak paths. The selector element typically possess a high ON/OFF ratio and switches from OFF to ON only when selected. Hence, it delivers the same functionality as the FET device while maintaining the same cell area and the BEOL compatibility which are critical in accomplishing the ultimate goal of 3D integration [63].

Two bias schemes exist in 1S1R arrays which are: 1/2 bias scheme and the 1/3 bias scheme depicted in Fig 2.9. In the 1/2 bias schemes, all unselected rows and columns are biased with $V_{wr}/2$ while in the 1/3 bias scheme the rows are biased with $V_{wr}/3$ and the columns are biased with $2V_{wr}/3$. This technique eliminates sneak path current from flowing while ensuring only $V_{wr}/2$ and $V_{wr}/3$ drop across the cells in the same row and column of the selected cell which is typically lower than the cell switching threshold.

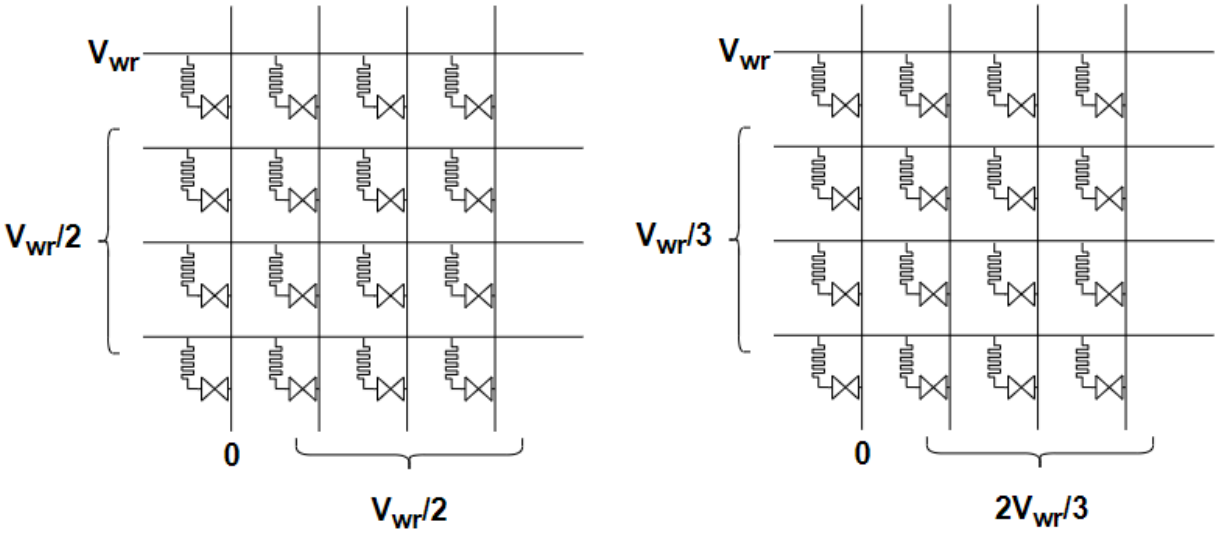

Figure 2.9: Bias schemes in 1S1R arrays during the write operation.

2.4.2 Crossbar Neuromorphic Arrays

Neuromorphic crossbar arrays adopt the same structure as Memory arrays. In the write operation, similar challenges are encountered as those found in memory arrays and, therefore, the same write techniques and methodologies are adopted. Unlike memory arrays, however, neuromorphic arrays store synaptic weights as opposed to bits and, hence, they are often referred to as synaptic arrays. The read operation, however, is somewhat different. During reading, the columns are connected to the output neurons and the weighted sum of the inputs is evaluated. Fig 2.10 depicts the write and read operation in neuromorphic crossbar arrays

2.5 Conclusions

This section reviewed some of the literature on memristive neuromorphic arrays. The theory of memristive devices was first introduced followed by memristor models for circuit simulation. Memristive crossbar arrays were then discussed in both memory and neuromorphic applications. Read and write techniques were also investigated and design challenges were demystified.

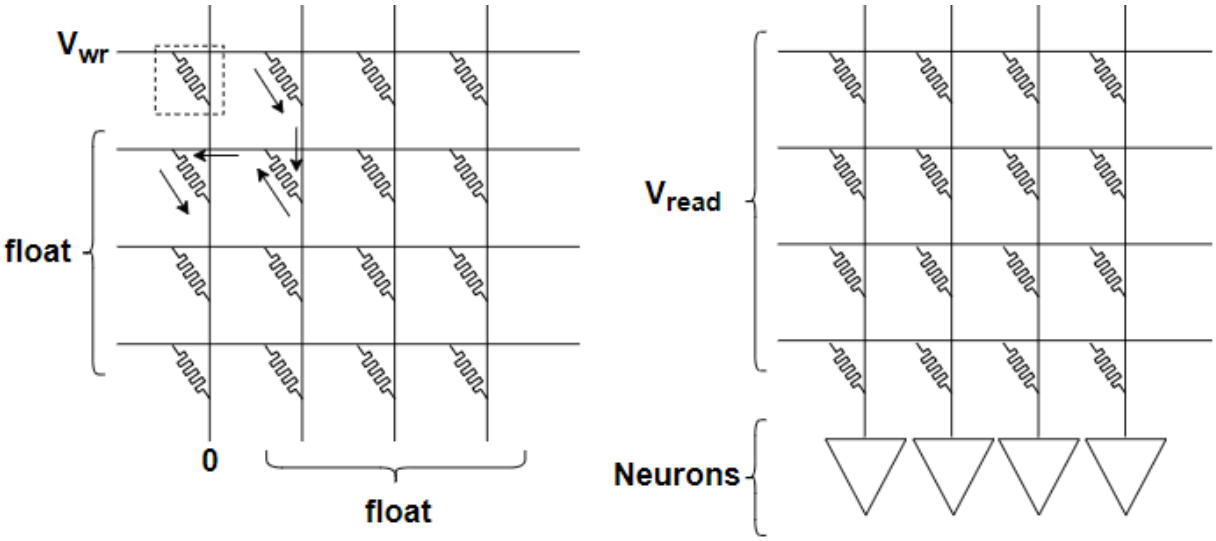

Figure 2.10: Write and read operation in neuromorphic crossbar arrays.

Chapter 3

A Practical Memristor Model Suitable for Circuit Design and Simulation

3.1 Introduction

Memristors have emerged as strong candidates for future computing paradigms. Their switching dynamics and small footprint have made them suitable candidates in applications such as memory and neuromorphic systems. Since the realization of the first physical memristor by HP in 2008, a surge of memristor based applications and architectures have been proposed which requires accurate yet computationally efficient compact models to enable the simulation of memristor based circuits in SPICE environment. Several memristor models have been proposed that range from simple behavioral models [2, 36] that are derived from the original equations proposed by Chua in [10] to complex physics-inspired phenomenological models [40, 39, 3] which are usually peculiar to some specific switching mechanism. The difficulty in modeling memristive devices lies in their dynamic behavior. Unlike most other semiconductor devices, memristor modeling relies on defining a state variable which can either be an abstract quantity in the case of behavioral models or a physical quantity in physical models. The challenge is that neither approach reflects measurable parameters that experiments report. Physical data is usually presented in terms of V-I sweeps from which the instantaneous resistance can be extracted. Connecting the resistance to the state variable is often challenging!

While it is necessary to build a generalized physical SPICE compatible memristor model that can model any generic memristive device based on its material characteristics and geometry, the lack of such a model has motivated the development of empirical models that are easily fitted to measurable parameters extracted from experimental data. In [52], Pino *et al.* developed an empirical model for a chalcogenide memristor. The model is a piece wise model which divides the switching operation into a subthreshold region where no change in resistance is allowed and an operation region where the change in resistance exhibits an exponential relation with the applied voltage. This model, however, did not account for the resistance saturation near the boundaries. Models in [6, 50] captured all the switching characteristics of the device but had too many fitting parameters. In these empirical models, unlike most other models, the state variable is the resistance of the device itself which makes them amenable to parameter extraction.

In this chapter, an empirical compact model for TMO memristors is proposed. The proposed model builds off the model used in [64] and is based on measurable parameters. The model parameters are chosen such that they can be easily tweaked to fit experimental data which facilitates parameter extraction.

3.2 Background

This section reviews few of the most commonly used models available in the literature.

3.2.1 Memristor Modeling

The existence of the memristor element was theoretically predicted in 1971 by Leon Chua [10]. It wasn't until 2008, however, that the first physical realization of a memristor was reported when HP announced the first recognized working memristor prototype based on a Titanium Dioxide process [2]. Memristors are generally modeled by a system of two coupled equations as expressed in (3.1) and (3.2), respectively:

$$I = G(w, V)V, \tag{3.1}$$

$$\frac{dw}{dt} = f(w, V), \tag{3.2}$$

where (3.1) is the output equation that describes the relationship between the voltage applied across the device and current flowing through the device and (3.2) is the state equation that describes the state dynamics of the memristance. A simple memristor model was proposed by HP for the TiO_2 device [2]. It assumes a linear ion drift of the oxygen vacancies and is described as shown in (3.3) and (3.4):

$$V = (R_{on}\frac{w(t)}{D} + R_{off}(1 - \frac{w(t)}{D}))I, \tag{3.3}$$

$$\frac{dw}{dt} = \mu_v \frac{R_{on}}{D} I, \tag{3.4}$$

where w is the length of the oxygen vacancy rich region, D is the total thickness of the switching layer of the memristor device, μ_v is the average mobility of oxygen vacancies and R_{on} is the low resistance state of the device. This model describes the memristor as two series resistors where R_{on} is the resistance of the oxygen vacancy rich region, R_{off} is the resistance of the oxygen vacancy deficient region and w is a state variable that modulates the memristance of the device based on the applied voltage. This model, however, failed to capture the experimental data which limited its validity [39, 40]. Specifically, the assumption that oxygen vacancies drift with constant velocity under the action of the electric field was proven to be inaccurate and that a nonlinear dependence is more likely to occur.

3.2.2 Physics-inspired Memristor Models

Several physical memristor models have been proposed in the literature. Those models can be divided into three categories based on their switching and conduction mechanisms. Table 3.1 summarizes the existing models and their respective switching and conduction mechanisms they assume.

Table 3.1: Existing Physical Models

Model	Stukov *et al.* [2, 39]	Pickett *et al.* Shahar *et al.*[40, 41, 42]	Yu *et al.* [3, 47]
Switching mechanism	Ion drift	Simmon's barrier tunneling	Arrhenius Law
Conduction mechanism	Ohmic+tunneling	tunneling	Ohmic+tunneling

Strukov *et al.* proposed the very first models which assumed that oxygen vacancies drift under the action of electric field [2, 39]. The drift velocity of oxygen vacancies can be a linear function of the electric field (i.e. $v = \mu E$) in linear ion drift models as in [2, 36] or an exponential function (i.e. $v = \mu E_0 e^{E/E_t}$) as in [39]. Models in [36, 37, 38] capitalized on the model in [2] and added window functions to capture the nonlinear dynamics near the boundaries. Pickett *et al.* [40] and Shahar *et al.* [41, 42] proposed models based on the Simmon's barrier tunneling model [44] where the device is modeled as a tunneling barrier modulated by the applied electric field. Models in [3, 47] are based on the reaction rate equation which assume that the creation and annhilation of oxygen vacancies follow Arrhenius law.

3.2.3 Resistance-based Memristor Models

The resistance-based approach to memristor modeling relies on using the instantaneous resistance, measured at a non-disturbing bias voltage, as the state variable. This approach alleviates the difficulty of finding a state variable that best reflects the switching dynamics of a particular memristor device and draws a direct connection between the applied stimulus and the resistance as illustrated in in Fig 3.1. This approach facilitates parameter extraction as will be shown in the later sections.

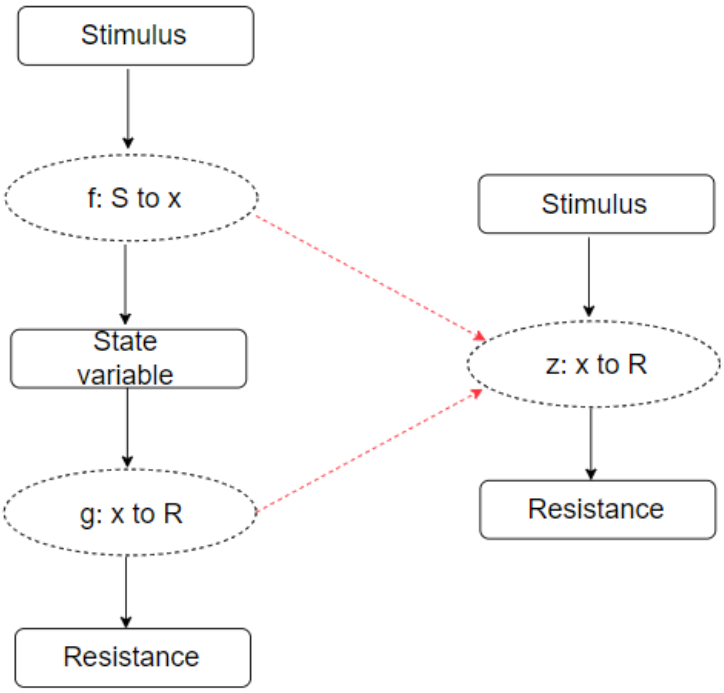

Figure 3.1: Physics-inspired models vs. resistance-based models

3.3 Proposed Memristor Model

The proposed model builds off a previously developed Piece Wise Linear (PWL) model first proposed in [65]. This section, therefore, starts by discussing the PWL model. Then, the proposed polynomial model is discussed.

3.3.1 PWL Model for HfO_2 Memristors

The PWL model considered in this section was originally used to model HfO_2 devices fabricated in-house at SUNY Polytechnic Institute. This model is adapted from another model developed by McDonald *et al.* in [65] and is expressed as:

$$\frac{dM}{dt} = \begin{cases} -\frac{\Delta r V(t)}{t_{swp} V_{tp}}, & V(t) > V_{tp} \\ \frac{\Delta r V(t)}{t_{swn} V_{tn}}, & V(t) < V_{tn} \\ 0, & \text{otherwise} \end{cases} \tag{3.5}$$

where $M_{t+1} = M_t + \frac{dM}{dt}\Delta t$ which captures the evolution dynamics of the resistance and the output equation is assumed to follow Ohm's law $V = IM$. The model clips the resistance at $HRS(LRS)$ if it goes above/below $HRS(LRS)$. Parameters $t_{swn}(t_{swp})$ capture the time taken to switch from LRS to $HRS(HRS$ to $LRS)$. Parameter Δr is the difference between HRS and LRS. Parameters $V_{tn}(V_{tp})$ are the negative (positive) thresholds. All parameters are measurable parameters and are extracted from physical data. The measured parameters for the fabricated HfO_2 devices are shown in Table 3.2.

3.3.2 Proposed Memristor Device Model

Despite the simplicity of the PWL model, physical measurements deviated significantly from the model predictions. Specifically, the PWL model failed to capture two main characteristics of the memristor switching dynamics which are:

- Nonlinear dependence of the rate of change of resistance on the applied voltage
- Plateauing of the resistance near the boundaries

Table 3.2: HfO_2 memristor parameters

Parameter	Value
LRS	$3k\Omega$
HRS	$45k\Omega$
Δr	$HRS - LRS$
V_{tp}	$0.75V$
V_{tn}	$-0.5V$
t_{swp}	$10ns$
t_{swn}	$1us$

These characteristics were added to the proposed model:

$$\frac{dM}{dt} = \begin{cases} -C_{LRS}(\frac{V(t)-V_{tp}}{V_{tp}})^{P_{LRS}} f_{LRS}(M(t)), & V(t) > V_{tp} \\ C_{HRS}(\frac{V(t)-V_{tn}}{V_{tn}})^{P_{HRS}} f_{HRS}(M(t)), & V(t) < V_{tn} \\ 0, & \text{otherwise} \end{cases} \tag{3.6}$$

where the $\frac{\Delta r}{t_{sw}}$ term is absorbed in the C coefficient. f_{HRS} and f_{LRS} capture the resistance saturation (commonly referred to as window functions). Expression (3.7) presents the proposed window function that can be easily fitted to measurable parameters.

$$f(M(t)) = \begin{cases} \frac{1}{1+e^{\frac{M(t)-\theta_{HRS}HRS}{\beta_{HRS}\Delta r}}}, V(t) < V_{tn} \\ \frac{1}{1+e^{\frac{\theta_{LRS}LRS-M(t)}{\beta_{LRS}\Delta r}}}, V(t) > V_{tp} \end{cases} \tag{3.7}$$

In the Verilog-A code provided in Appendix A, $M(t)$ is still clipped to either HRS or LRS should the resistance reach either boundary for modeling convenience.

3.3.3 Comparison Between the PWL and proposed Models

In order to show the improvement posed by our proposed model over the PWL model, both models are compared against experimental data. Fig 3.2 depicts the I-V sweeps for both simulation and experimental measurements from [5]. It is readily shown that both models exhibit hysteresis in the V-I plane - a fingerprint of memristive devices. Fig 3.3 captures the change in resistance with respect to the applied voltage for both the PWL model and the proposed model. It is readily observed that the PWL model fails to capture the non-linearity exhibited by the device. The proposed model, however, captures such non-linearity with a control parameter P which can be tweaked to fit any memristive device. Another drawback with the PWL is the high discontinuity of the model around the memrsitor threshold. This discontinuity not only hampers the accuracy of the model, as shown in Fig 3.3, but also results in convergence difficulties during circuit simulation which will be addressed in the next section.

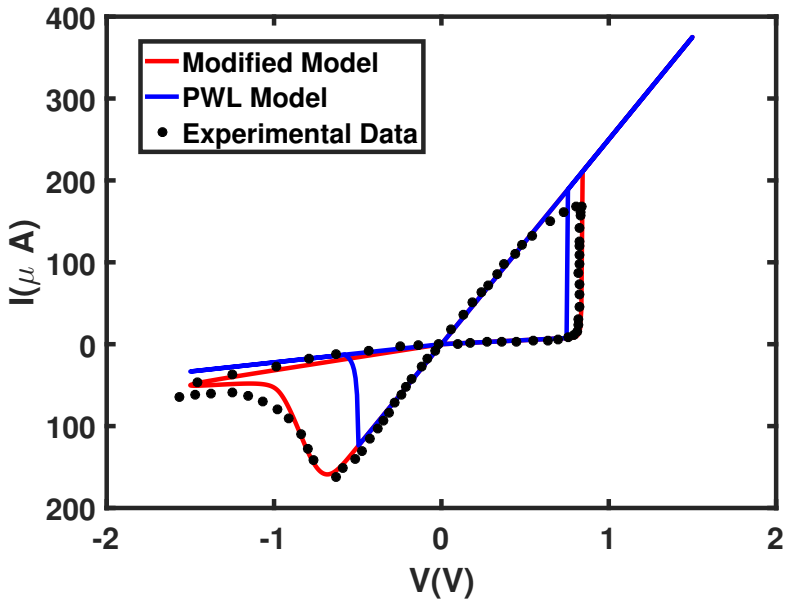

Figure 3.2: I-V plots of the linear and polynomial model against experimental data. Experimental data was extracted from [5].

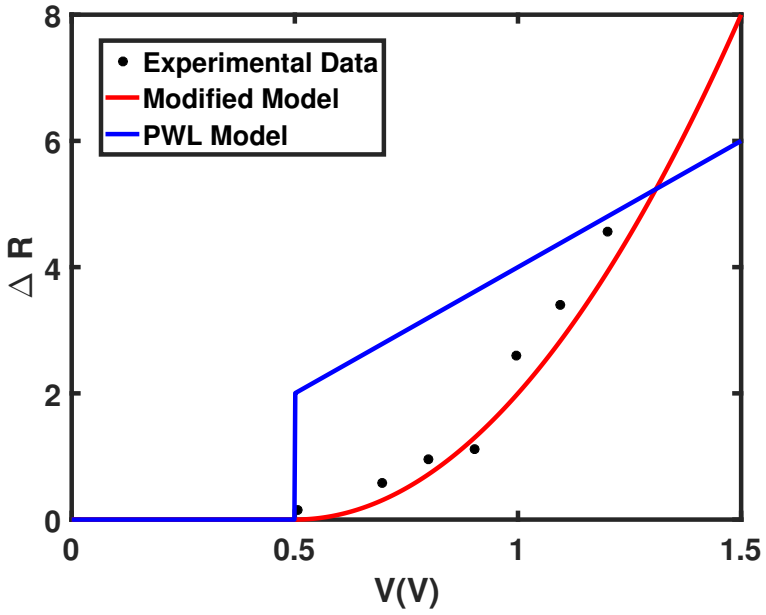

Figure 3.3: Change in resistance vs applied voltage for both the proposed model and the PWL model. Experimental data was extracted from [6].

.

Oxide based memrisors are characterized by resistance saturation near the boundaries (i.e. HRS and LRS). This phenomenon has been conventionally captured by window functions. In this work, a sigmoid window function with two fitting parameters θ and β is employed. Fig. 3.4 depicts the resistance evolution with time for the PWL model and the proposed model. Pulses with equal magnitude and widths were applied across the device yielding and incremental increase in the resistance from $3k\Omega$ to $18k\Omega$. As expected, experimental data shows that the pulse response plateaus. Unlike the PWL model which predicts a linear increase in resistance over time, the proposed model captures the plateauing effect. Parameters θ and β were selected such that the model accurately fits the experimental data. A more detailed explanation about how these parameters are selected is provided in the next section.

3.4 Parameter Extraction

This section presents the parameter extraction methodology for the proposed model. The model contains six parameters: HRS/LRS, V_{tp}/V_{tn}, C_{LRS}/C_{HRS}, P_{LRS}/P_{HRS}, $\theta_{HRS}/\theta_{LRS}$ and β_{HRS}/β_{LRS}.

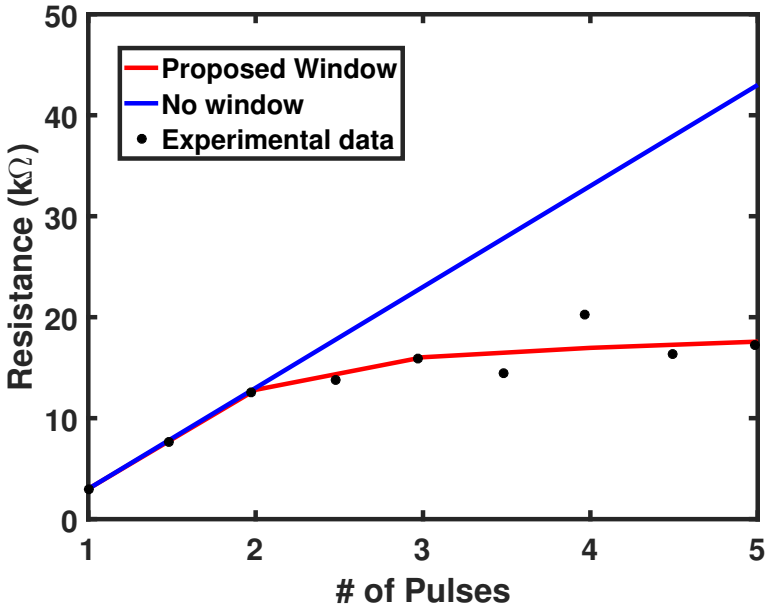

Figure 3.4: Resistance vs.time plots of both models against experimental data.

Two of the parameters, namely: HRS/LRS and V_{tp}/V_{tn}, can be extracted from DC sweeps while the rest of the parameters require transient testing as shown in the next subsections. The experimental data used for parameter extraction is extracted from a hybrid CMOS/memristor chip available at UTK and fabricated at SUNY Polytechnic.

3.4.1 Hafnium Oxide Device Structure

The Hafnium Oxide memristor device is integrated between the first and second metal layers. Fig 3.5 depicts the physical structure of the device. Fig 3.6 shows an image of the memristor device under test extracted from the probe station. Fig 3.7 depicts DC I-V sweeps of the device under test. A forming step is first applied to electroform the device. The forming voltage for this particular device is around $1.6V$. Multiple Set and Reset cycles are then applied demonstrating reliable operation.

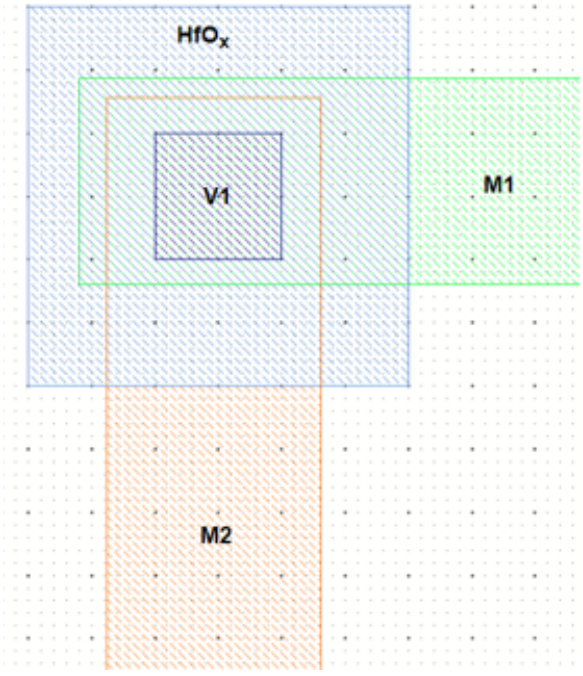

Figure 3.5: Hafnium Oxide device physical structure.

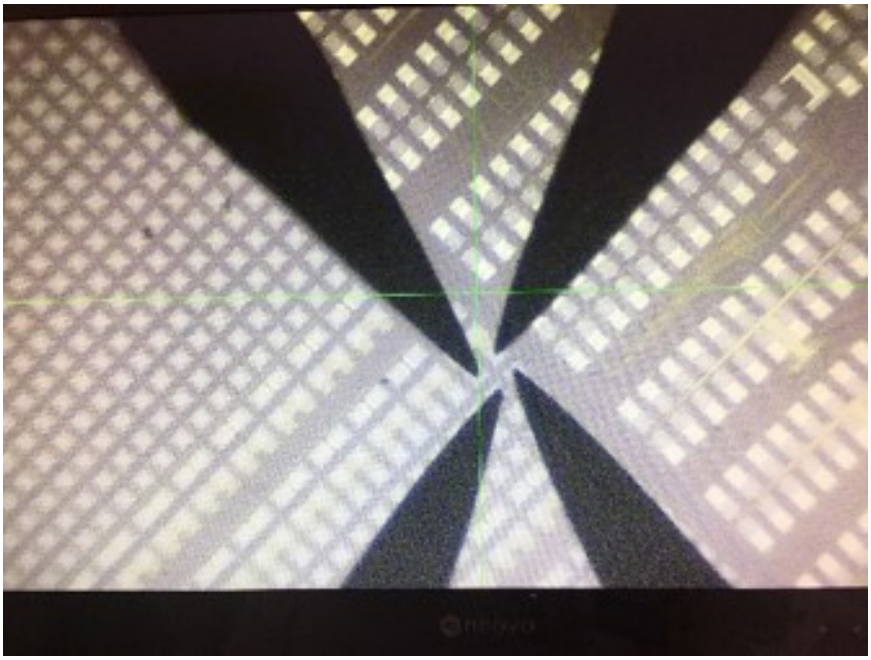

Figure 3.6: Image taken from the probe station while testing memristor devices.

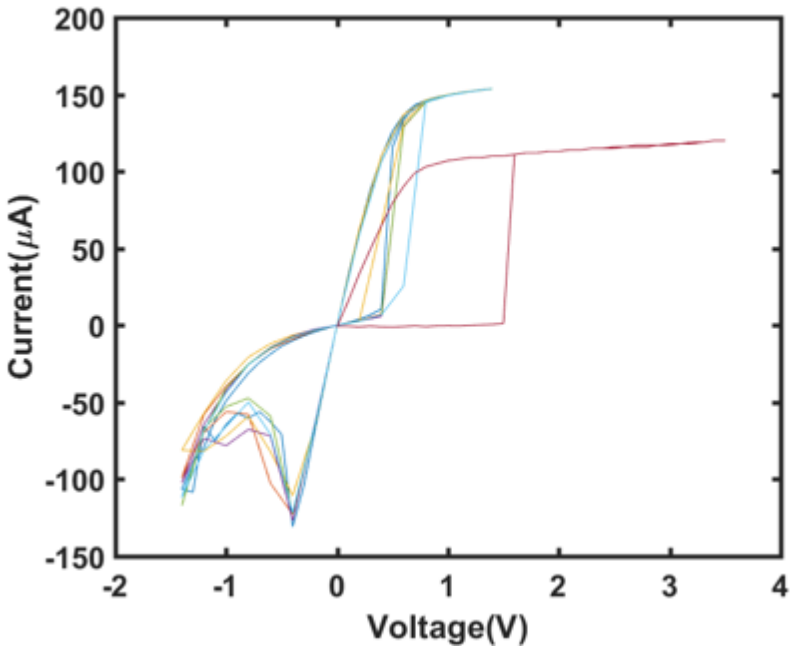

Figure 3.7: Characterization of the memristor device. Multiple SET/RESET cycles were executed after an initial forming step.

3.4.2 Parameter Extraction from DC sweeps

Parameters HRS/LRS and V_{tp}/V_{tn} can be extracted from the hysteresis loop achieved by running a dual sweep on the memristor device. As mentioned earlier, the device switches between HRS and LRS states when the voltage applied across the device reaches the threshold voltage. Since DC sweeps report V-I plots, HRS and LRS can be extracted from the hysteresis loop where the line with the higher slope represents LRS and the line with the lower slope represents HRS. The voltages at which the transition happens between the two lines are the threshold voltages. Fig 3.8 depicts the DC parameter extraction.

3.4.3 Parameter Extraction from Transient tests

The other parameters cannot be extracted from DC sweeps but require pulse based testing. Equation (3.6) can be expressed as follows:

$$\frac{dM}{dt} = g(V(t))f(M(t)), \tag{3.8}$$

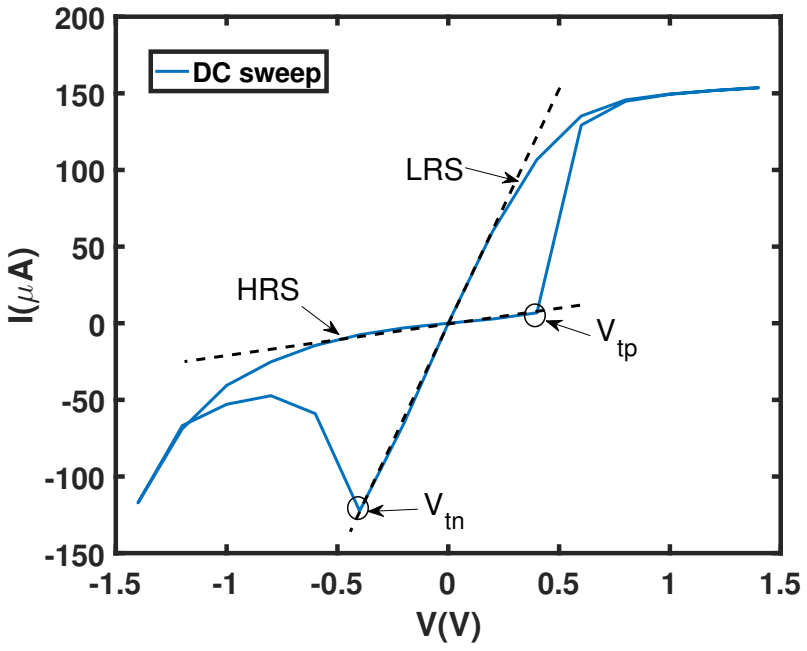

Figure 3.8: DC parameter extraction.

The rate of the change of the memristance is a function of both the applied voltage and the current resistance state. The extraction procedure is divided into two phases such that only one variable is changed at a time:

- Phase I: extract $g(V(t))$ while $f(M(t)) = 1$.
- Phase II: extract $f(M(t))$ while $g(V(t)) = constant$.

The following subsections discuss phases I and II of the transient parameter extraction procedure.

Phase I: Extracting $g(V(t))$

With $f(M(t)) = 1$, one can write $\frac{dM}{dt} = g(V(t))$. Hence, $g(V(t))$ can be extracted from the line slope of memristance vs. time (or pulse number) plots. However, for $f(M(t)) = 1$ to hold, this test should be executed before the plateauing effect takes place. Fig 3.9 depicts the resistance vs. number of pulses plot for $V = 1.2V$ and $V = 1.3V$.

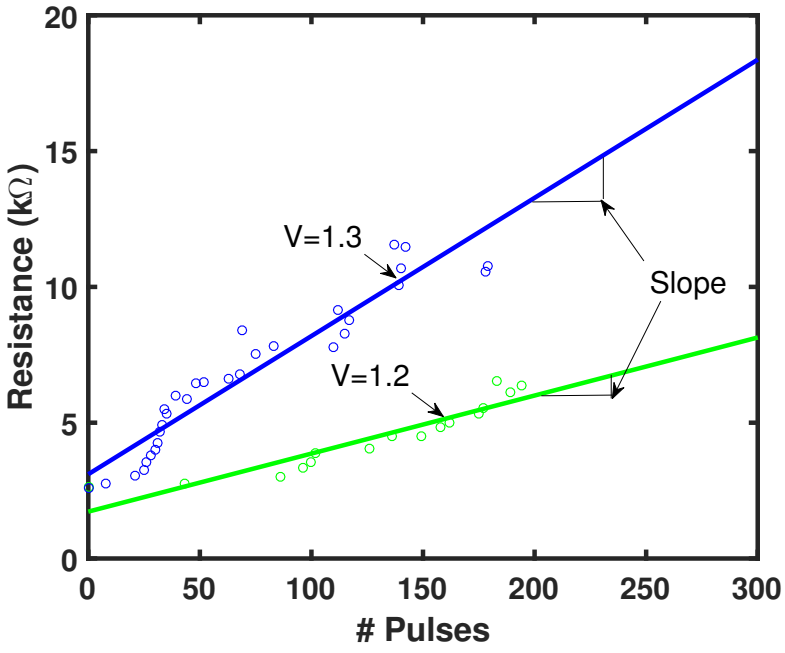

Figure 3.9: Resistance vs. number of pulses for different applied voltage magnitudes. Experimental data (symbols) and line slope (solid).

After extracting the slopes for multiple voltage magnitudes, $\frac{dM}{dt}$ vs. V can be plotted. Note, however, that the extracted slope from Fig 3.9 reflects $\frac{dM}{dN}$, where N is the pulse number. dt, however, can be extracted from dN such that $dt = dN * PW$, where PW is the pulse width. The pulse width used in this work is $1.5ns$. parameters C_{HRS}/C_{LRS} and P_{HRS}/P_{LRS} are then extracted from $\frac{dM}{dt}$ vs. V plot as shown in Fig 3.10.

Phase II: Extracting $f(M(t))$

This test requires $g(V(t))$ to be held constant such that $M(t) = k\Delta t f(M(t))$ where $g(V(t)) = k$. $f(M(t)$ can then be extracted from the resistance vs. time (or number of pulses as mentioned earlier) plot as shown in Fig 3.11. Parameter θ captures the plateauing point and β captures the sharpness of the plateauing.

3.5 Comparison with other Models

Table 3.3 compares the proposed model to other resistance based models in the literature.

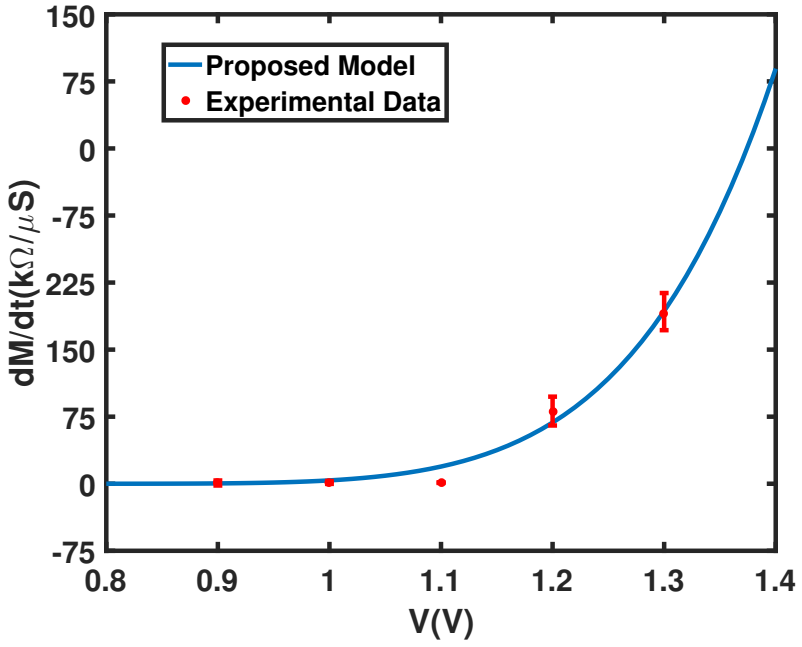

Figure 3.10: Rate of Change of resistance vs. voltage plot.

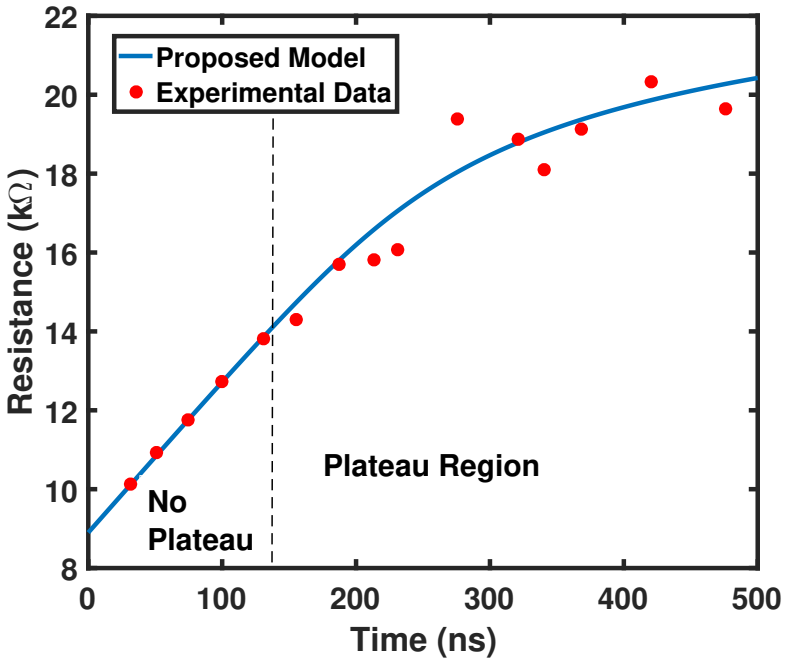

Figure 3.11: Resistance evolution with time for a fixed voltage amplitude.

Table 3.3: Model Comparisons

Model	Threshold	$\frac{dM}{dt}$ vs. voltage	Plateauing	no. of parameters
PWL	YES	Linear	No	1
Pino [52]	YES	non-Linear	No	2
Bayat [50]	YES	non-Linear	YES	6
Proposed	YES	non-Linear	YES	4

3.6 Model Convergence in Circuit Simulation

Memristive circuits are typically dense networks and, hence, might be computationally demanding in circuit simulation. Thus, the convergence of the proposed model is assessed and compared against the previous PWL model via a benchmark circuit proposed in [7] in which the authors compare the performance of various models for large networks of memristors. The advantage of this circuit is:

- It has no other devices but memristors. This makes the convergence of the simulation highly dependent on the memristor model.
- All memristors are updated periodically since all node voltages are updated for every transient step.

The circuit under consideration has 840 memristive elements. A $10kHz$ sinusoidal signal was applied at the input and the simulation was run for various transient times according to the applied test. All models are implemented in Verilog-A and simulations were executed on Spectre circuit simulator from Cadence . The proposed benchmark is depicted in Fig. 3.12.

A major drawback of the PWL model is the discontinuity of the model about the memristor threshold. It is not unusual in compact models to split the domain of operation into multiple regions with model equations specific to each region. It is important, however, to ensure that the model is smooth at the region boundaries to facilitate convergence during circuit simulation. Circuit simulators use the Newton-Raphson method to find the DC operating point of the circuit which is an iterative numerical method that requires at least first order continuity: the model equations and their first derivatives are continuous across the boundaries. The PWL model is not even zeroth order continuous around the memristor threshold: below V_t, the change in memristance is forced to zero while at V_t, the change in memristance is $\Delta r/t_{sw}$. This abrupt transition at the memristor threshold causes convergence difficulties. In our modified polynomial model, this problem was resolved which improved the convergence notably. Table 3.4 depicts the simulation times for both the PWL model and the proposed model. A significant difference between the simulation times of the both models is observed. This difference is mainly attributed to the model smoothness.

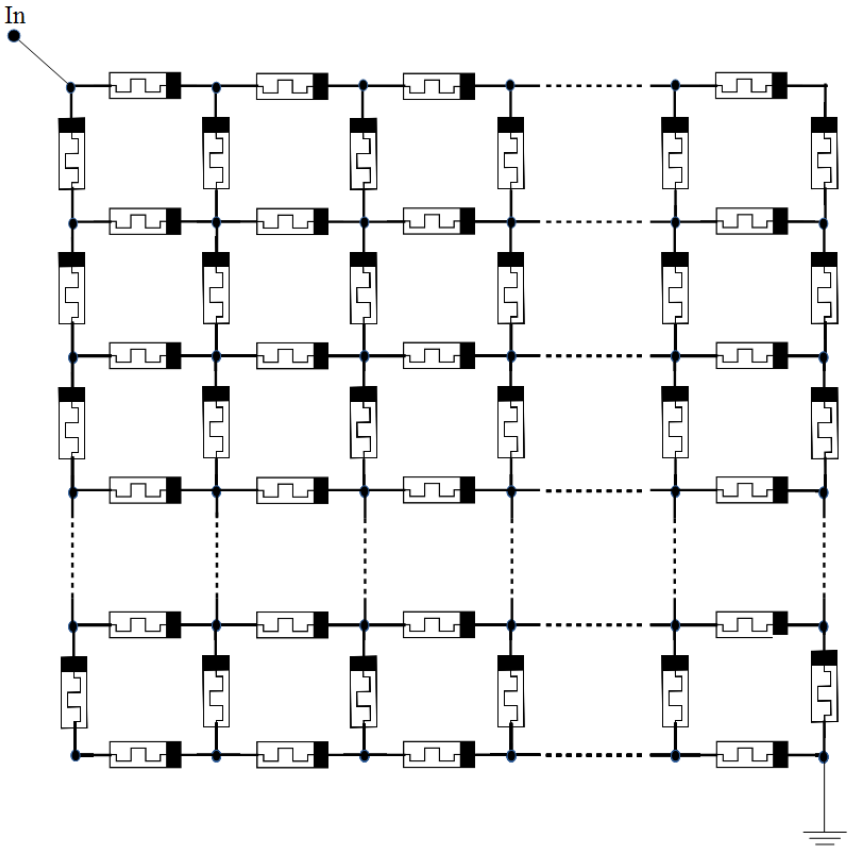

Figure 3.12: Benchmark circuit for the convergence test [7].

Table 3.4: Effect of the model smoothness on the simulation time

Model	PWL	Proposed
Simulation time	4m	< 10s

3.7 Conclusions

In this work, an empirical memrsitor model was proposed. The proposed model builds off a previously proposed piece wise linear model and uses the instantaneous resistance as the state variable. The proposed model captures the three main characteristics of oxide based memristors and is amenable to parameter extraction. The model equations are smooth across the different regions of operation to facilitate convergence during circuit simulation.

Chapter 4

Modeling Electroforming in Transition Metal Oxide Memristors

4.1 Introduction

Memristors based on Transition Metal Oxides (TMOs) materials [2] are widely used in nanoelectronic structures such as crossbar arrays. The requirement of a one time electroforming process, however, makes it challenging to integrate such devices in state-of-the-art CMOS processes. Electroforming typically requires higher than nominal voltages which bring about significant design challenges [66] and area constraints due to the introduction of a forming circuit that ultimately degrades the density advantage of crossbars. Efforts have been made to lower the forming voltages down to a level compatible with standard CMOS processes to enable seamless integration of TMO devices. Experiments have shown successful reduction in the forming voltage with device scaling and/or varying process parameters such as local field enhancement [67, 9, 8] achieved by structural modification of the oxide material. Yet, little has been done towards modeling this relationship. In this work, a physical model of electroforming in TMOs is proposed. The developed model identifies key geometric and material characteristics that impact the forming voltage. It is shown that there exists a linear dependence of the forming voltage on the oxide thickness and a logarithmic dependence on the oxide area. Local enhancement may also play a key role in

lowering the forming voltage. The proposed model provides insight into the key material and geometric characteristics that can be varied to reduce the forming voltage.

4.2 Forming Voltage model derivation

It was shown in [68] that, similar to gate oxide breakdown in FET devices [69], electroforming can be modeled using Poisson statistics as follows:

$$P(k, A) = \frac{(DA)^k}{k!} e^{-DA}, \tag{4.1}$$

where k is the number of breakdown paths, A is the area of the oxide and D is the density of the breakdown paths. Electroforming follows the weakest link character where formation is accomplished once a single filament is formed between both electrodes [70, 68] as shown in Fig 4.1. Hence, prior to forming, no path has yet been formed and, accordingly, one can write:

$$P(D, A) = e^{-DA}, \tag{4.2}$$

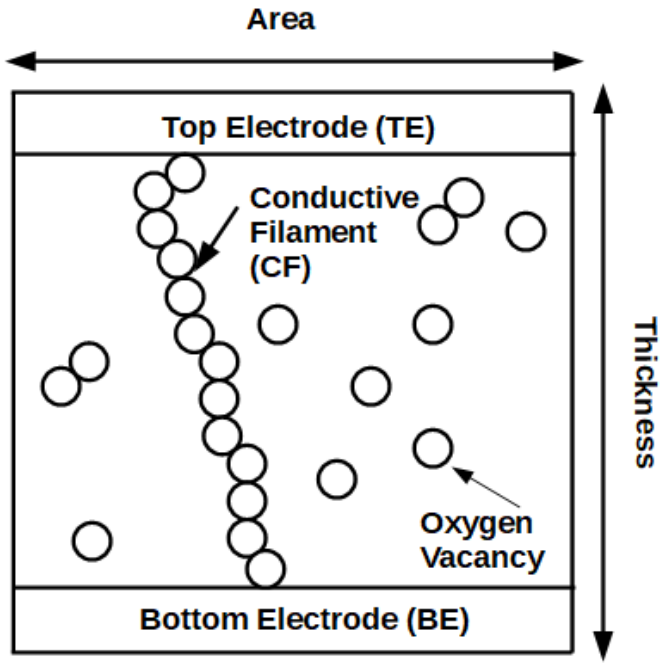

Figure 4.1: Oxygen Filament structure in TMO devicecs.

Expression (4.2) models the probability of having no forming paths between the top and the bottom electrodes. According to Poisson statistics, expression (4.1) captures the probability that k breakdown paths exist in an area A while expression (4.2) captures the area needed to form one breakdown path and, hence, can be used to model the formation of the percolating path between both electrodes in electroforming of TMOs. From a reliability modeling standpoint, expression (4.2) can be interpreted as oxide reliability as a function of area $R(A)$ with parameter D. Therefore, one can define Mean Area To Failure (MATF) as the average area required for the oxide to break down, formally derived as follows:

$$A_f = \int_0^\infty R(A)dA = \int_0^\infty e^{-DA}dA = \frac{1}{D}, \tag{4.3}$$

where A_f is the MATF. During electroforming, vacancies are induced based on the applied electric field which can be described using Arrhenius law [48, 71, 72]:

$$r = \nu e^{-\frac{(E_A - \alpha E)}{k_b T}} = r_0 e^{K\alpha \frac{V_F}{t_{ox}}}, \tag{4.4}$$

where K is $1/k_bT$, r is the generation probability of oxygen vacancies, ν is a characteristic frequency of generation, E_A is the average activation energy of oxygen vacancy generation, α is a barrier lowering coefficient reflecting the local field enhancement, E is the applied electric field across the oxide which can be described as the forming voltage divided by the oxide thickness such that $E = V_F/t_{ox}$, k_b and T are the Boltzmann constant and temperature, respectively.

By definition, D is the density of breakdown paths. Hence, D can be viewed as a chain of oxygen vacancies from the top to the bottom electrode whose length is proportional to the oxide thickness and, accordingly, D can be approximated as follows $D \approx r^{t_{ox}}$. This can be formally described according to Poisson statistics by the following expression:

$$P(k, dA) \approx \begin{cases} 1 - DdA, & k = 0 \\ DdA, & k = 1 \\ 0, & k > 1 \end{cases} \tag{4.5}$$

where in any arbitrary area ΔA, multiple paths to ground can exist. However, as ΔA becomes infinitesimally small such that ΔA approaches dA, one path can exist and the Poisson distribution can be approximated according to the above expression. From (4.3) and (4.4), the following forming voltage model can be derived:

$$V_F = \frac{ln(1/r_0)}{K\alpha} t_{ox} - \frac{ln(A_f)}{K\alpha}, \tag{4.6}.$$

4.3 Model Validation

The proposed model is validated against experimental data drawn from the literature as well as Monte Carlo simulations.

4.3.1 Model Verification Against Experimental Data

To validate the proposed model, two essential features need to be captured, namely: the linear dependence of the forming voltage on the oxide thickness and the logarithmic dependence of the forming voltage on the oxide area.

Fig 4.2 plots the forming voltage against the oxide thickness. The experimental data is drawn from the work in [8]. It is shown that the linear relationship between the forming voltage and the oxide thickness is captured by the proposed model.

Fig 4.3 plots the forming voltage against the oxide area for different oxide thicknesses. The experimental data is drawn from the work in [9]. Its is shown that the proposed model captures the logarithmic relationship between the forming voltage and the oxide area. The proposed model is fitted to the device with $t_{ox} = 5.2nm$ to extract r_o and $K\alpha$. Using the extracted model parameters (r_o and $K\alpha$), the model is used to predict the forming voltage for the two other oxide thicknesses. A decent match between the model prediction and the experimental data is observed.

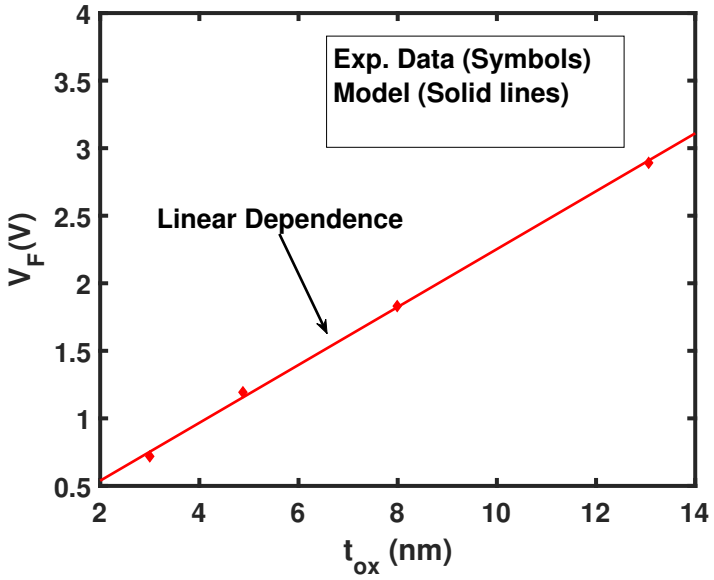

Figure 4.2: Forming Voltage vs. Oxide thickness for constant Oxide area. The experimental data is drawn from [8]. Experimental data (symbols) and proposed model (dashed line).

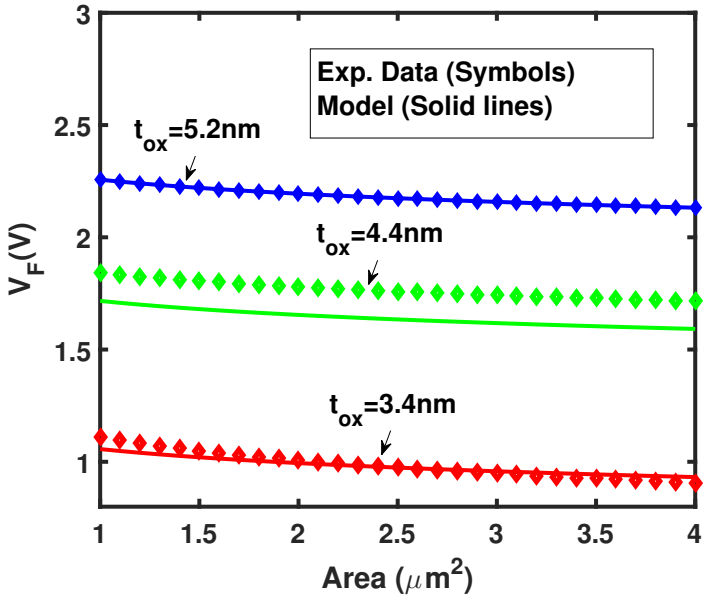

Figure 4.3: Forming Voltage vs. Oxide Area for various oxide thicknesses. The experimental data is drawn from [9]. Experimental data (symbols) and proposed model (dashed line).

4.3.2 Model Verification Against Monte Carlo Simulations

A numerical simulation framework is also developed in order to provide further validation of the analytic model. Numerical simulations allow better control over the model parameters, thus, helping with the investigation of the impact of each parameter on the forming voltage separately. In this work, Monte Carlo simulation is used as the numerical simulation vehicle since the governing equation, Arrhenius law, needs to be solved in a statistical framework. The Monte Carlo simulation algorithm is based on the work in [48]. First, the device is modeled on a 2D grid as shown in Fig 4.4 where each point on the grid represents a viable location for an oxygen vacancy. The length of the grid corresponds to the oxide thickness while the width of the grid corresponds to the oxide area. The presence/absence of oxygen vacancy is then determined by comparing the voltage dependent switching probability derived from Arrhenius law with a random test number. Finally, the algorithm checks if a continuous path of oxygen vacancies is formed between the top and bottom electrodes and a forming event is registered should this condition be met. Fig 4.5 depicts a flow chart of the Monte Carlo simulation framework developed in this work.

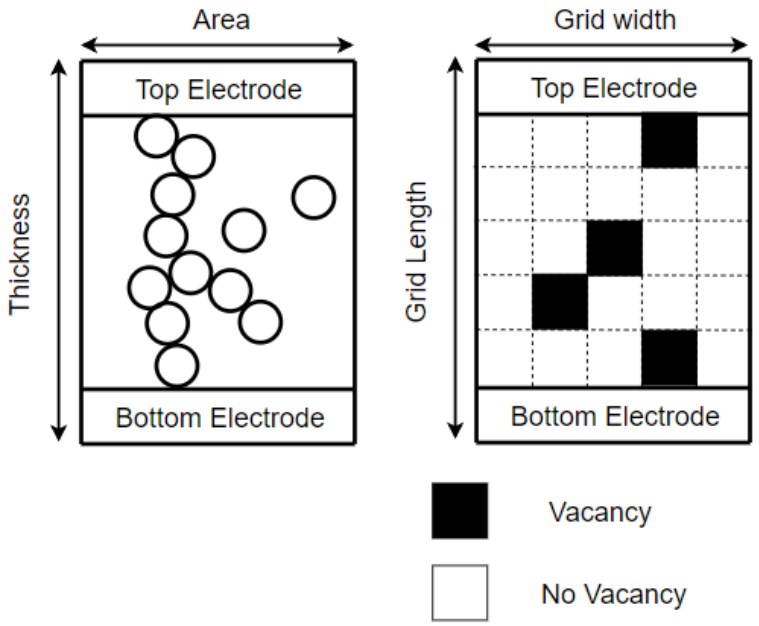

Figure 4.4: Grid model for the Oxide layer. Each square represents a viable location of an Oxygen vacancy. Vacancy (black) and No Vacancy (white).

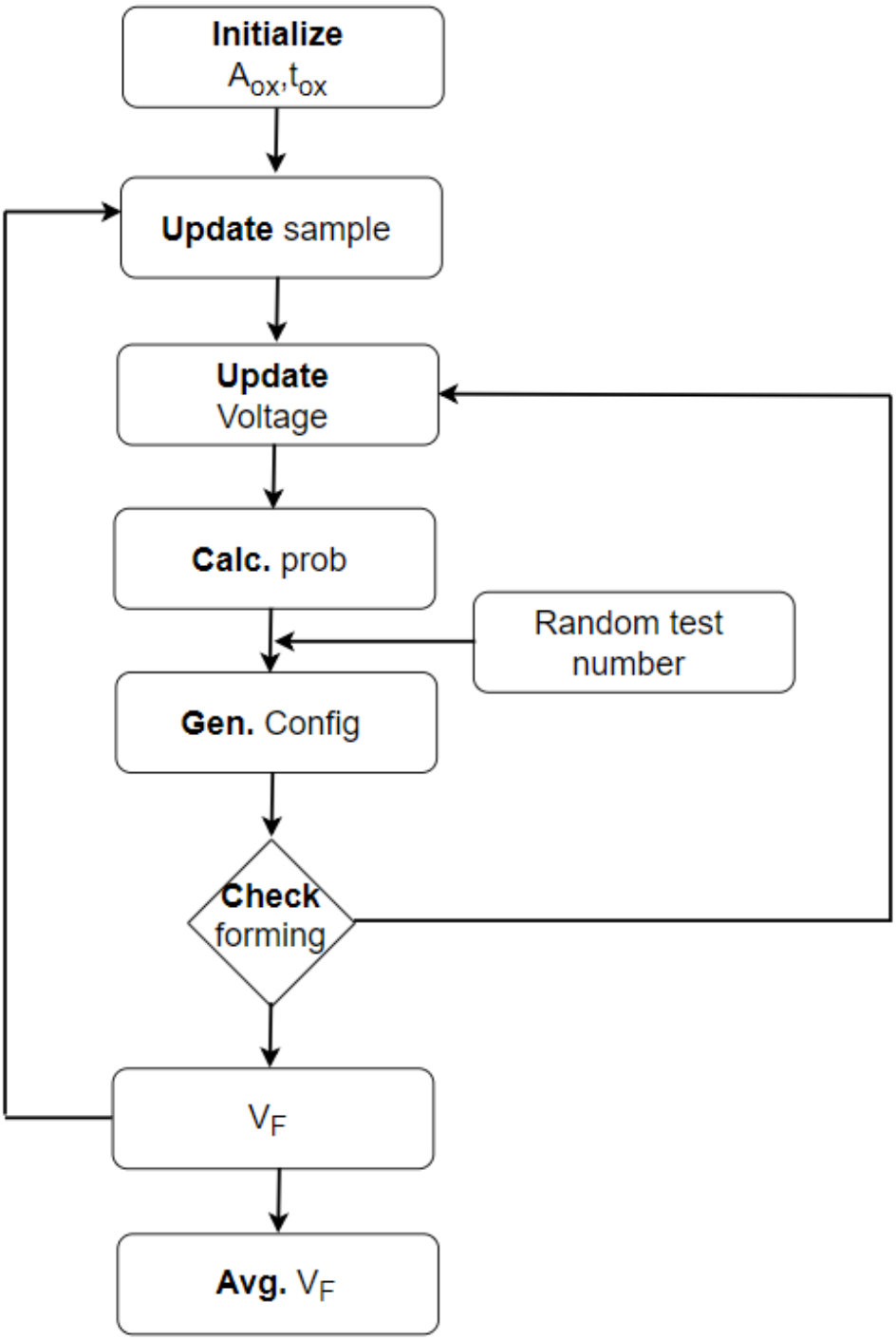

Figure 4.5: Monte Carlo Simulation framework.

Fig 4.6 plots the forming voltage against the grid area (grid width for a 2D model) for various grid thicknesses. The logarithmic relationship is still captured by the proposed model for the various grid thicknesses. Fig 4.7 plots the forming voltage against the grid thickness for various grid areas. The linear relationship between the forming voltage and the grid thickness is also captured for the various grid areas.

An important device parameter that can play a key role in lowering the forming voltage is the local field enhancement factor captured by the parameter α. While scaling the oxide thickness can lower the forming voltage, it comes at the expense of degrading the the ON/OFF ratio of the device which is critical to many applications. Scaling the oxide area might not yield significant improvement due to the weak (logarithmic) dependence of the forming voltage on the oxide area. Thus, local field enhancement may be a viable option for reducing the forming voltage. Fig 4.8 plots the forming voltage against grid area for different local enhancement factors. Higher local enhancement weakens the sensitivity of the forming voltage to the grid area. This characteristic is captured by the model as the the local enhancement factor controls the slope of the forming voltage versus grid area line plot.

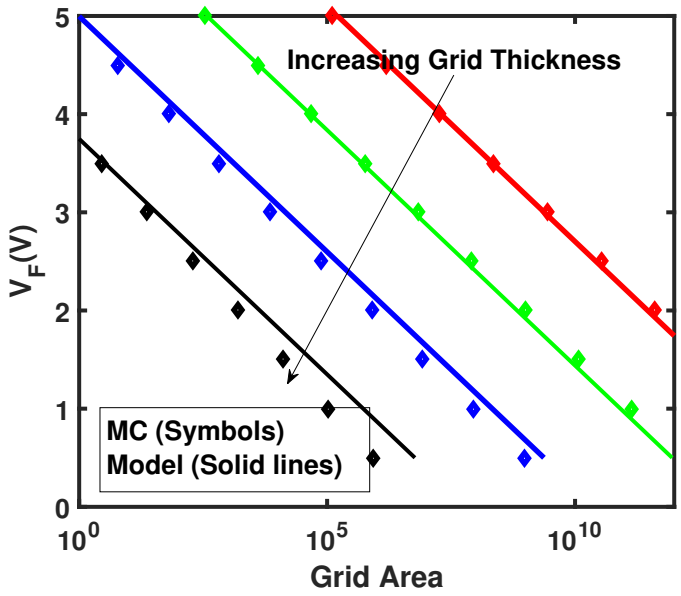

Figure 4.6: Forming Voltage vs. grid area for various grid thicknesses. Monte Carlo simulation results (symbol) and Analytic model (solid).

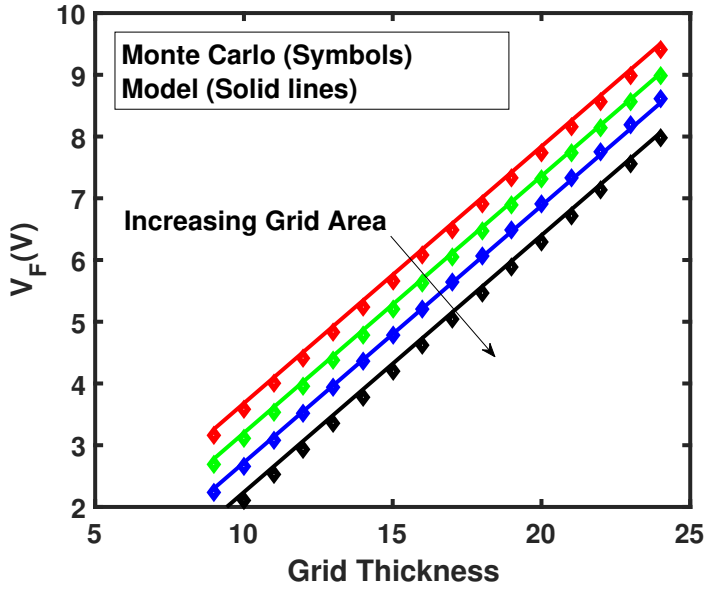

Figure 4.7: Forming voltage vs. grid thickness for various grid areas. Monte Carlo simulation results (symbol) and Analytic model (solid).

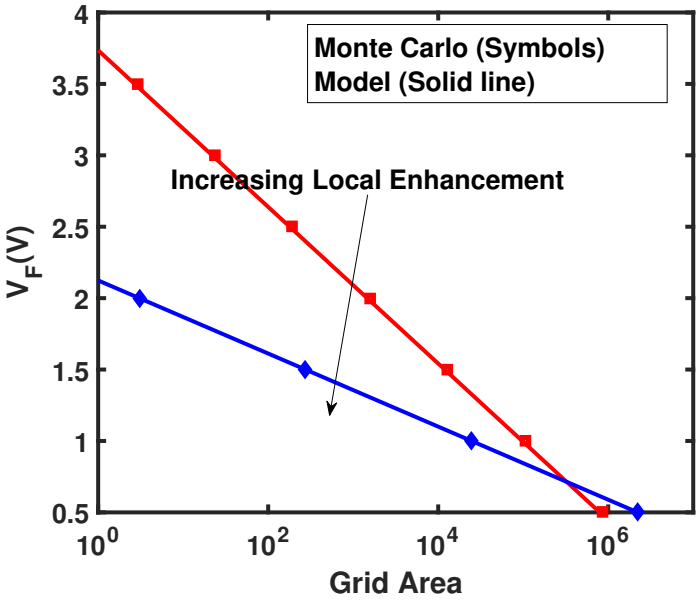

Figure 4.8: Grid Area vs. Forming Voltage. Impact of the local field enhancement. Monte Carlo simulation results (symbol) and Analytic model (solid).

4.4 Conclusions

In this chapter, impact of the Transition Metal Oxide device parameters on the forming voltage was investigated. An analytic model is first developed. The proposed model is based on the reaction rate equation and uses reliability modeling techniques to arrive at a closed form expression. The model was validated against experimental data drawn from multiple sources in the literature and shows decent results. The proposed model is also validated against Monte Carlo simulation for further validation. It is shown that the forming voltage linearly decreases with increasing oxide thickness while it logarithmically decreases with increasing oxide area. Local field enhancement can play a key role in reducing the forming voltage. This model can be used by device designers to identify process parameters that can be changed in order to lower the forming voltage.

Chapter 5

Circuit Techniques for Robust and Energy Efficient Synaptic Arrays in Neuromorphic Systems

Neuromororphic systems based on memristive devices have emerged as promising hardware platforms for implementing emerging computing architectures. A typical neuromorphic system consists of a synaptic array wherein synaptic elements are organized in a crossbar structure and neurons connected to the rows and columns of the synaptic array [17]. Several studies have shown that device and circuit level constraints may impact the performance of the neuromorphic crossbar array. In [73, 74], effect of the selector non-linearity on the performance of the neuoromorphic array was investigated. Impact of the IR drop on neuromorphic arrays was also investigated in [75, 76, 77, 78] and system level solutions were provided. This work, however, focuses on providing design solutions at the circuit level to mitigate the effects of crossbar array parasitics. Two parasitic sources are identified at the circuit level, namely: sneak paths (leakage paths) and line resistance [15, 79].These parasitic sources result in voltage degradation across the selected cell which ultimately hampers the performance of the synaptic array.

To this end, a multi-driver write scheme is proposed to improve voltage delivery across the selected cell. The proposed write scheme reduces the effective line resistance and number of leakage current paths thereby boosting the voltage delivery across the selected cell.

5.1 The Crossbar Array as a Synaptic Memory

Synaptic arrays adopt the same crossbar structure as memory arrays wherein each cross point represents a memory element. Unlike memory arrays where each memory element represents one bit of information, the full analog resistance range of the memory element is utilized in synaptic arrays to represent a synaptic weight. Hence, in layman terms, one can think of synaptic arrays as analog crossbar memory arrays. Synaptic weights are programmed into synaptic arrays via similar write techniques as those used in conventional memory arrays. In this work, we adopt the V/2 write scheme [13] for its lower energy consumption. Fig 5.1 depicts the synaptic array and Fig 5.2 depicts a circuit model of the synaptic array under the half bias scheme. Similar to memory arrays, the synaptic array exhibits voltage degradation across the selected cell due to (I) the voltage drop across the interconnect resistance of each segment, R_{int}, of the lines connecting the row and column drivers to the selected cell and (II) leakage current flow through the unselected cells on the same row and column of the selected cell. The combined effect of these two parasitic sources results in significant degradation in voltage delivery [80].

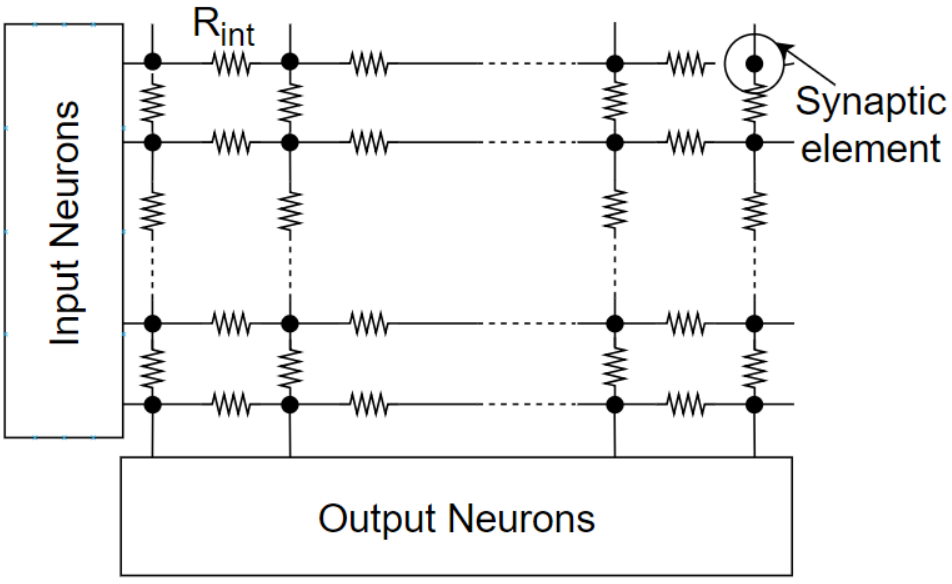

Figure 5.1: Synaptic Crossbar Array.

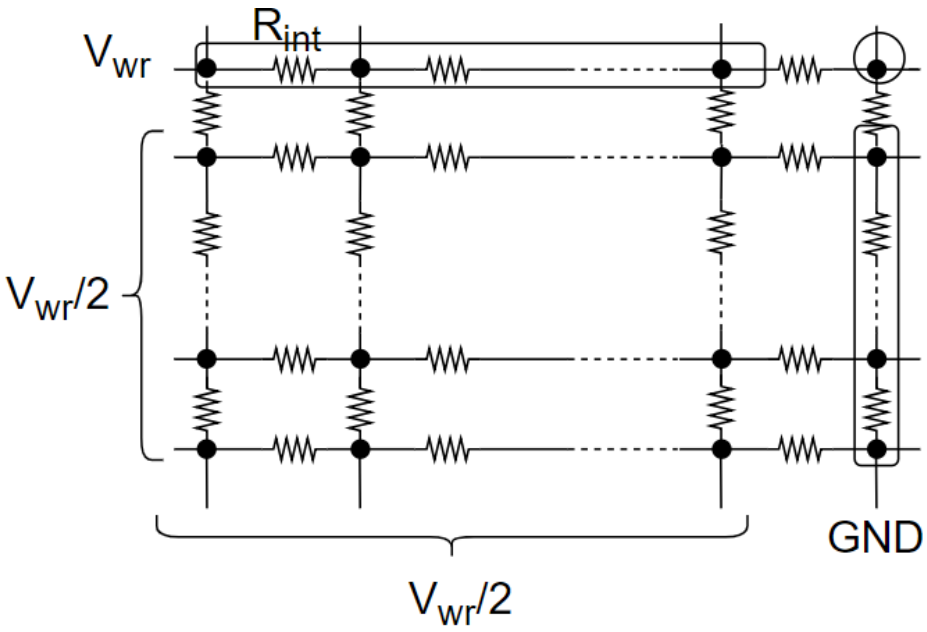

Figure 5.2: Synaptic array under the half bias scheme.

5.2 The Multi-Driver Write scheme

The write operation in the work considered herein involves writing the synaptic weights of the pre-trained network. A pre-trained network assumes that the synaptic weight values are determined offline using a software algorithm and does not require online update (i.e. during regular operation) of the synaptic weights. Due to leakage paths and line resistance, the voltage delivered to the designated synaptic element may be insufficient to effect a resistance change which could hamper the performance of the neuromorphic array. In the multi-driver design approach, the crossbar array is driven from all four sides reducing the effective leakage paths and line resistance, thus, boosting the voltage delivered to the selected cells. It is important, however, to note that the each driver is halved and distributed on both side of the array, thus, maintaining the same overall driver size. Fig 5.3 depicts the circuit model of the synaptic array under the multi-driver write scheme. The half bias scheme is assumed in this circuit model and the worst case cell is selected (the circled cell) for write operation.

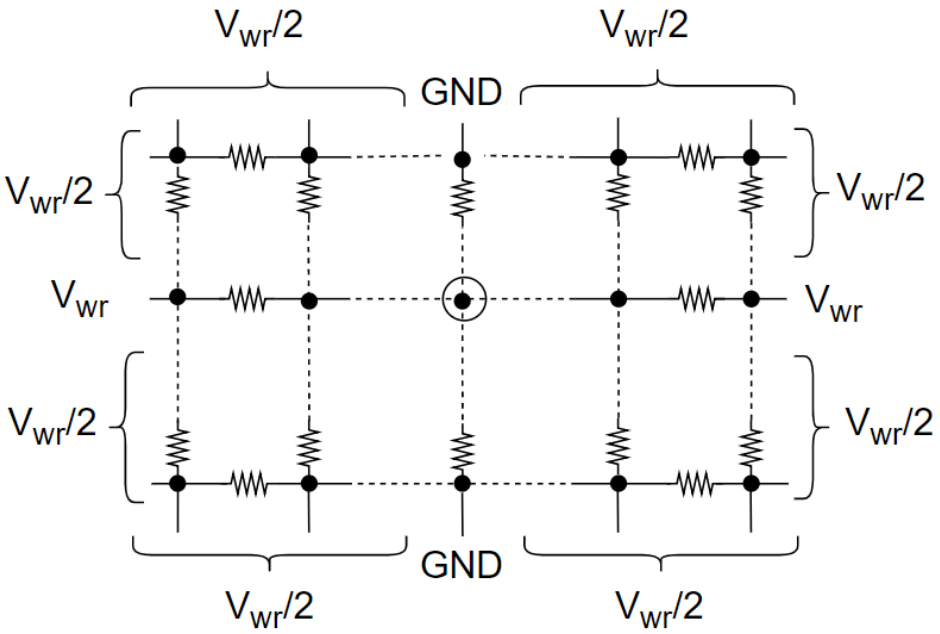

Figure 5.3: Synaptic array under the multi-driver write scheme

Fig 5.4 depicts the equivalent circuit of the crossbar array with the worst case cell being selected for write operation. R_H and R_V represent the horizontal and vertical resistances, respectively, which togther model the impact of the line resistance and leakage current. This circuit model is derived using Delta-to-Wye conversion as shown in Appendix C. It is readily shown in Fig 5.5 that as the crossbar size increases, R_H increases while R_V decreases which result in degrading the voltage delivered across the selected cell. In the multi-driver write scheme, however, the advantage is twofold: (1) only half the distance is traversed between the driver and the worst case selected cell compared to the single driver scheme and (2) two branches are driving the crossbar array in parallel which further strengthens the driving capability. In the single driver scheme, the selected cell, R_{cell}, is driven by branches A_{Row} and A_{Col}. In the multi-driver write scheme, on the other hand, branches A_{Row} and A_{Col} are connected in parallel with branches B and C, respectively. In addition, one can readily see that the equivalent driver resistance in the multi-driver scheme is halved or, alternatively, the the same driver resistance can be maintained if the driver size is halved, thus, maintaining the same overall driver area.

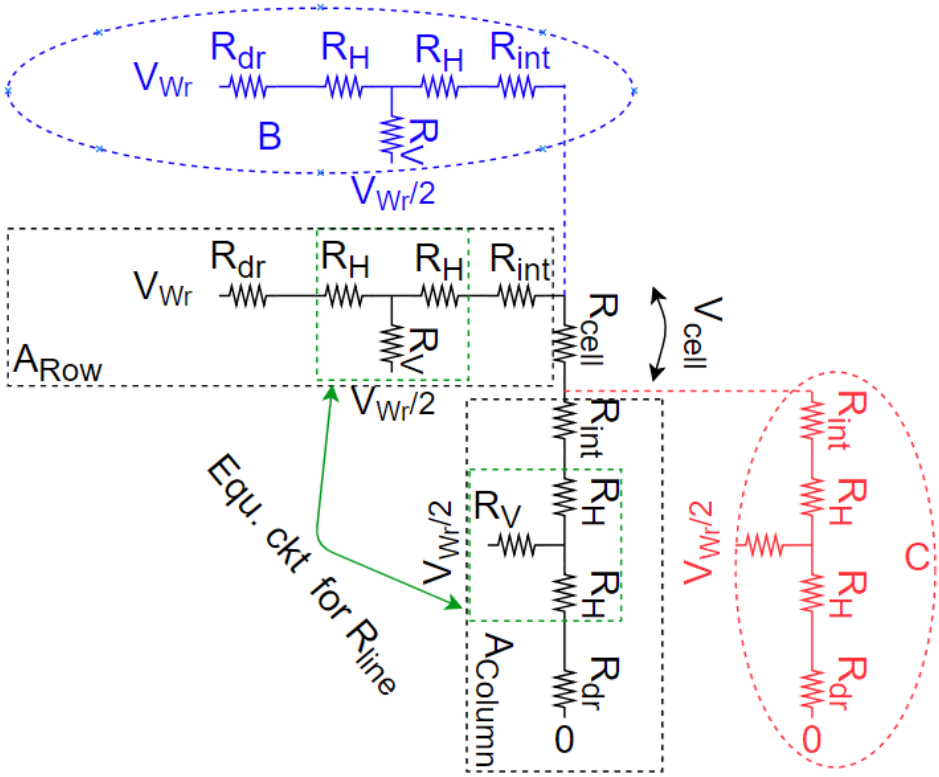

Figure 5.4: Equivalent circuit of the crossbar array under worst case cell write condition.

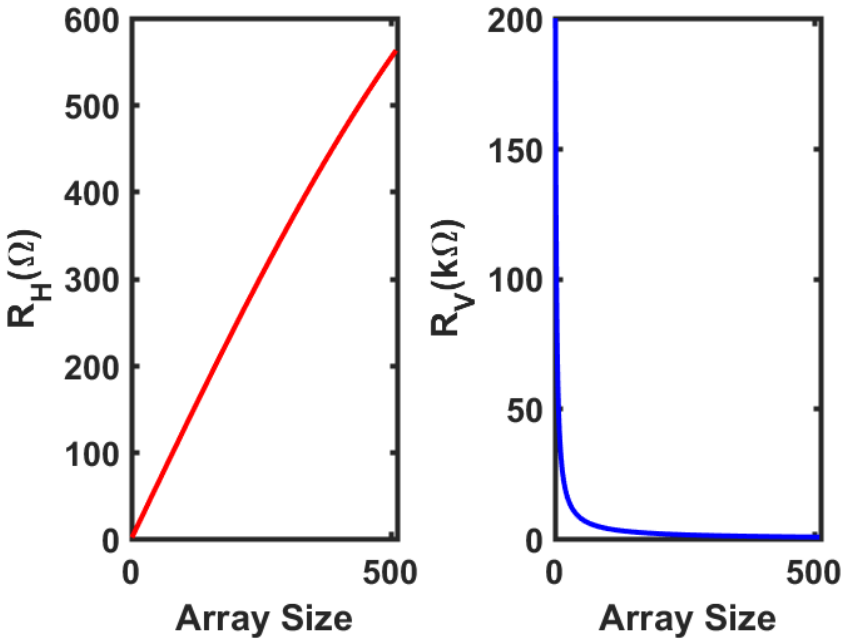

Figure 5.5: R_H and R_V vs. Crossbar array size

5.3 Performance of the Multi-Driver Write Scheme

In this section, the impact of the multi-driver design is investigated. In a synaptic array, weights are typically programmed using a sequence of pulses with constant magnitude and width until each weight reaches its designated value. The pulse magnitude and width used in this work are $1.2V$ and $100ns$, respectively. Hence, in order to evaluate the performance of the synaptic array, we use the change in resistance per spike $\Delta R/spike$ as a figure of merit. The larger the $\Delta R/spike$ the better since the designated weight values can be reached faster which eventually reduces the energy consumed during programming. It is shown in Fig 5.6 that the multi-driver approach outperforms the conventional approach. This improvement is attributed to the enhanced voltage delivery across the selected cell as shown in Fig 5.7. It is also shown that the improvement becomes more pronounced as the crossbar array size increases as the parasitic resistance begin to dominate the driver resistance.

5.4 Conclusions

This chapter investigated the impact of crossbar array parasitics on the performance of synaptic arrays in neuromorphic systems. The line resistance and leakage path current result in voltage degradation across the selected cell. A multi-driver write scheme was proposed that improves voltage delivery via reducing the effective line resistance and number of leakage current paths. This enhancement in voltage delivery leads to better performance while maintaining the same overall driver area.

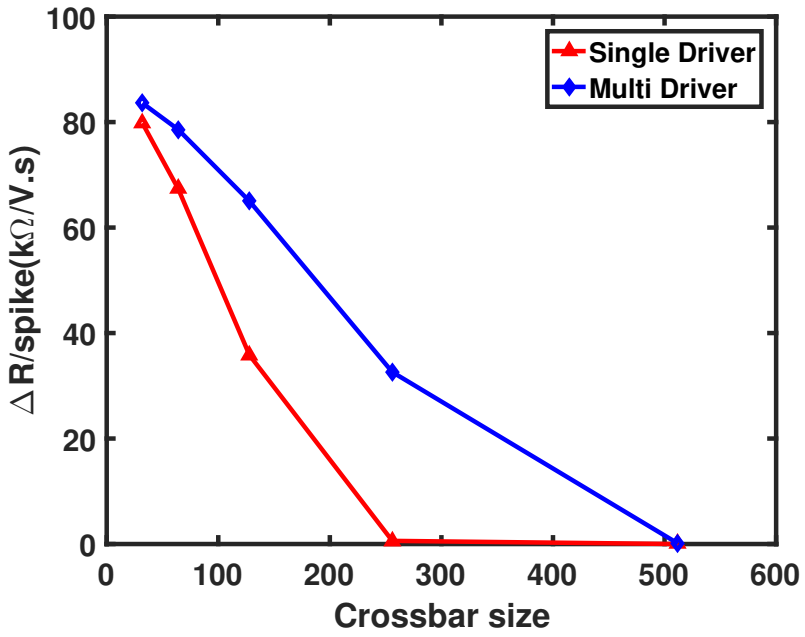

Figure 5.6: Change in resistance per spike versus crossbar array size. $V = 1.2V$ and the pulse width is $100ns$.

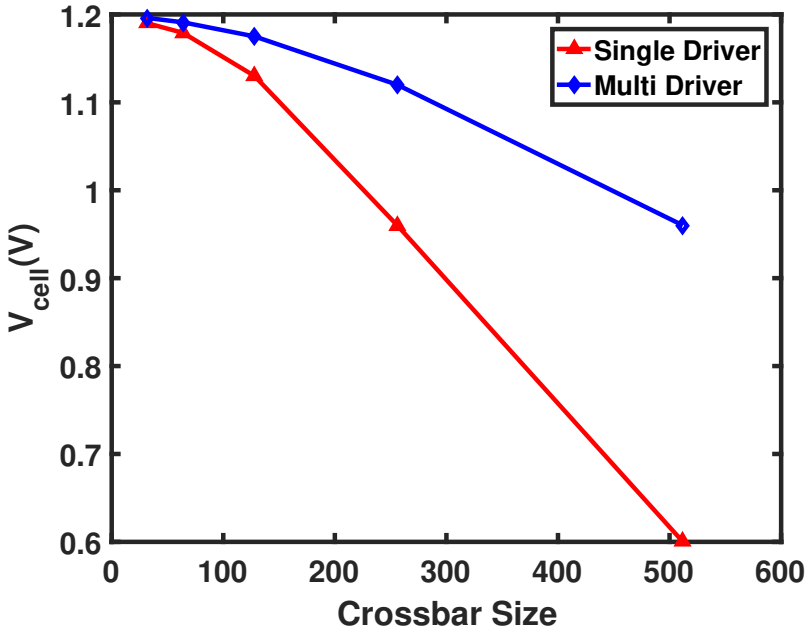

Figure 5.7: Voltage delivered across the worst case selected cell versus crossbar array size.

.

Chapter 6

Modeling Insulator Metal Transition Devices for Circuit Simulation

Insulator Metal Transition (IMT) devices have attracted significant interest in the research community [81, 28] owing to their switching dynamics that have shown to be suitable for applications such as neuromorphic circuits and memory arrays. The high ON/OFF ratio of IMT devices makes them good candidates for use as selector devices in memory arrays. In addition, their Back-End-Of-Line compatibility (BEOL) helps achieve the ideal $4F^2$ density of crossbar arrays [82, 83]. On the neuromorhpic front, it has been shown that IMTs can be used in building Integrate-And-Fire neurons while alleviating the need for complex CMOS circuitry, thus, providing a significant density advantage [32, 23].

A decent body of work has been presented on IMT devices trying to understand the underlying physical mechanisms contributing to resistive switching. Several studies have shown that temperature is the main cause of resistive switching such as the work in [57, 58] while others have attributed it to the electric field [59] with temperature playing a collateral role. A more in depth study about the switching mechanism is presented in [60, 61] which show that Joule heating may not be sufficient for resistive switching and an electric field assisted switching is more plausible. The authors in [60] hypothesize that a certain threshold voltage is required to effect a phase transition which decreases with increasing temperature. In [84], the authors have classified IMT devices into two categories: Thermally-driven IMT (T-IMT) and Electronic IMT (E-IMT) and the characteristics of each type has been studied.

The lack of a physics-inspired SPICE model, however, has prohibited the full exploration of IMT devices in circuit applications. Specifically, understanding the interplay between temperature and electric field has been the main barrier to the development of such a model [84]. In [32], an IMT electro-thermal model was developed that leverages the positive electro-thermal feedback to effect a phase transition of the device. The model was validated against vanadium oxide (VO_2) experimental data and could reproduce the data with sufficient accuracy. The model in [62] followed similar lines and is based on Mott insulator theory.

In this chapter, an IMT SPICE model is proposed and implemented in Verilog-A. The proposed model describes the IMT device as a memristive system with the local temperature of the device acting as the internal state variable. Spectre from Cadence is used to simulate the model and shows a close match to experimental data and device simulations based on the models in [32, 62].

6.1 Background

Two IMT device models were proposed in [32, 62]. The model in [32] used an elaborate thermal model to capture the temperature evolution of the IMT device along with a look up table that captures the relationship between the device's resistance and temperature. This look up table based approach is not very popular in SPICE models which are often represented in a closed form functional form.

The model in [62] built off the previous model in [32] but presented a more physical picture based on band theory. The IMT device is modeled as a low bandgap semiconductor where the bandgap of the device decreases with increasing temperature. This reduction in the bandgap increases the carrier concentration which ultimately results in decreasing the device resistance. A model is also presented which captures the change in the thermal conductivity with temperature. Both models are then solved in a self consistent fashion to effect a phase transition as a function of temperature. This model, similar to the previous one, is best used in a TCAD simulation flow and is not well-suited for SPICE level simulators.

The SPICE model proposed in this work build off both models while employing some simplifications to enable its seamless implementation in SPICE environment. A lumped

element thermal model is used to describe the temperature dynamics of the device. An empirical function is also employed to describe the evolution of the device's resistance with temperature.

6.2 The proposed IMT SPICE model

IMT's resistive switching has been attributed to the interplay between the electric field applied across the device and the change in the device's local temperature resulting from Joule heating. As the current flows through the device, the device's temperature rises until it reaches a critical temperature at which point the device's resistance switches from a high resistance state to a low resistance state. The resistance relaxes back to its initial high resistance state as the temperature of the device drops below the critical temperature.

Here we leverage the memristor theory [2, 10, 11] to describe the IMT device. The memristive dynamics of the IMT device can be described as follows [12]:

$$I = G(x).V, \tag{6.1}$$

$$\frac{dx}{dt} = g(x, V), \tag{6.2}$$

where (6.1) and (6.2) describe the output and state equations, respectively, and x is the internal state variable. The proposed model has two main governing equations: (I) the resistance change equation that corresponds to the output equation (here the resistance is used for modeling convenience) and (II) the temperature evolution equation that corresponds to the state equation, with the temperature being the state variable such that $x = T(t)$.

The relationship between the device's resistance and local temperature is captured by two empirical functions: (I) a thermistor function that captures the resistance evolution with temperature before and after switching and (II) a sigmoid function that captures the device's switching. The two thermistor states are expressed as exponential functions of the temperature such that $R_{LRS} = R_{LRS_F} e^{-B_{LRS}(T(t)-T_F)}$ and $R_{HRS} = R_{HRS_0} e^{-B_{HRS}(T(t)-T_0)}$. R_{LRS_F} is the low resistance state defined at temperature T_F (a reference temperaure) and R_{HRS_0} is the high resistance state defined at the ambient temperature T_0. B_{LRS} and B_{HRS}

are the temperature coefficients which are extracted from the slope of the thermistance vs. temperature plot and the negative sign describes Negative Temperature Coefficient (NTC) thermistors. This implementation, however, requires clipping of the R_{LRS} and R_{HRS} at some minimum and maximum values to avoid any unphysical behavior during circuit simulation. Clipping, however, requires the use of conditionals which hamper the "smoothness" of the model yielding potential convergence difficulties during circuit simulation. Hence, the model equations are reformulated such that R_{LRS} and R_{HRS} smoothly plateau to R_{LRS_F} and R_{HRS_0} at high and low temperatures, respectively.

This relationship between the temperature and the resistance can be expressed as follows:

$$R_{LRS} = R_{LRS_F}(1 + {K_{LRS}}^{A})^{\frac{1}{A}}, \tag{6.3a}$$

$$R_{HRS} = R_{HRS_0}(\frac{K_{HRS}}{(1 + {K_{HRS}}^{A})^{\frac{1}{A}}}), \tag{6.3b}$$

$$R_{IMT} = R_{LRS} + \frac{(R_{HRS} - R_{LRS})}{1 + e^{\frac{T(t)-T_c}{T_x}}}, \tag{6.3c}$$

where $K_{LRS} = e^{-B_{LRS}(T(t)-T_F)}$ and $K_{HRS} = e^{-B_{HRS}(T(t)-T_0)}$. T_x is a fitting parameter that captures the sharpness of the resistive transition. T_c is the critical temperature which is around 340K in the case of VO_2 devices [23]. R_{LRS} and R_{HRS} are the Low Resistance State and High Resistance State, respectively. A is a conrol parameter which governs how the two thermistor states approach the asymptotes [85]. In this work, $A = 10^4$ is used. However, this parameter can be varied by the user as needed. While the model might seem complicated at first glance, the principal equations are simple exponential functions as aforementioned. This formulation is only employed to abide by compact modeling practices as suggested in [86, 85]. Section 6.3 provides a more thorough explanation for (6.3) and describes the parameter extraction procedure.

The temperature evolution dynamics are described by the lumped element thermal model presented in [3] as expressed in (6.4):

$$C_{th}\frac{dT(t)}{dt} = V_{IMT}I_{IMT} - \frac{(T(t) - T_0)}{R_{th}}, \tag{6.4}$$

where $V_{IMT}I_{IMT}$ is the Joule heating, C_{th} and R_{th} are the effective thermal capacitance and the effective thermal resistance, respectively, and T_0 is the ambient temperature.

Listing 6.1 depicts Verilog-A code snippet of the core model equations. Suggested compact modeling techniques are adopted based on the work in [87, 88]. Expressions (6.3a) and (6.3b) are each divided into two expressions. This model formulation helps avoiding numerical overflow as the values for K_{LRS} and K_{HRS} become significantly large; see [85] for more elaborate discussion on this point. The expressions in each conditional are, however, identical and, therefore, do not cause any discontinuities during the model's execution.

Listing 6.1: Verilog-A code snippet

```
Iwr = I(p,n);
 V(p,n) <+ Iwr*Rm;
K_HRS=exp(-B_HRS*(tem-T_0));
K_LRS=exp(-B_LRS*(tem-T_F));
if (tem>T_F) begin
   LRS=LRSF*pow((1+pow(K_LRS,A)),1/A);
end
else begin
   LRS=LRSF*K_LRS*pow((1+pow(K_LRS,-A)),1/A);
end

if (tem>T_0) begin
   HRS=HRS0*K_HRS/(pow((1+pow(K_HRS,A)),1/A));
end
else begin
   HRS=HRS0/(pow((1+pow(K_HRS,-A)),1/A));
end
Rm= LRS + (HRS-LRS)/(1+exp((tem-Tc)/Tx));
Pwr(temp)<+ ddt(Temp(temp));
Pwr(temp)<+ -pow(Iwr,2)*Rm/Cth;
Pwr(temp)<+(Temp(temp)-T0)/(Rth*Cth);
```

6.3 IMT Model Parameter Extraction

In this section, we develop the parameter extraction procedure for the proposed model. The parameter extraction procedure extracts the model parameters in expressions (6.3) and (6.4) representing the output and state equations, respectively.

6.3.1 Extracting Model Parameters for the Output Equation

Parameters B_{HRS} and B_{LRS} are first extracted from the output equation. The simplified form of equation (6.3) is used in the extraction procedure, described in section III, which can be expressed in the following form:

$$ln(R_{LRS}) = ln(R_{LRS_F}) - B_{LRS}(T(t) - T_F), \tag{6.5}$$

$$ln(R_{HRS}) = ln(R_{HRS_0}) - B_{HRS}(T(t) - T_0), \tag{6.6}$$

four data points are then used to extract the thermal coefficients (B_{HRS} and B_{LRS}) as shown in Fig 6.1. The thermal coefficients can be expressed as follows:

$$B_{LRS} = \frac{ln(R_{LRS_F}) - ln(R_{LRS})}{T_2 - T_F}, \tag{6.7}$$

$$B_{HRS} = \frac{ln(R_{HRS_0}) - ln(R_{HRS})}{T_1 - T_0}, \tag{6.8}$$

As alluded to before, the value chosen for A in this work is $A = 10^4$. The higher the value of A, the faster R_{HRS} and R_{LRS} saturate to R_{HRS_0} and R_{LRS_0} beyond T_0 and T_F, respectively.

The sigmoidal function contains only one parameter, T_x which captures the steepness of the resistive transition about the critical temperature. This parameter can be formally extracted from the line slope of at the transition temperature are simply chosen such the the model accurately fits the the experimental data.

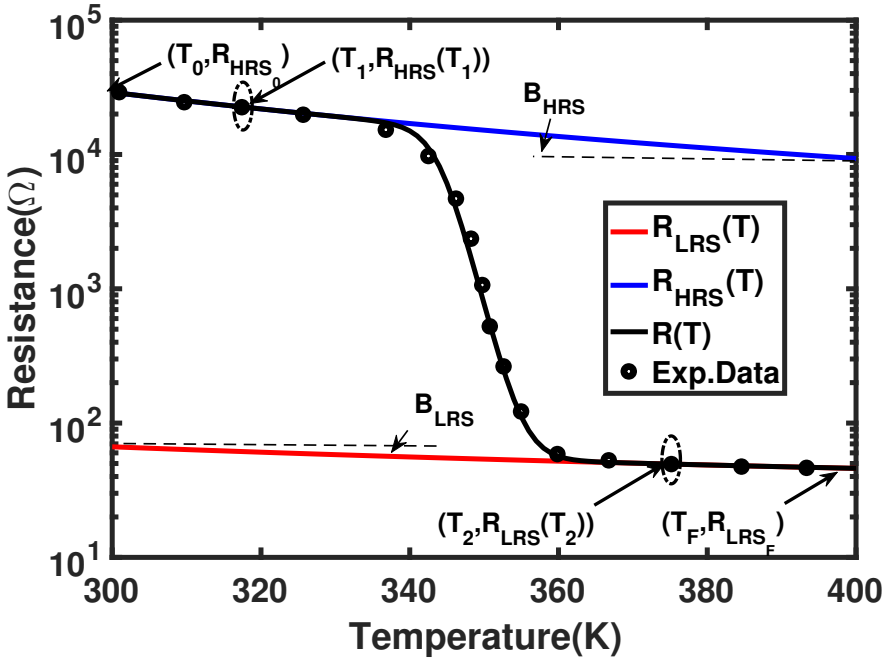

Figure 6.1: Thermal coefficients extraction

6.3.2 Extracting Model Parameters for the State Equation

Parameters R_{th} and C_{th} are then extracted from the state equation. Fig. 6.2 depicts the temperature evolution with time. Note that the temperature does not reach the critical temperature and, accordingly, the IMT electric resistance R_{IMT} does not switch. Hence, we chose this specific curve for parameter extraction since the electrical resistance of the device is constant throughout the simulation. First, the steady state solution of the temperature equation is solved and expressed as follows:

$$T_{SS} = T_0 + R_{th}\frac{V_{IMT}^2}{R_{IMT}}, \tag{6.9}$$

Knowing the applied voltage across the IMT device V_{IMT} and the device's resistance R_{IMT}, R_{th} can be extracted. C_{th} is then extracted from the transient solution to fit the curve. The transient solution for the temperature evolution can be expressed as follows:

$$T(t) = T_0 + R_{th}\frac{V_{IMT}^2}{R_{IMT}}(1 - e^{\frac{-t}{R_{th}C_{th}}}), \tag{6.10}$$

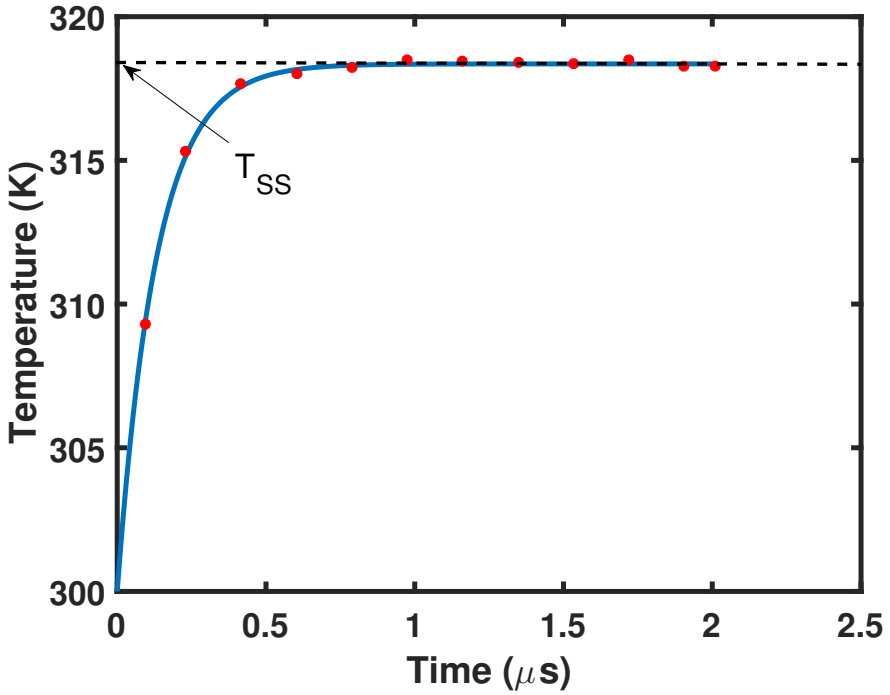

Figure 6.2: Thermal resistance and capacitance extraction

6.4 Model Validation

The proposed model is validated against experimental results drawn from [32]. Fig. 6.3, 6.4 and 6.5 depict the resistive transition about the critical temperature which is about $340K$ in VO_2 devices for three device samples with different high and low resistance states. Fig. 6.6 depicts the hysteresis in the V-I plane (a fingerprint of memristive systems) exhibited by the IMT device as shown in [23, 32] and fitted against the experimental data from [32].

Figures 6.7 and 6.8 depict the time dependence of temperature and resistance evolution, respectively, fitted against electrothermal simulations from [32]. Three voltage levels, based on the values used from [32], were applied across the device: $1.4V$, $1.6V$ and $1.8V$. One can readily observe in Fig 6.7 that the local temperature of the device saturates at a higher temperature value for higher voltages due to increased Joule heating. In Fig 6.8, higher voltages result in faster transition time due to faster rate of joule heating.

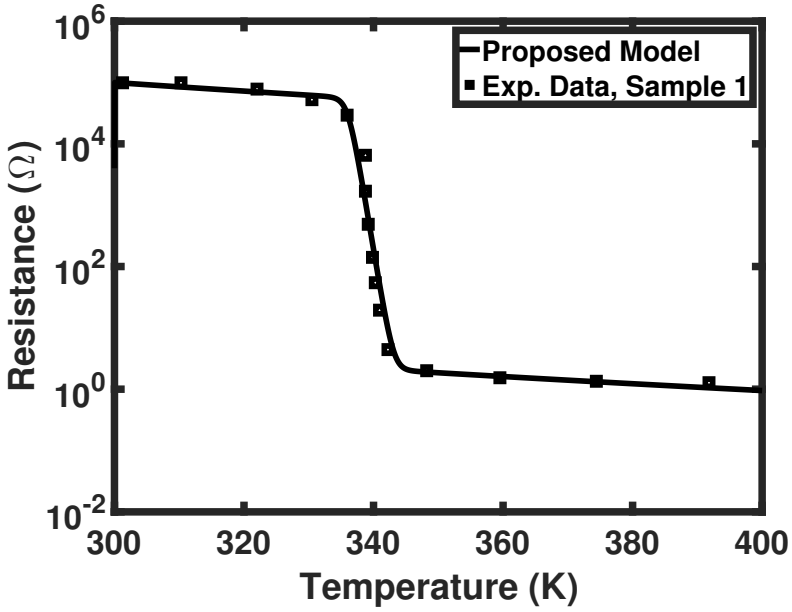

Figure 6.3: Model fitting against experimental data

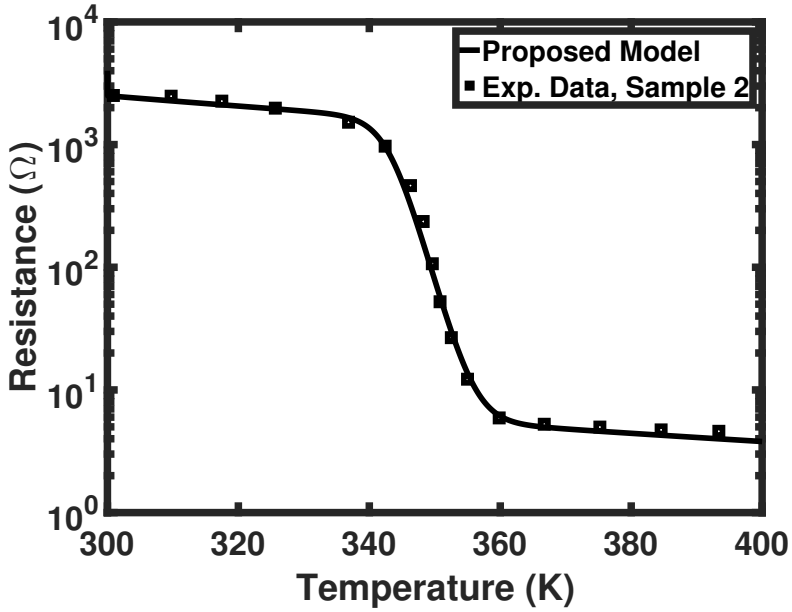

Figure 6.4: Model fitting against experimental data

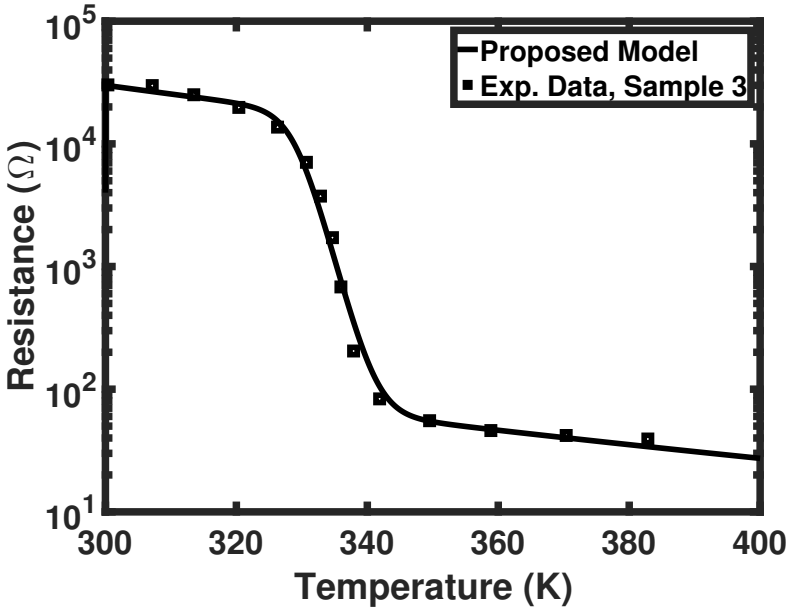

Figure 6.5: Model fitting against experimental data

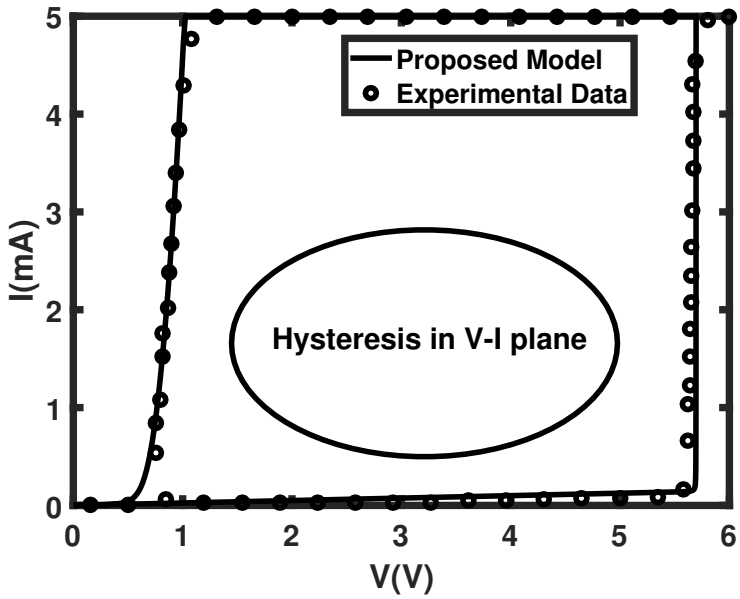

Figure 6.6: Demonstration of device Hysteresis in the V-I plane which demonstrates the memristive dynamics of the IMT device.

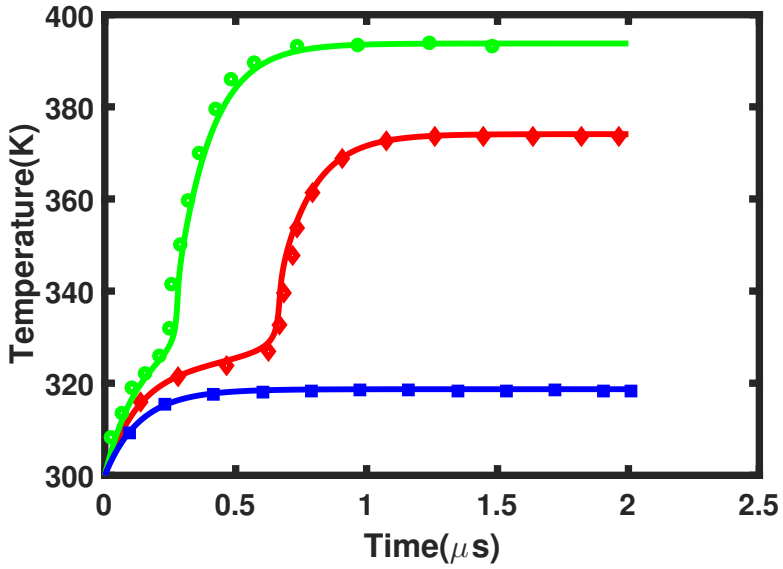

Figure 6.7: Model (solid line) fitting against electro-thermal simulations (markers). Plotting the Device local temperature against time for three applied voltage values.

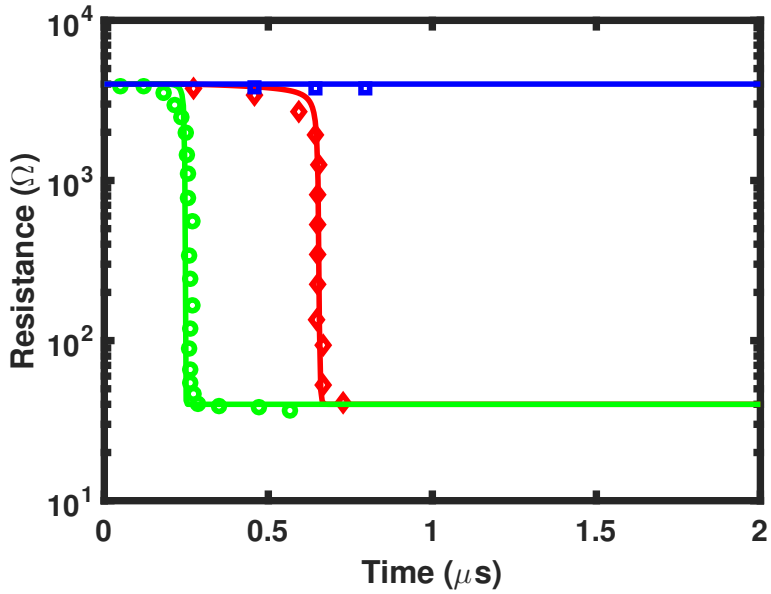

Figure 6.8: Model (solid line) fitting against elector-thermal simulations (markers). Plotting the Device resistance against time for three applied voltage values.

6.5 Discussions and Future prospects

The values selected for R_{HRS} and R_{LRS} in this work are extracted from the devices used in [32]. These values are considered relatively low and may not be compatible with many of the applications that employ IMT devices. For example, in memory arrays, IMT devices are often used as selector elements to circumvent sneak path current. This requires the selector device to possess a high on/off ratio which may not be accomplished with the resistance values used in this work. In neuromorphic arrays, on the other hand, IMTs are used as neurons. The neuron firing relies on the IMT switching from R_{HRS} to R_{LRS} as the local temperature of the device exceeds the critical temperature by means of Joule heating. This condition, however, may not be achieved if the R_{HRS} is high as will be shown on the next chapter.

Models for IMT devices also still require constant refinements as more investigation of these devices are conducted. The lack of a complete understanding of their switching dynamics necessitate the inclusion of a significant empirical content in IMT models to capture the observed behavior, specifically, compact models. This empiricism, however, hinders the model's predictability and scalability.

6.6 Conclusions

This work presented a SPICE compatible compact model for Insulator Metal Transition devices validated against experimental data and electrothermal simulations from the literature. The proposed model describes the IMT device as a memristive system and captures the role of temperature and electric field in the resistive transition of the device. A lumped element thermal model was employed to capture the temperature evolution of the IMT device resulting from Joule heating and an empirical model was developed to capture the functional relationship between the device's resistance and temperature. This model can be used by circuit designers who wish to explore the use of IMT devices in designing nanoelectonic circuits in SPICE environment.

Chapter 7

Design of Insulator Metal Transition based Integrate And Fire Neurons

The switching dynamics of Insulator Metal Transition (IMT) devices can be leveraged in realizing Integrate And Fire (IAF) neurons. An IMT-based oscillating neuron was proposed in [23] along with a behavioral IMT model used to simulate the neuron. The proposed model, however, did not capture the role of temperature in IMT switching and, accordingly, did not study how the temperature dependent resistive switching can be leveraged to deliver the functionality of an IAF neuron. In [32], an IMT-based IAF neuron was designed that captures the temperature dynamics of the device. That design, however, is basic and was studied in isolation without insight into how the proposed neuron circuit can be included in a fully fledged neuromorphic crossbar array.

This work proposes an IMT-based IAF neuron [89]. The proposed design capitalizes on the design proposed in [32] and introduces an output buffer to enable driving other stages. Impact of the IMT devices parameters on the operation of the IAF neuron is also studied. It is shown that, unlike CMOS neurons, the properties of the IMT neuron are dependent on the IMT device parameters.

7.1 The IAF and LIF Neurons

IAF neurons are characterized by three phases of operation: (I) accumulation, (II) fire and (III) refractory period. The neuron integrates the incoming current from the synaptic network $I_{in} = \Sigma_i V_i G_i$ during accumulation and, consequently, the membrane potential V_{mem} increases. The neuron then fires should V_{mem} exceed some threshold voltage V_{th} and V_{mem} is reset to its resting potential. In the case of ideal IAF neurons, if no inputs are accumulated from the synaptic network, the membrane potential remains unchanged and no charge is leaked as predicted by the governing equation $I_{in} = C_{mem}\frac{dVmem}{dt}$. Biological neurons, however, are leaky due to their finite membrane resistance R_{mem} and their governing equation can be expressed as follows: $I_{in} - V_{mem}/R_{mem} = C_{mem}\frac{dV_{mem}}{dt}$. This type of IAF neurons is often referred to as Leaky IAF neuron or simply: LIF. Fig 7.1 depicts the equivalent circuit of both IAF and LIF neurons. Linear circuit techniques such as Thevinin and Norton equivalent circuits were employed to develop these circuit models for the integration stage. Fig 7.2 depicts the SPICE simulation of the schematic in Fig 7.1. Current pulses were fed into the neuron from an ideal current source. Both neurons experience an increase in the membrane potential V_{mem} when a current pulse is fed to the neuron input. When the neuron is idle, however, V_{mem} in the IAF neuron remains constant while it drops gradually (leaks) in the case of the LIF neuron. This leakage can be readily explained by the RC discharge present in the equivalent circuit of the LIF neuron.

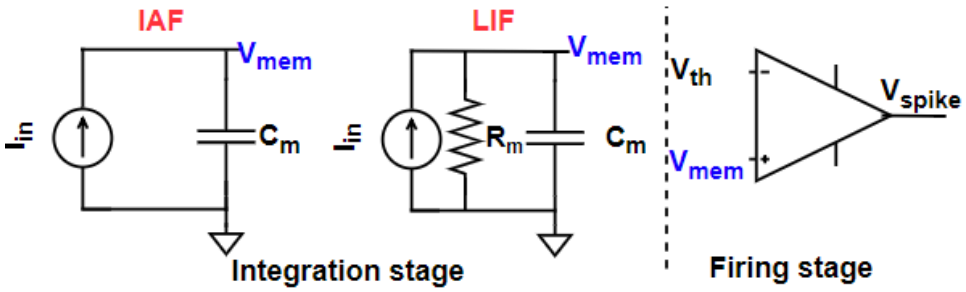

Figure 7.1: Equivalent circuit models of IAF and LIF neurons

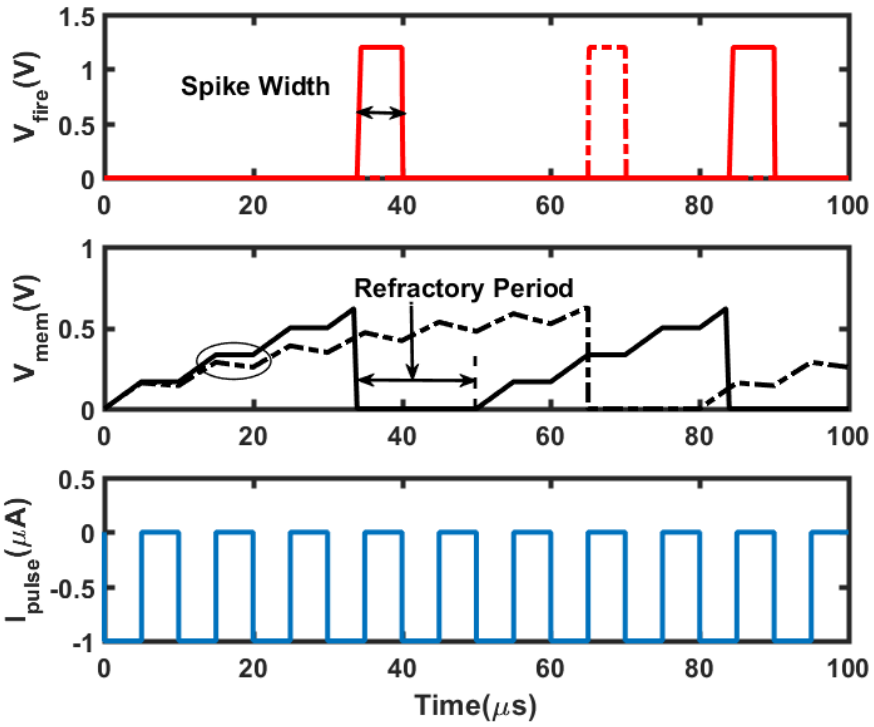

Figure 7.2: Simulation of IAF (solid) and LIF (dashed) neurons. V_{mem} decreases in LIF neurons when no spike arrive while it remains constant in the case of IAF neurons (circled).

7.2 Proposed IMT-based IAF Neuron

The IMT IAF neuron presented in this work leverages the switching physics of the IMT device to deliver the functionality of a spiking neuron. Neuromorphic circuits consist of input neurons, a synaptic network organized in a crossbar structure and output neurons. Non-volatile resistive devices are often used as synaptic devices along with selector devices to boost the cell non-linearity and suppress sneak path current. In this specific design, selectors are also required to prohibit any current flow back from the neuron to the synaptic network. This can be accomplished by employing either FET transistors or diodes as selector elements.

This work focuses on the analysis and design of the IMT neuron and, hence, we consider one column of the synaptic crossbar array driving the IMT neuron. Fig. 7.3a depicts a crossbar array neuromorphic structure and Fig. 7.3b presents the circuit under consideration.

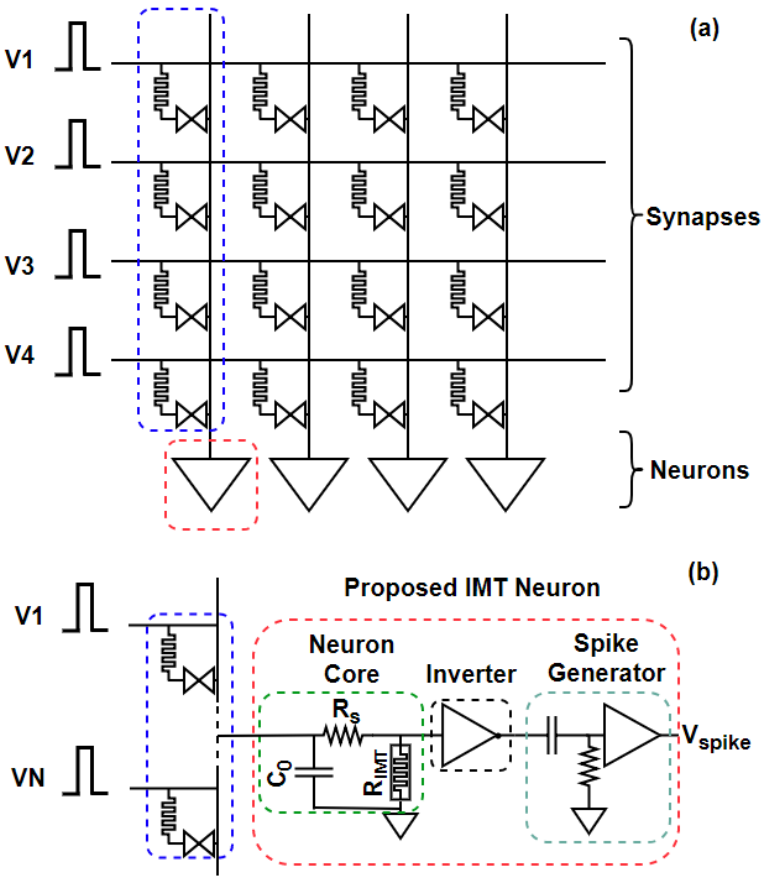

Figure 7.3: schematic of the proposed IMT IAF neuron. (a) A synaptic netwrok driving neurons. (b) Circuit under consideration.

CMOS neurons are typically comprised of two operational amplifiers [90, 91]. The first acts as an integrator to integrate the incoming current from the synaptic network while the other acts as a comparator to compare the membrane voltage with the threshold voltage. A feedback circuit is usually employed to implement the refractory period which prohibits the neuron from accumulating the incoming current from the synaptic network.

The proposed IMT neuron, unlike CMOS neurons that process voltage information, accumulates, compares to a threshold and fires, and implements the refractory period via processing thermal information. The neuron's thermal dynamics, however, are still dependent on the applied electrical signals since they're governed by Joule heating. This electro-thermal coupling is what distinguishes IMT neurons from conventional CMOS neurons and will be discussed in more detail in the next section. The core of the neuron lies in the parallel combination of the IMT device and the capacitor as depicted in Fig 7.3b.

As the input spikes are fed to the neuron, the device's temperature gradually increases due to Joule heating, thereby, accumulating inputs from the synaptic network. The IMT then switches from high resistance state to low resistance state once the temperature exceeds the critical temperature, thereby, firing. Lastly, the neuron does not accumulate inputs from the synaptic network until it relaxes back to the high resistance state, thereby, implementing a refractory period.

Fig. 7.4 depicts the simulation of the proposed IMT neuron. In this simulation, two identical input voltage pulses were fed to the neuron with magnitude $2V$. The device's temperature rises as more pulses are fed to the neuron until it reaches the critical temperature. At this point, the IMT device switches from a high resistance state to a low resistance state and a current spike is generated. To enable the neuron to drive other stages, an output buffer needs to be employed. The output buffer is comprised of two blocks: a CMOS inverter and a spike generator circuit. The CMOS inverter converts the current spike into a voltage pulse. The spike generator circuit then modulates the pulse width of the spike depending on the designer's choice.

7.3 Impact of Device parameter on the Properties of the IMT Neuron

As mentioned earlier, the dynamics of the IMT neuron are governed by electro-thermal coupling. First, the circuit in Fig 7.3b is simplified to the circuit shown in Fig 7.5. The input voltage pulses and the input synapses are represented by an equivalent voltage source $V_{in}(t)$ and an equivalent synapse resistance R_w. These simplifications enable the development of closed form expressions that help provide insight into the properties of the IMT neuron. Similar to the analysis conducted on the ideal neuron, The equivalent circuit models developed in this section are based on Thevenin and Norton equivalent circuits and assumes quasi-static conditions (i.e. the analysis is conducted when the resistance is at either HRS or LRS without considering the transition state) to ensure that the analyzed circuits are linear.

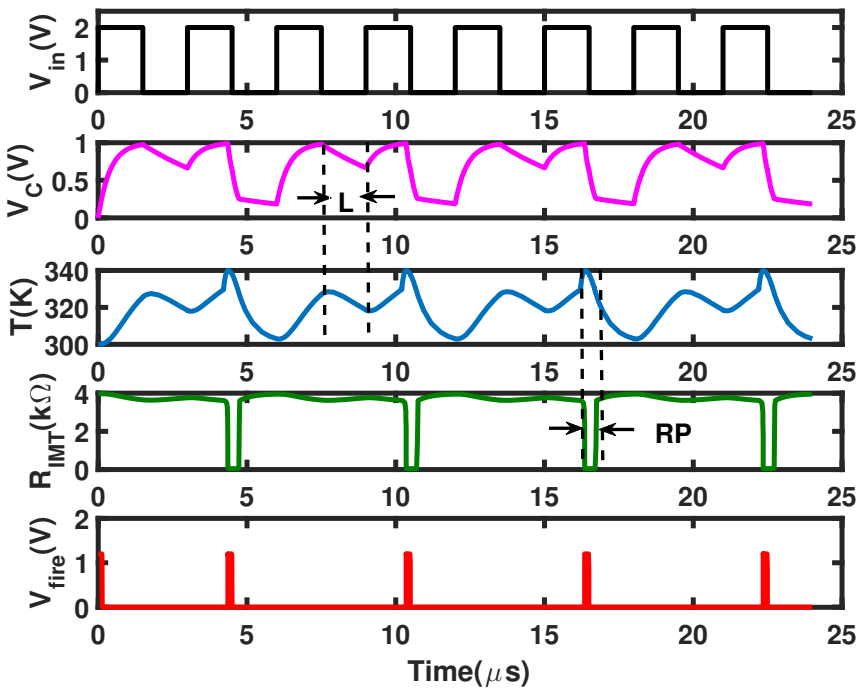

Figure 7.4: Simulation of the proposed IMT IAF neuron. L denotes the neurons leakage and RP denotes the refractory period.

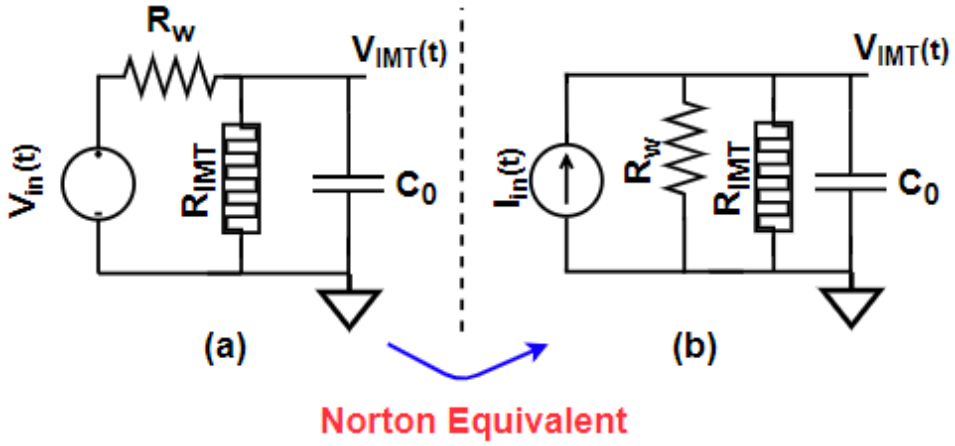

Figure 7.5: Schematic of the simplified IMT Neuron.

To enable better understanding of dynamics of the proposed IMT neuron, equivalent circuit models were derived as shown in Fig 7.6. These models help describe the coupling taking place between thermal and electrical domains during the neuron's operation. The electrical circuit model is the Norton equivalent of the circuit in Fig. 7.5 and can be described by the following differential equation:

$$C_0 \frac{dV_{IMT}(t)}{dt} = I_{in} - \frac{V_{IMT}(t)}{R_{IMT}//R_w}, \tag{7.1}$$

such that $I_{in} = V_{in}/R_w$. The thermal circuit model, on the other hand, can be described by the lumped element thermal model presented in expression (7.2).

7.3.1 Accumulation

This analysis applies to both accumulation and idle states. In these phases, $R_{IMT} = HRS$. The IMT neuron is inherently leaky as shown on Fig 7.4. Leakage, however, takes place in both thermal and electrical domains. Leakage in the thermal domain results from the finite thermal resistance R_{th} of the IMT device. In the electrical domain, on the other hand, leakage results from the parallel combination of the synaptic network equivalent resistance R_w and the IMT device resistance: $R_{IMT}//R_w$.

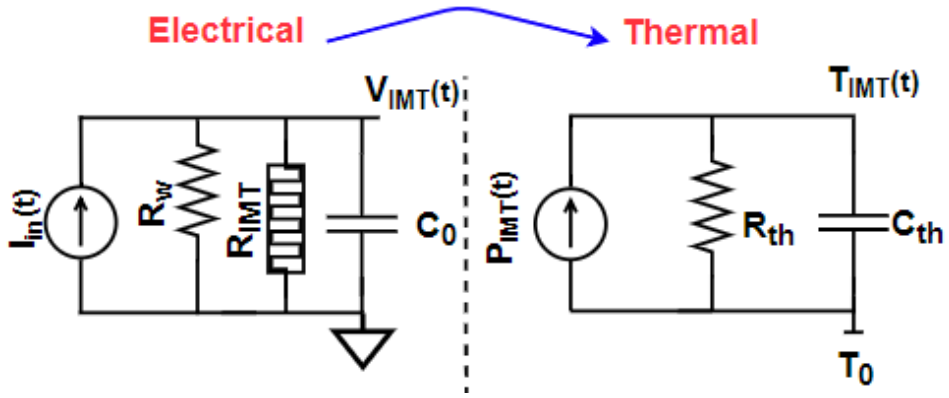

Figure 7.6: Equivalent circuit model of the simplified IMT neuron. The electrical model is the Thevenin equivalent circuit of the simplified neuron shown in Fig 7.5. The thermal model is based on equation (2).

7.3.2 Firing

In this phase, $R_{IMT} = HRS$. The IMT neuron fires when the device's temperature exceeds the critical temperature. Hence, the critical temperature here serves as the threshold voltage in the case of CMOS neuron. To enable fair comparison with CMOS neuron, however, the neuron's threshold is expressed in terms of a threshold voltage as shown in the following equations:

$$V_{IMT}(acc) = \frac{HRS}{HRS + R_w} V_{in}, \tag{7.2}$$

The steady state temperature can be expressed as follows:

$$T_{IMT} = T_0 + R_{th} \frac{V_{IMT}^2}{HRS}, \tag{7.3}$$

Substitute $T_{IMT} = T_c$ and $V_{IMT} = V_{th}$ and, thus, the threshold voltage can be expressed as:

$$V_{th} = \sqrt{\frac{HRS}{R_{th}}(T_c - T_0)}, \tag{7.4}$$

Therefore, the neuron fires when $V_{IMT}(acc) > V_{th}$. The neuron's threshold in IMT based neurons is a function of the IMT device parameters and cannot be controlled by the designer.

The neuron's firing rate also, similar to CMOS neurons, increases with increasing the current coming from the synaptic network as shown in Fig. 7.7. Unlike CMOS neurons, however, the relationship is not linear due to the serial connection of R_w and R_{IMT}.

7.3.3 Refractory Period

In this phase, the device's resistance is at LRS and the capacitor does not accumulate any charge. The device's temperature, accordingly, starts to decrease as the leakage mechanism dominates Joule heating. It is critical, however, that the steady state temperature at LRS drops below the critical temperature for the device to relax back to HRS and allow the neuron to accumulate for the next cycles. Using steady state analysis, one can derive the following expression:

$$V_{IMT}(postfire) = \frac{LRS}{LRS + R_w} V_{in}, \tag{7.5}$$

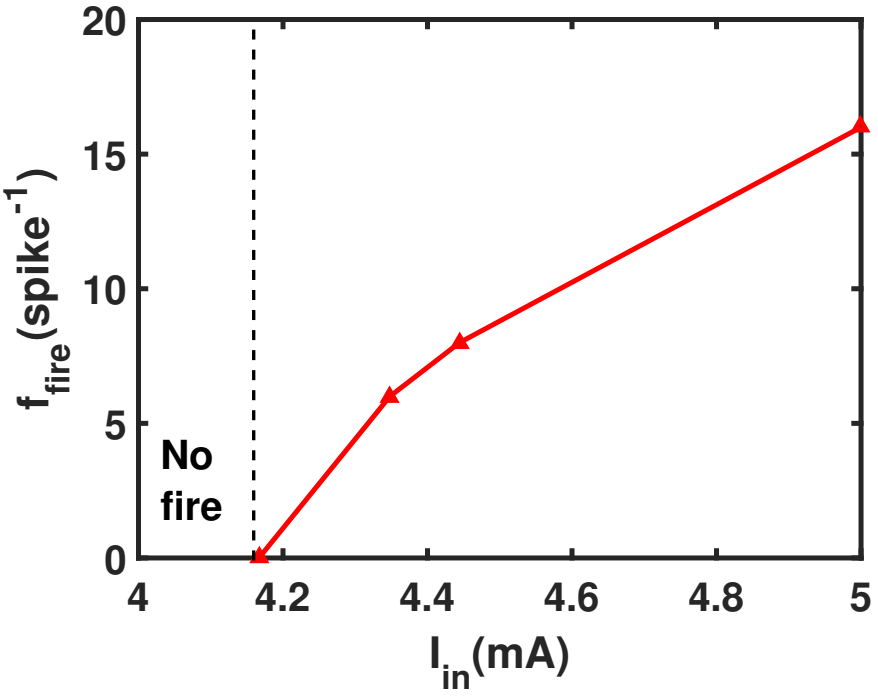

Figure 7.7: Relationship between the proposed IMT neuron's firing rate and the input current.

Therefore, to enable robust operation of the neuron, the following condition has to be met: $V_{IMT}(postfire) < V_{th}$

In addition, unlike CMOS neurons where the duration of the refractory period can be controlled by the designer, the refractory period here depends on the thermal time constant $\tau_{th} = R_{th}C_{th}$ and the electrical time constant $\tau_{elec} = (R_w//LRS)C_0$.

7.4 Comparison between the Conventional CMOS IAF Neuron and the Proposed IMT Neuron

Table 7.1 summarizes the differences between the conventional CMOS IAF neuron and the proposed IMT neuron with respect to each neuron property. IMT neurons, as mentioned before, are more compact than their CMOS counterparts for they do not require dedicated CMOS circuitry for each phase of operation. This, however, comes at the expense of limited design options as the neuron properties are governed by not only the device parameters of

Table 7.1: CMOS-based neurons vs. IMT-based neurons

Property	CMOS Neuron	IMT Neuron
Accumulation	$R_m = \infty$, IAF $R_m = R_{mo}$, LIF	$R_m(elec) = R_{IMT}//R_w$ $R_m(th) = R_{th}$
Fire	$V_{th} = V_{tho}$	$V_{th} = \sqrt{\frac{HRS}{R_{th}}(T_c - T_a)}$
Refractory Period	$T_{ref} = T_{refo}$	$T_{ref} = f(\tau_{th}, \tau_{elec})$

the IMT, but also circuit variables such as synaptic weights. This dependency on synaptic weights results in the neuron properties, leakage in this case, changing for each set of programmed weights. These limitations should be carefully studied should the designer choose to employ and IMT neuron in a neuromorphic circuit design.

7.5 Conclusions

This chapter proposed a Leaky Integrate And Fire (LIF) spiking neuron based on IMT devices. The proposed neuron leverages the switching physics of the device which alleviates the need for complex CMOS circuitry. The neuron properties, however, are function of the device parameters. Design expressions were derived to help with the design space exploration of IMT based LIF neurons.

Chapter 8

Conclusions and Future Prospects

8.1 Conclusions

Neuromorphic computing has emerged as a promising alternative for conventional Von Neumann architectures. The inherent parallelism it provides due to the distributed nature of memory (synapses) and processors (neurons) alleviates the memory wall problem and promises significant advances in computing. Several neuromorphic hardware platforms have been proposed each with its merits and demerits. Memristive devices, in particular, have attracted significant interest owing to their electrical characteristics that make them suitable for neuromorphic hardware. To this end, this dissertation investigated the design of neuromorphic crossbar arrays using emerging memristive devices. Two memristive devices were studied, namely, Transition Metal Oxide (TMO) devices and Insulator Metal Transition (IMT) devices. This work adopted a bottom-up approach and focused on the device and circuit abstractions. Compact models for both devices were first developed and verified against experimental data followed by proposing circuit design techniques for integrating those devices in neuromorphic circuits. The contributions presented in this work can be summarized as follows:

- A compact model of TMO memristors was first developed and implemented in Verilog-A. The proposed model is compatible with SPICE simulators and based on measurable parameters which facilitates parameter extraction. The model was verified against

experimental data extracted from TMO devices available at UTK as well as data drawn from the literature.

- Transition Metal Oxide (TMO) devices require a one step electroforming process before the device can be used for regular operation. The forming voltage of TMO devices is typically high which hinders their integration with state-of-the-art CMOS technologies. A model for the forming voltage is presented in this work and validated against experimental data and Monte Carlo simulations. It is shown that decreasing/increasing the device thickness/area results in reducing the forming voltage. Local field enhancement may also help reducing the device's forming voltage.

- The integration of TMO devices as synapses in neuromorphic arrays is then investigated. It is shown that the line resistance and the leakage currents present in syanptic arrays hamper the voltage delivery to the selected cells which results in performance degradation. A multi-driver write scheme is proposed that enhances voltage delivery via reducing the effective line resistance and leakage current paths. This enhancement in voltage delivery ultimately improves the performance of the synaptic array while maintaining the same overall crossbar area.

- Insulator Metal transition (IMT) devices are also studied. A SPICE model is proposed that captures the electro-thermal coupling involved in the resistive transition of the device. The proposed model is validated against experimental data and electro-thermal device simulations.

- An IMT-based Integrate-And-Fire (IAF) neuron is then proposed which leverages the IMT switching dynamics to deliver the neuron's functionality. Impact of IMT device parameters on the proposed neuron is also studied and design guidelines are presented. It is shown that, unlike CMOS-based neurons, the IMT neuron parameters such as the firing threshold and the refractory period depend on the device parameters as well as circuit variables.

8.2 Future Prospects

The work presented in this dissertation addressed some of the existing challenges at both device and circuit levels that hinder the realization of robust and reliable memristive neuromorphic systems. Solutions at both levels of abstraction were provided but there is obviously still more room for improvement. The author would advise further exploration of the following:

- The high forming voltages of TMO devices hinders the seamless integration of such devices with state-of-the-art CMOS technology. Currently, techniques such as the ones presented in [66] may be required to form TMO devices in-field should the forming voltage of the TMO devices exceed that of the CMOS devices. Otherwise, old CMOS nodes may be deployed to solve this compatibility issue at the cost of increased area occupancy and potentially power consumption. The author would suggest two future directions to address this limitation: (I) work on reducing the forming voltage at the device level. This may include structural modification of the device or exploring new materials and (II) compare the performance of All-CMOS neuromorphic circuit implemented in a modern CMOS node to the performance of a hybrid CMOS/memristor neurmorphic circuit implemented in an older CMOS node. This study can help determine the necessity of integrating TMO devices with advanced CMOS nodes to outperform their All-CMOS counterparts or the lack thereof.

- The work presented in this dissertation on IMT devices is a first step towards the integration of IMT devices in neuromorphic circuit applications. More efforts should be undertaken to further explore the use of IMTs in that direction. This includes: (I) working on developing more accurate device and SPICE models that better reflect the switching physics of the device, (II) explore the use of IMT devices in synaptic networks to provide short term plasticity and (III) develop high level models of IMTs to explore their impact at the application level.

Bibliography

[1] M. E. Fouda, A. M. Eltawil, and F. Kurdahi, "Modeling and analysis of passive switching crossbar arrays," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 1, pp. 270–282, 2018. xii, 3

[2] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *nature*, vol. 453, no. 7191, p. 80, 2008, DOI:10.1038/nature06932. xii, 1, 10, 12, 14, 22, 23, 24, 25, 40, 59

[3] P.-Y. Chen and S. Yu, "Compact Modeling of RRAM Devices and its Applications in 1T1R and 1S1R Array Design," *IEEE Transactions on Electron Devices*, vol. 62, no. 12, pp. 4022–4028, 2015, DOI: 10.1109/TED.2015.2492421. xii, 12, 13, 15, 22, 25, 60

[4] S. Ambrogio, S. Balatti, A. Cubeta, A. Calderoni, N. Ramaswamy, and D. Ielmini, "Statistical fluctuations in hfo x resistive-switching memory: Part i-set/reset variability," *IEEE Transactions on Electron Devices*, vol. 61, no. 8, pp. 2912–2919, 2014. xii, 13, 15

[5] K. Beckmann, J. Holt, H. Manem, J. Van Nostrand, and N. C. Cady, "Nanoscale Hafnium Oxide RRAM Devices Exhibit Pulse Dependent Behavior and Multi-level Resistance Capability," *MRS Advances*, vol. 1, no. 49, pp. 3355–3360, May 2016, DOI:10.1557/adv.2016.377. xii, 28, 29

[6] I. Messaris, S. Alexander, S. Stathopoulos, A. Khiat, S. Nikolaidis, and T. Prodromakis, "A data-driven verilog-a reram model," *IEEE Transactions on Computer Aided Design*, vol. 37, no. 12, pp. 3151–3162, 2018. xii, 15, 23, 29

[7] D. Biolek, Z. Kolka, V. Biolkova, and Z. Biolek, "Memristor models for spice simulation of extremely large memristive networks," in *Circuits and Systems (ISCAS), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 389–392. xiii, 37, 38

[8] W. Kim, D. J. Wouters, S. Menzel, C. Rodenbücher, R. Waser, and V. Rana, "Lowering forming voltage and forming-free behavior of Ta_2O_5 ReRAM devices," in *IEEE European Solid-State Device Research Conference (ESSDERC)*, Sept 2016, pp. 164–167, DOI:10.1109/ESSDERC.2016.7 599 612. xiii, 16, 40, 43, 44

[9] S. Koveshnikov, K. Matthews, K. Min, D. Gilmer, M. Sung, S. Deora, H. Li, S. Gausepohl, P. Kirsch, and R. Jammy, "Real-Time Study of Switching Kinetics in Integrated 1T/HfO_x 1R RRAM: Intrinsic Tunability of Set/Reset Voltage and Trade-off with Switching Time," in *IEEE International Electron Devices Meeting (IEDM)*, Dec 2012, pp. 20.4.1–20.4.3, DOI:10.1109/IEDM.2012.6 479 080. xiii, 40, 43, 44

[10] L. Chua, "Memristor-the missing Circuit Element," *IEEE Transactions on circuit theory*, vol. 18, no. 5, pp. 507–519, 1971, DOI: 10.1109/TCT.1971.1083337. 1, 8, 22, 23, 59

[11] L. O. Chua and S. M. Kang, "Memristive devices and systems," *Proceedings of the IEEE*, vol. 64, no. 2, pp. 209–223, 1976, DOI: 10.1109/PROC.1976.10092. 1, 8, 11, 59

[12] C. Leon, "Everything you wish to know about memristors but are afraid to ask," *Radioengineering*, vol. 24, no. 2, p. 319, 2015. 1, 16, 59

[13] A. Ciprut and E. G. Friedman, "Energy-efficient write scheme for nonvolatile resistive crossbar arrays with selectors," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 4, pp. 711–719, 2018. 2, 51

[14] S. Kim, J. Zhou, and W. D. Lu, "Crossbar rram arrays: Selector device requirements during write operation," *IEEE Transactions on Electron Devices*, vol. 61, no. 8, pp. 2820–2826, 2014. 2

[15] A. Ciprut and E. G. Friedman, "Modeling size limitations of resistive crossbar array with cell selectors," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 25, no. 1, pp. 286–293, 2017. 2, 50

[16] J. Zhou, K.-H. Kim, and W. Lu, "Crossbar rram arrays: Selector device requirements during read operation," *IEEE Transactions on Electron Devices*, vol. 61, no. 5, pp. 1369–1376, 2014. 4

[17] S. Yu, "Neuro-inspired computing with emerging nonvolatile memorys," *Proceedings of the IEEE*, vol. 106, no. 2, pp. 260–285, 2018. 4, 5, 50

[18] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, B. L. Jackson, N. Imam, C. Guo, Y. Nakamura *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, 2014. 5

[19] Y.-H. Chen, T. Krishna, J. S. Emer, and V. Sze, "Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 1, pp. 127–138, 2016. 5

[20] B. Moons, R. Uytterhoeven, W. Dehaene, and M. Verhelst, "14.5 envision: A 0.26-to-10tops/w subword-parallel dynamic-voltage-accuracy-frequency-scalable convolutional neural network processor in 28nm fdsoi," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE, 2017, pp. 246–247. 5

[21] G. Desoli, N. Chawla, T. Boesch, S.-p. Singh, E. Guidetti, F. De Ambroggi, T. Majo, P. Zambotti, M. Ayodhyawasi, H. Singh *et al.*, "14.1 a 2.9 tops/w deep convolutional neural network soc in fd-soi 28nm for intelligent embedded systems," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE, 2017, pp. 238–239. 5

[22] J. Sim, J.-S. Park, M. Kim, D. Bae, Y. Choi, and L.-S. Kim, "14.6 a 1.42 tops/w deep convolutional neural network recognition processor for intelligent ioe systems," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE, 2016, pp. 264–265. 5

[23] P.-Y. Chen, J.-s. Seo, Y. Cao, and S. Yu, "Compact Oscillation Neuron Exploiting Metal-Insulator-Transition for Neuromorphic Computing," in *Proceedings of the ACM International Conference on Computer-Aided Design*, 2016, pp. 1–6, DOI: 10.1145/2 966 986.2 967 015. 5, 11, 16, 57, 60, 64, 69

[24] J. J. Yang, M. D. Pickett, X. Li, D. A. Ohlberg, D. R. Stewart, and R. S. Williams, "Memristive switching mechanism for metal/oxide/metal nanodevices," *Nature nanotechnology*, vol. 3, no. 7, p. 429, 2008. 10

[25] A. Pirovano, A. Lacaita, D. Merlani, A. Benvenuti, F. Pellizzer, and R. Bez, "Electronic switching effect in phase-change memory cells," in *Electron Devices Meeting, 2002. IEDM'02. International*. IEEE, 2002, pp. 923–926. 10

[26] A. G. Qoutb and E. G. Friedman, "Mtj magnetization switching mechanisms for iot applications," in *Proceedings of the 2018 on Great Lakes Symposium on VLSI*. ACM, 2018, pp. 347–352. 10

[27] H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. T. Chen, and M.-J. Tsai, "Metal–oxide rram," *Proceedings of the IEEE*, vol. 100, no. 6, pp. 1951–1970, 2012. 11, 16

[28] Y. Zhou and S. Ramanathan, "Mott Memory and Neuromorphic Devices," *Proceedings of the IEEE*, vol. 103, no. 8, pp. 1289–1310, 2015, DOI: 10.1109/JPROC.2015.2431914. 11, 16, 57

[29] A. Chen, "Memory selector devices and crossbar array design: a modeling-based assessment," *Journal of Computational Electronics*, vol. 16, no. 4, pp. 1186–1200, 2017. 11

[30] J. S. Najem, G. J. Taylor, N. Armendarez, R. J. Weiss, M. S. Hasan, G. S. Rose, C. D. Schuman, A. Belianinov, S. A. Sarles, and C. P. Collier, "Assembly and characterization of biomolecular memristors consisting of ion channel-doped lipid membranes," *JoVE (Journal of Visualized Experiments)*, no. 145, p. e58998, 2019. 11

[31] J. S. Najem, G. J. Taylor, R. J. Weiss, M. S. Hasan, G. Rose, C. D. Schuman, A. Belianinov, C. P. Collier, and S. A. Sarles, "Memristive ion channel-doped biomembranes as synaptic mimics," *ACS nano*, vol. 12, no. 5, pp. 4702–4711, 2018. 11

[32] J. Lin, S. Sonde, C. Chen, L. Stan, K. Achari, S. Ramanathan, S. Guha *et al.*, "Low-Voltage Artificial Neuron using Feedback Engineered Insulator-to-Metal-Transition Devices," in *IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 1–4, DOI: 10.1109/IEDM.2016.7 838 541. 11, 16, 57, 58, 64, 68, 69

[33] J. H. Yoon, Z. Wang, K. M. Kim, H. Wu, V. Ravichandran, Q. Xia, C. S. Hwang, and J. J. Yang, "An artificial nociceptor based on a diffusive memristor," *Nature communications*, vol. 9, no. 1, p. 417, 2018. 11

[34] Z. Wang, S. Joshi, S. E. Savelev, H. Jiang, R. Midya, P. Lin, M. Hu, N. Ge, J. P. Strachan, Z. Li *et al.*, "Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing," *Nature materials*, vol. 16, no. 1, p. 101, 2017. 11

[35] S. Amer, M. S. Hasan, and G. S. Rose, "Analysis and modeling of electroforming in transition metal oxide-based memristors and its impact on crossbar array density," *IEEE Electron Device Letters*, vol. 39, no. 1, pp. 19–22, 2017. 11

[36] D. Biolek, V. Biolkova, and Z. Biolek, "Spice model of memristor with nonlinear dopant drift," *Radioengineering*, 2009. 12, 22, 25

[37] T. Prodromakis, B. P. Peh, C. Papavassiliou, and C. Toumazou, "A versatile memristor model with nonlinear dopant kinetics," *IEEE transactions on electron devices*, vol. 58, no. 9, pp. 3099–3105, 2011. 12, 25

[38] Y. N. Joglekar and S. J. Wolf, "The elusive memristor: properties of basic electrical circuits," *European Journal of Physics*, vol. 30, no. 4, p. 661, 2009. 12, 25

[39] D. B. Strukov and R. S. Williams, "Exponential ionic drift: fast switching and low volatility of thin-film memristors," *Applied Physics A: Materials Science & Processing*, vol. 94, no. 3, pp. 515–519, 2009. 12, 22, 24, 25

[40] M. D. Pickett, D. B. Strukov, J. L. Borghetti, J. J. Yang, G. S. Snider, D. R. Stewart, and R. S. Williams, "Switching dynamics in titanium dioxide memristive devices," *Journal of Applied Physics*, vol. 106, no. 7, p. 074508, 2009. 12, 22, 24, 25

[41] S. Kvatinsky, M. Ramadan, E. G. Friedman, and A. Kolodny, "Vteam: A general model for voltage-controlled memristors," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, no. 8, pp. 786–790, 2015. 12, 25

[42] S. Kvatinsky, E. G. Friedman, A. Kolodny, and U. C. Weiser, "Team: Threshold adaptive memristor model," *IEEE transactions on circuits and systems I: regular papers*, vol. 60, no. 1, pp. 211–221, 2012. 12, 25

[43] A. M. Hassan, H. A. Fahmy, and N. H. Rafat, "Enhanced model of conductive filament-based memristor via including trapezoidal electron tunneling barrier effect," *IEEE Transactions on Nanotechnology*, vol. 15, no. 3, pp. 484–491, 2016. 12

[44] J. G. Simmons, "Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film," *Journal of applied physics*, vol. 34, no. 6, pp. 1793–1803, 1963. 12, 25

[45] C. Yakopcic, T. M. Taha, G. Subramanyam, and R. E. Pino, "Generalized memristive device spice model and its application in circuit design," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, no. 8, pp. 1201–1214, 2013. 12

[46] C. Yakopcic, T. M. Taha, G. Subramanyam, R. E. Pino, and S. Rogers, "A memristor device model," *IEEE electron device letters*, vol. 32, no. 10, pp. 1436–1438, 2011. 12

[47] X. Guan, S. Yu, and H.-S. P. Wong, "A SPICE Compact Model of Metal Oxide Resistive Switching Memory with Variations," *IEEE Electron Device Letters*, vol. 33, no. 10, pp. 1405–1407, 2012, DOI: 10.1109/LED.2012.2210856. 12, 25

[48] S. Yu, Y. Yin Chen, X. Guan, H.-S. Philip Wong, and J. A. Kittl, "A Monte Carlo study of the low resistance state retention of HfOx based resistive switching memory," *Applied Physics Letters*, vol. 100, no. 4, p. 043507, 2012, DOI:10.1063/1.3679610. 13, 42, 45

[49] S. Yu, X. Guan, and H.-S. P. Wong, "Understanding metal oxide rram current overshoot and reliability using kinetic monte carlo simulation," in *2012 International Electron Devices Meeting*. IEEE, 2012, pp. 26–1. 13

[50] F. M. Bayat, B. Hoskins, and D. B. Strukov, "Phenomenological modeling of memristive devices," *Applied Physics A*, vol. 118, no. 3, pp. 779–786, 2015. 15, 23, 36

[51] S. Amer, S. Sayyaparaju, G. S. Rose, K. Beckmann, and N. C. Cady, "A practical hafnium-oxide memristor model suitable for circuit design and simulation," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2017, pp. 1–4. 15

[52] R. E. Pino, J. W. Bohl, N. McDonald, B. Wysocki, P. Rozwood, K. A. Campbell, A. Oblea, and A. Timilsina, "Compact method for modeling and simulation of memristor devices: Ion conductor chalcogenide-based memristor devices," in *Proceedings of the 2010 IEEE/ACM International Symposium on Nanoscale Architectures*. IEEE Press, 2010, pp. 1–4. 15, 23, 36

[53] S. Jung, J. Kong, T.-W. Kim, S. Song, K. Lee, T. Lee, H. Hwang, and S. Jeon, "In-depth study on the effect of active-area scale-down of solution-processed tiox," *IEEE Electron Device Letters*, vol. 33, no. 6, pp. 869–871, 2012. 16

[54] F. Pan, S. Yin, and V. Subramanian, "A detailed study of the forming stage of an electrochemical resistive switching memory by kmc simulation," *IEEE Electron Device Letters*, vol. 32, no. 7, pp. 949–951, 2011. 16

[55] J. Guy, G. Molas, P. Blaise, C. Carabasse, M. Bernard, A. Roule, G. Le Carval, V. Sousa, H. Grampeix, V. Delaye *et al.*, "Experimental and theoretical understanding of forming, set and reset operations in conductive bridge ram (cbram) for memory stack optimization," in *2014 IEEE International Electron Devices Meeting*. IEEE, 2014, pp. 6–5. 16

[56] E. Abbaspour, S. Menzel, A. Hardtdegen, S. Hoffmann-Eifert, and C. Jungemann, "Kmc simulation of the electroforming, set and reset processes in redox-based resistive switching devices," *IEEE Transactions on Nanotechnology*, vol. 17, no. 6, pp. 1181–1188, 2018. 16

[57] D. Ruzmetov, S. D. Senanayake, V. Narayanamurti, and S. Ramanathan, "Correlation between metal-insulator transition characteristics and electronic structure changes in vanadium oxide thin films," *Physical Review B*, vol. 77, no. 19, p. 195442, 2008, DOI: 10.1103/PhysRevB.77.195442. 16, 57

[58] E. Freeman, A. Kar, N. Shukla, R. Misra, R. Engel-Herbert, D. Schlom, V. Gopalan, K. Rabe, and S. Datta, "Characterization and modeling of metal-insulator transition (mit) based tunnel junctions," in *Device Research Conference (DRC), 2012 70th Annual*. IEEE, 2012, pp. 243–244. 16, 57

[59] D. Ruzmetov, G. Gopalakrishnan, J. Deng, V. Narayanamurti, and S. Ramanathan, "Electrical triggering of metal-insulator transition in nanoscale vanadium oxide junctions," *Journal of Applied Physics*, vol. 106, no. 8, p. 083702, 2009, DOI: 10.1063/1.3245338. 16, 57

[60] Z. Yang, S. Hart, C. Ko, A. Yacoby, and S. Ramanathan, "Studies on electric triggering of the metal-insulator transition in vo2 thin films between 77 k and 300 k," *Journal of Applied Physics*, vol. 110, no. 3, p. 033725, 2011, DOI: 10.1063/1.3619806. 16, 57

[61] G. Gopalakrishnan, D. Ruzmetov, and S. Ramanathan, "On the triggering mechanism for the metal–insulator transition in thin film vo2 devices: electric field versus thermal effects," *Journal of Materials Science*, vol. 44, no. 19, pp. 5345–5353, 2009, DOI: 10.1007/s10853-009-3442-7. 16, 57

[62] K. Karda, C. Mouli, S. Ramanathan, and M. A. Alam, "A Self-Consistent, Semiclassical Electrothermal Modeling Framework for Mott Devices," *IEEE Transactions on Electron Devices*, vol. 65, no. 5, pp. 1672–1678, 2018, DOI: 10.1109/TED.2018.2817604. 16, 58

[63] M. Mao, S. Yu, and C. Chakrabarti, "Design and analysis of energy-efficient and reliable 3-d reram cross-point array system," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2018. 19

[64] M. Uddin, M. B. Majumder, G. S. Rose, K. Beckmann, H. Manem, Z. Alamgir, and N. C. Cady, "Techniques for improved reliability in memristive crossbar puf circuits," in *VLSI (ISVLSI), 2016 IEEE Computer Society Annual Symposium on*. IEEE, 2016, pp. 212–217. 23

[65] N. R. McDonald, R. E. Pino, P. J. Rozwood, and B. T. Wysocki, "Analysis of dynamic linear and non-linear memristor device models for emerging neuromorphic

computing hardware design," in *Neural Networks (IJCNN), The 2010 International Joint Conference on*. IEEE, 2010, pp. 1–5. 26

[66] S. Amer, K. Beckmann, N. C. Cady, and G. S. Rose, "Design Techniques for In-Field Memristor Forming Circuits," in *IEEE Midwest Symposium on Circuits and Systems (MWSCAS)*, Aug 2017, pp. 1224–1227, DOI:10.1109/MWSCAS.2017.8 053 150. 40, 81, 108

[67] B. Govoreanu, G. Kar, Y. Chen, V. Paraschiv, S. Kubicek, A. Fantini, I. Radu, L. Goux, S. Clima, R. Degraeve, N. Jossart, O. Richard, T. Vandeweyer, K. Seo, P. Hendrickx, G. Pourtois, H. Bender, L. Altimime, D. J. Wouters, J. A. Kittl, and M. Jurczak, "$10\times10\text{nm}^2$ Hf/HfO_x Crossbar Resistive RAM with Excellent Performance, Reliability and low-Energy Operation," in *IEEE International Electron Devices Meeting (IEDM)*, Dec 2011, pp. 31.6.1–31.6.4, DOI:10.1109/IEDM.2011.6 131 652. 40

[68] E. Y. Wu and R.-P. Vollertsen, "On the Weibull Shape Factor of Intrinsic Breakdown of Dielectric Films and its Accurate Experimental Determination. Part I: Theory, Methodology, Experimental Techniques," *IEEE Transactions on Electron Devices*, vol. 49, no. 12, pp. 2131–2140, Dec 2002, DOI:10.1109/TED.2002.805612. 41

[69] S. Hofstein and F. Heiman, "The silicon insulated-gate field-effect transistor," *Proceedings of the IEEE*, vol. 51, no. 9, pp. 1190–1202, Sept 1963, DOI:10.1109/PROC.1963.2488. 41

[70] A. Chen, "Area and Thickness Scaling of Forming Voltage of Resistive Switching Memories," *IEEE Electron Device Letters*, vol. 35, no. 1, pp. 57–59, Jan 2014, DOI: 10.1109/LED.2013.2288262. 41

[71] J. S. Lee, S. Lee, and T. W. Noh, "Resistive switching phenomena: A review of statistical physics approaches," *Applied Physics Reviews*, vol. 2, no. 3, p. 031303, 2015, DOI:10.1063/1.4929512. 42

[72] D. Ielmini, F. Nardi, and C. Cagli, "Universal Reset Characteristics of Unipolar and Bipolar Metal-Oxide RRAM," *IEEE Transactions on Electron Devices*, vol. 58, no. 10, pp. 3246–3253, July 2011, DOI:10.1109/TED.2011.2161088. 42

[73] 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)*. IEEE, 2018, pp. 1–4. 50

[74] B. Yan, A. M. Mahmoud, J. J. Yang, Q. Wu, Y. Chen, and H. H. Li, "A neuromorphic asic design using one-selector-one-memristor crossbar," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2016, pp. 1390–1393. 50

[75] B. Liu, H. Li, Y. Chen, X. Li, T. Huang, Q. Wu, and M. Barnell, "Reduction and ir-drop compensations techniques for reliable neuromorphic computing systems," in *Proceedings of the 2014 IEEE/ACM International Conference on Computer-Aided Design*. IEEE Press, 2014, pp. 63–70. 50

[76] S. B. Eryilmaz, D. Kuzum, S. Yu, and H.-S. P. Wong, "Device and system level design considerations for analog-non-volatile-memory based neuromorphic architectures," in *2015 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 2015, pp. 4–1. 50

[77] S. B. Eryilmaz, S. Joshi, E. Neftci, W. Wan, G. Cauwenberghs, and H.-S. P. Wong, "Neuromorphic architectures with electronic synapses," in *2016 17th International Symposium on Quality Electronic Design (ISQED)*. IEEE, 2016, pp. 118–123. 50

[78] W. Wen, C.-R. Wu, X. Hu, B. Liu, T.-Y. Ho, X. Li, and Y. Chen, "An eda framework for large scale hybrid neuromorphic computing systems," in *2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*. IEEE, 2015, pp. 1–6. 50

[79] A. Ciprut and E. G. Friedman, "Design models of resistive crossbar arrays with selector devices," in *Circuits and Systems (ISCAS), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 1250–1253. 50

[80] S. Amer and G. S. Rose, "A multi-driver write scheme for reliable and energy efficient 1s1r reram crossbar arrays," in *20th International Symposium on Quality Electronic Design (ISQED)*. IEEE, 2019, pp. 64–69. 51

[81] A. Pergament, G. Stefanovich, and A. Velichko, "Oxide Electronics and Vanadium Dioxide Perspective: A Review," *Journal on Selected Topics in Nano Electronics and Computing*, vol. 1, no. 1, pp. 24–43, 2013, DOI: 10.15393/j8.art.2013.3002. 57

[82] M. Son, J. Lee, J. Park, J. Shin, G. Choi, S. Jung, W. Lee, S. Kim, S. Park, and H. Hwang, "Excellent Selector Characteristics of Nanoscale VO2 for High-Density Bipolar ReRAM Applications," *IEEE Electron Device Letters*, vol. 32, no. 11, pp. 1579–1581, 2011, DOI: 10.1109/LED.2011.2163697. 57

[83] L. Gao, K. E. Holbert, and S. Yu, "Total Ionizing Dose Effects of Gamma-Ray Radiation on NbO x-Based Selector Devices for Crossbar Array Memory," *IEEE Transactions on Nuclear Science*, vol. 64, no. 6, pp. 1535–1539, 2017, DOI: 10.1109/TNS.2017.2700434. 57

[84] J. Lin, K. Alam, L. Ocola, Z. Zhang, S. Datta, S. Ramanathan, and S. Guha, "Physics and Technology of Electronic Insulator-to-Metal Transition (E-IMT) for Record High On/Off ratio and Low Voltage in Device Applications," in *IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 1–4, DOI: 10.1109/IEDM.2017.8 268 446. 57, 58

[85] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*. Oxford Univ. Press, 2011. 60, 61

[86] C. C. McAndrew, G. J. Coram, K. K. Gullapalli, J. R. Jones, L. W. Nagel, A. S. Roy, J. Roychowdhury, A. J. Scholten, G. D. Smit, X. Wang *et al.*, "Best practices for compact modeling in verilog-a," *IEEE Journal of the Electron Devices Society*, vol. 3, no. 5, pp. 383–396, 2015. 60

[87] T. Wang and J. Roychowdhury, "Well-posed models of memristive devices," *arXiv:1605.04897v1 [cs.ET]*. 61

[88] T. Wang, "Modelling multistability and hysteresis in esd clamps, memristors and other devices," in *IEEE Custom Integrated Circuits Conference (CICC)*, 2017, pp. 1–10, DOI: 10.1109/CICC.2017.7 993 681. 61

[89] S. Amer, M. S. Hasan, M. M. Adnan, and G. S. Rose, "Design considerations for insulator metal transition based artificial neurons," in *2019 IEEE Midwest Symposium on Circuits and Systems (MWSCAS)*. IEEE, 2019, pp. 1–4. 69

[90] G. Chakma, M. M. Adnan, A. R. Wyer, R. Weiss, C. D. Schuman, and G. S. Rose, "Memristive mixed-signal neuromorphic systems: Energy-efficient learning at the circuit-level," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 8, no. 1, pp. 125–136, 2018. 72

[91] G. Chakma, S. Sayyaparaju, R. Weiss, and G. S. Rose, "A mixed-signal approach to memristive neuromorphic system design," in *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*. IEEE, 2017, pp. 547–550. 72

Appendices

Appendix A

RRAM Model Verilog-A code

Listing A.1: RRAM Verilog-A code

```
// VerilogA for demo, memr_TMO_switching, veriloga
//Model Developed by: Sherif Amer
//This model only includes memrsitor switching equations. It does not include
    any secondary effects such as variations, temperature dependence or aging.
//Copyright of the model is maintained by the developers. This model is
    distributed under the terms of the Creative Commons Attribution-NonCommercial
    4.0
//International Public License
    https://creativecommons.org/licenses/by-nc/4.0/legalcode. If you choose to
    use this model, you are kindly requested to cite.
//the following paper: Sherif Amer, Sagarvarma Sayyaparaju, Karsten Beckmann,
    Nathaniel C. Cady and Garrett S. Rose, "A Practical Hafnium Oxide Memristor
//Model suitable for cirucit design and simulation", in International Symposium
    in Circuits and Systems (ISCAS), May 2017, pp. 1-4, DOI:10.1109/ISCAS.2017.8
    050 790.
`include "constants.vams"
`include "disciplines.vams"

module memr_TMO_switching(p,n,r);
```

```
 inout         p;         //positive pin
 inout         n;         //negative pin
 inout         r;         //Resistance terminal. This is not a physical terminal
    of the device. It is just there to facilitate the computaion of resistance.
  electrical    p, n, r;

// global parameters

  parameter real window    =  0;

// model parameters
  parameter real HRS = 1.5e5;  // high resistance state
  parameter real LRS = 1e4;    // low resistance state
  parameter real Vtp = 0.75;   // positive Voltage below which the change in
     resistance is zero
  parameter real Vtn = -1.0;    // negative Voltage below which the change in
     resistance is zero. Must be a negative value.
  parameter real tsw_p = 1e-8;    // time to switch under +V bias
  parameter real tsw_n = 1e-6;    // time to switch under -V bias

//window parameters

  parameter real theta_HRS = 0.85; // transition boundary at HRS
  parameter real beta_HRS = 0.2; // transition sharpness at HRS
  parameter real theta_LRS = 2.1; // transition boundary at LRS
  parameter real beta_LRS = 0.05; // transition sharpness at LRS

//Fitting parameters

   parameter real CLRS = 1; // speed parameter while transitioning to LRS
   parameter real CHRS = 1; // speed parameter while transitioning to HRS
```

```
   parameter real P_LRS = 3; // non-linearity parameter while transitioning to LRS
   parameter real P_HRS = 3; // non-linearity parameter while transitioning to HRS
   parameter real Rinit = 1e4; //initial resistance

real delR;
real time_last;
real Vwr;
real delt;
real Rm;
real Rm_tmp;


analog begin
  @ ( initial_step or initial_step("dc") ) begin

    delt = 0;
    time_last = 0;
    Rm  = Rinit;
    delR = HRS - LRS;

   end

  delt      = $abstime - time_last;
  time_last = $abstime;
  Vwr       = V(p,n);
//////////////////////////// Model equations//////////////////////////////////
// window == 0, window function is deactivated

 if(window == 0)begin
   if (Vwr >= Vtp && Rm != LRS) begin
      Rm_tmp = Rm - delt* CLRS* (delR/tsw_p)*( pow(((Vwr-Vtp)/Vtp), P_LRS));
         //switching equation while transitioning from HRS to LRS
```

```
      if (Rm_tmp <= LRS) begin
         Rm_tmp = LRS; //clipping the resistance at LRS
      end
   end
   else if (Vwr < Vtn && Rm != HRS) begin
      Rm_tmp = Rm + delt* CHRS * (delR/tsw_n) *(pow(((Vwr-Vtn)/Vtn), P_HRS)) ;
          //switching equation while transitioning from LRS to HRS
      if (Rm_tmp >= HRS) begin
         Rm_tmp = HRS; //clipping the resistance at HRS
      end
   end
   else begin
      Rm_tmp = Rm;
   end
 end
// window == 1, window function is activated
  if(window == 1)begin
   if (Vwr >= Vtp && Rm != LRS) begin
      Rm_tmp = Rm - delt* CLRS* (delR/tsw_p)* pow(((Vwr-Vtp)/Vtp), P_LRS)
          /(1+exp((theta_LRS*LRS-Rm)/(delR)/beta_LRS));
      if (Rm_tmp <= LRS) begin
         Rm_tmp = LRS;
      end
   end
   else if (Vwr <= Vtn && Rm != HRS) begin
      Rm_tmp = Rm + delt* CHRS* (delR/tsw_n)* pow(((Vwr-Vtn)/Vtn),
          P_HRS)/(1+exp((Rm-theta_HRS*HRS)/(delR)/beta_HRS));
      if (Rm_tmp >= HRS) begin
         Rm_tmp = HRS;
      end
   end
   else begin
```

```
      Rm_tmp = Rm;
   end
 end
//////////////////////End Model Equations//////////////////////////
Rm = Rm_tmp;
I(p,n) <+ Vwr / Rm;
V(r) <+ Rm; //This is not a physical terminal
 end        // end analog
endmodule
```

Appendix B

IMT Model Verilog-A code

Listing B.1: IMT Verilog-A code

```
// VerilogA for demo, IMT_newmodel_simple, veriloga

`include "constants.vams"
`include "disciplines.vams"

module IMT_newmodel_simple (p,n,tempV,r,temp);

  inout         p;        //positive pin
  inout         n;        //negative pin
  inout         tempV;        //negative pin
  inout         r;        //negative pin
  electrical    p, n,r,tempV;
thermal temp;

//parameters

      parameter real HRS0        = 4e3;
      parameter real LRSF       = 4e1;
```

```
      parameter real B_HRS=0.0035;
      parameter real B_LRS=0.0025;
      parameter real A=1e2;

      parameter real Cth        = 3.174e-12;
      //parameter real tau_th     = 2.3e-7*0.53;
      parameter real Rth        = 4.1667e4;
      parameter real T0        = 300;
      parameter real Rinit      =  4e3;

       parameter real Tinit=300;
      parameter real T_0=300;
      parameter real T_F=400;

      parameter real Tc         = 330;
      parameter real Tx         = 2;

//variables

real Vwr;
real Iwr;
real Rm;
real K_HRS,K_LRS,HRS,LRS;
real tem;

analog begin

tem=Temp(temp);
Rm=V(r);
Iwr=I(p,n);
V(p,n) <+ Iwr*Rm;
```

```
//Resistance evolution
K_HRS=exp(-B_HRS*(tem-T_0));
K_LRS=exp(-B_LRS*(tem-T_F));

if (tem>T_F) begin
   LRS=LRSF*pow((1+pow(K_LRS,A)),1/A);
end
else begin
   LRS=LRSF*K_LRS*pow((1+pow(K_LRS,-A)),1/A);
end

if (tem>T_0) begin
   HRS=HRS0*K_HRS/(pow((1+pow(K_HRS,A)),1/A));
end
else begin
   HRS=HRS0/(pow((1+pow(K_HRS,-A)),1/A));
end
// Thermal Model
Rm= LRS + (HRS-LRS)/(1+exp((tem-Tc)/Tx));
Pwr(temp)<+ ddt(Temp(temp));
Pwr(temp)<+ -pow(Iwr,2)*Rm/Cth;
Pwr(temp)<+(Temp(temp)-T0)/(Rth*Cth);

V(r) <+ Rm;
V(tempV)<+tem;

end
endmodule
```

Appendix C

Crossbar Reduction Algorithm Using Delta-to-Wye Conversion

To enable better understanding and faster analysis of the multi-driver write scheme, it is important to develop a technique to reduce the crossbar array into a tractable circuit for which circuit analysis concepts can be applied. The algorithm developed herein is based on Delta-to-Wye conversion. The algorithm executes iterative Delta-to-Wye conversions on the target circuit to help arriving at a tractable circuit from which one can gain insight about the the circuit under consideration. It is important, however, to note that the algorithm derived in this work is exclusively applicable to the work discussed in this dissertation and may not be generalized to other problems involving crossbar arrays.

The equivalent circuit of the crossbar array under the $V/2$ bias scheme during the write operation is depicted in Fig C.1. This analysis considers the worst case cell defined as the cell farthest from the row and column drivers. Driving this cell results in the highest leakage current through the half selected cells and the highest drop across the line resistance and, hence, exhibits the worst voltage delivery.

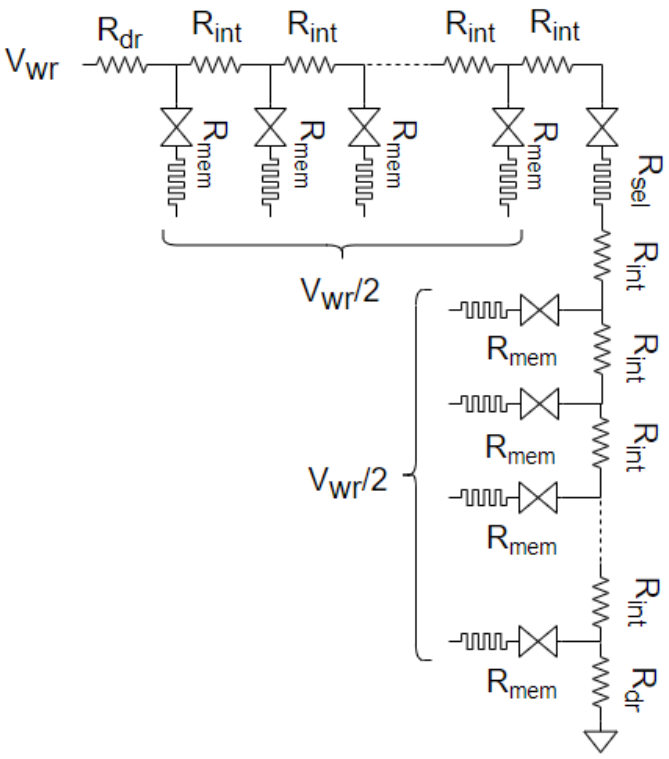

Figure C.1: Equivalent circuit of the crossbar array under worst case write operation.

The first step in the algorithm is to reduce the equivalent circuit in Fig C.1 to a tractable circuit for which closed form expressions for the voltage and currents can be derived. This reduction is executed via an iterative Delta-to-Wye conversion as shown in fig C.2. For N branches, $N-1$ half selected branches are reduced in $(N-1)-1$ steps. The reduction algorithm is applied for both rows and columns.

Fig C.3 depicts the reduced circuit of the crossbar array where R_H and R_V represent the line resistance and the leakage path resistance (resistance of the half selected cells). Fig C.4 depicts R_H and R_V versus the arrays size.

Kirchhoff Voltage Law (KVL) is then applied to the three loops depicted in Fig C.3 to find the current flowing in the circuit. The voltage across the selected cell V_{cell} is then calculated.

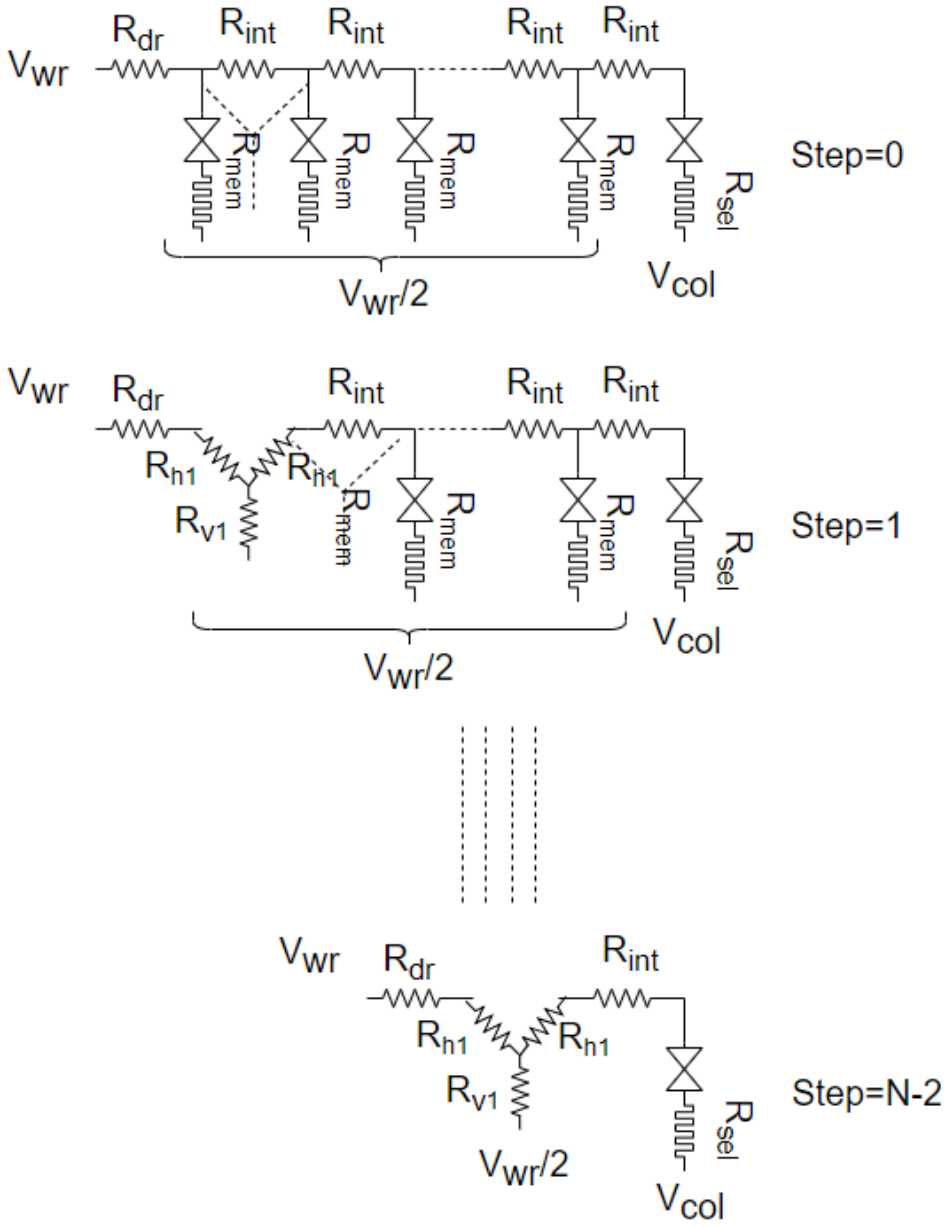

Figure C.2: Visual representation of circuit reduction using Delta-to-Wye conversion.

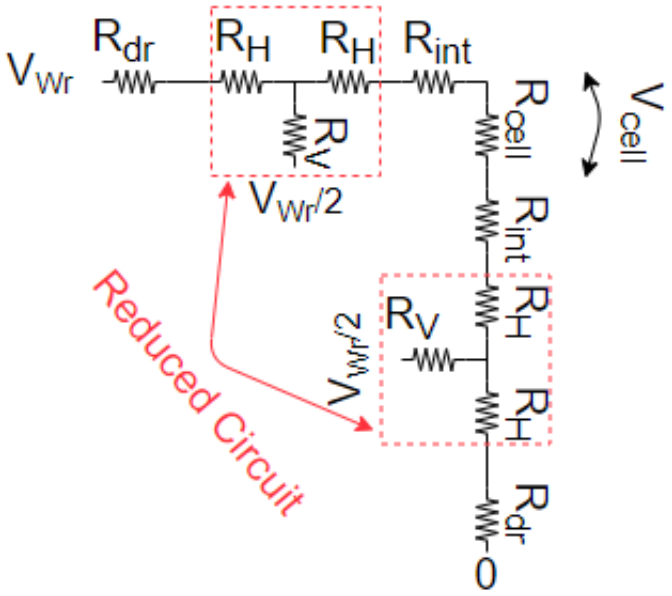

Figure C.3: Reduced equivalent circuit of the crossbar array.

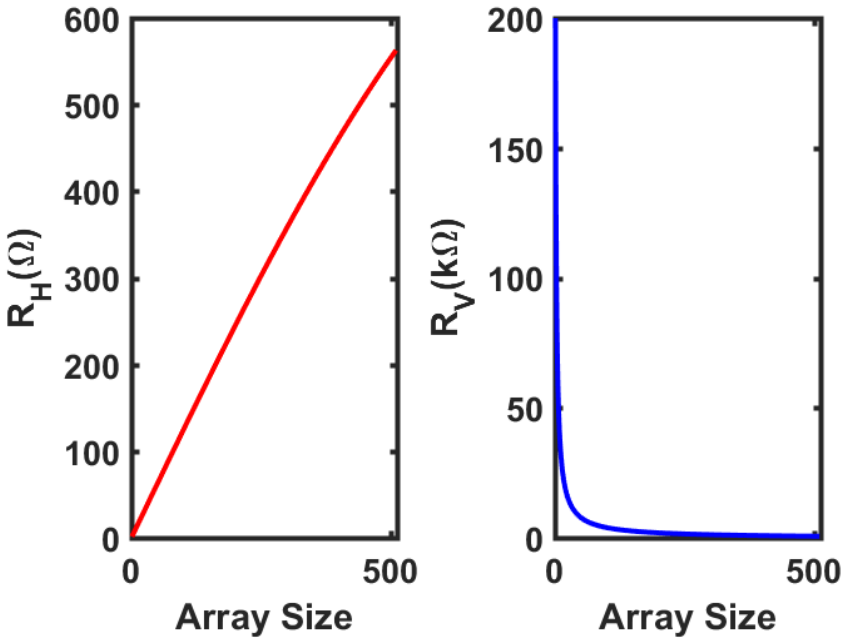

Figure C.4: R_H and R_V vs. array size.

Appendix D

Forming Circuit

Forming TMO devices in-field is necessary when memristors are integrated in the CMOS process flow. The process considered in this work is a hybrid CMOS/memristor flow developed by SUNY where Hf0x devices are integrated in the 65nm CMOS process available at SUNY. The forming voltage, however, of the Hf0x devices is higher than the nominal voltage of the the 65nm node available at SUNY which requries dedicated circuitry to execute in-field forming of the TMO devices.

Two forming circuits were designed in this work. The first forming circuit is presented in A and only considers forming the devices in-field. The second is presented in B and considers in-field forming as well as programming. The forming circuit in B was fabricated by SUNY Polytechnic institute as a part of a collaborative project.

A Forming Circuit I

The circuit presented in this section is based on the work in [66] and is shown in Fig D.1. The circuit operates in two phases: ($\phi 1$) operation phase and ($\phi 2$) forming phase. In $\phi 1$, the forming circuit is isolated and the device is connected to the pre and post neuron circuits (could be any other circuit depending on the application. In this work, we consider neuromorphic circuits). In In $\phi 2$, the forming circuit is activated and the pre and post neuron circuits are isolated. The voltages used for forming are typically higher than the nominal operating voltages and, therefore, the devices used for forming are DGXFET available in

the Process Design Kit (PDK) used at SUNY which can take up to 3.3V. The nominal voltages of the devices used in the pre and post neuron circuits is 1.2V. During forming, two pull down FET devices are activated to protect the pre and post neuron circuits. Also, a current compliance mechanism is activated to limit the current through the TMO device. $\phi 1$ and $\phi 2$ are controlled by a non-overlapping clock circuit to ensure that both paths are not simultaneously activated. The non-overlapping clock circuit is shown in Fig D.2.

Figs D.3, D.4 and D.5 depict the operation of the proposed circuit. In Fig D.3, during the forming phase, the pre and post neuron nodes are pulled down to $0V$ and the voltage across the TMO device exceeds $2.1V$ - the forming voltages of the devices used in this work. Fig D.4 shows that the current does no exceed $72\mu A$ which is the compliance current used for this design. Fig. D.5 shows successful forming for a wide range of pre-forming resistances.

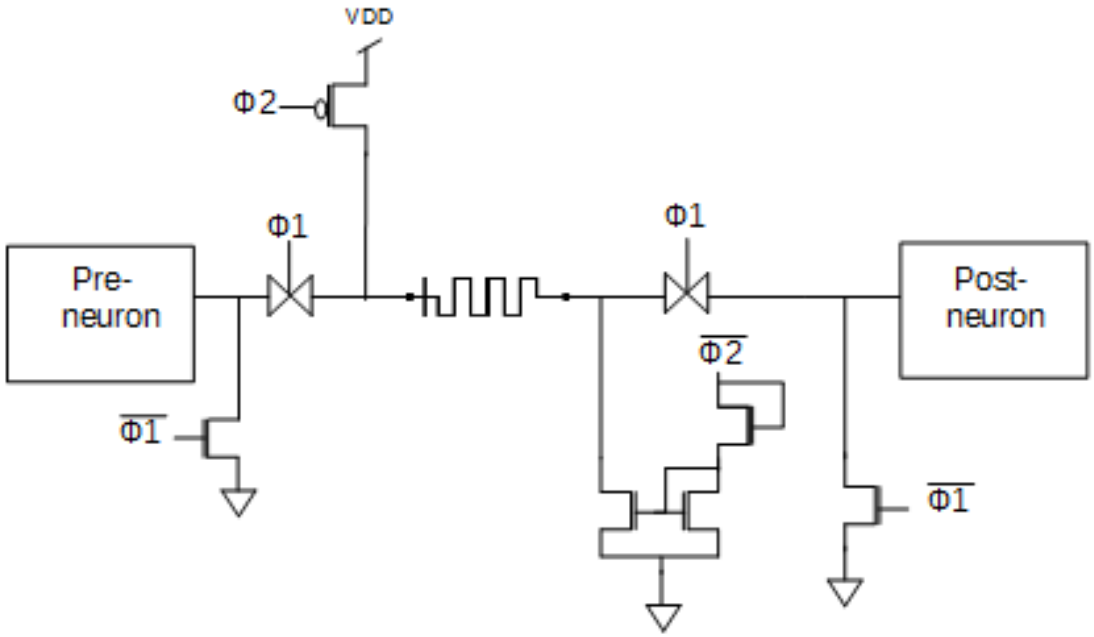

Figure D.1: Proposed in-field forming circuit

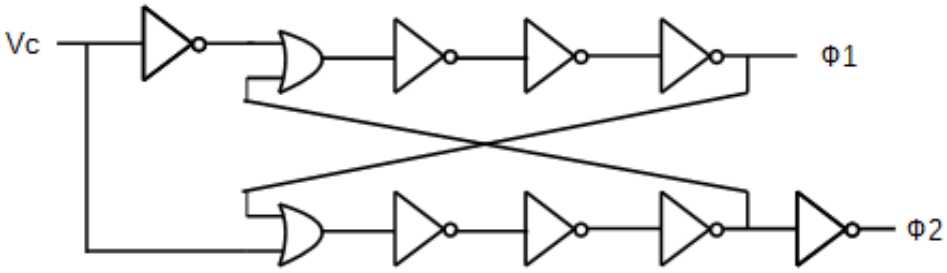

Figure D.2: Non-overlapping clock generator

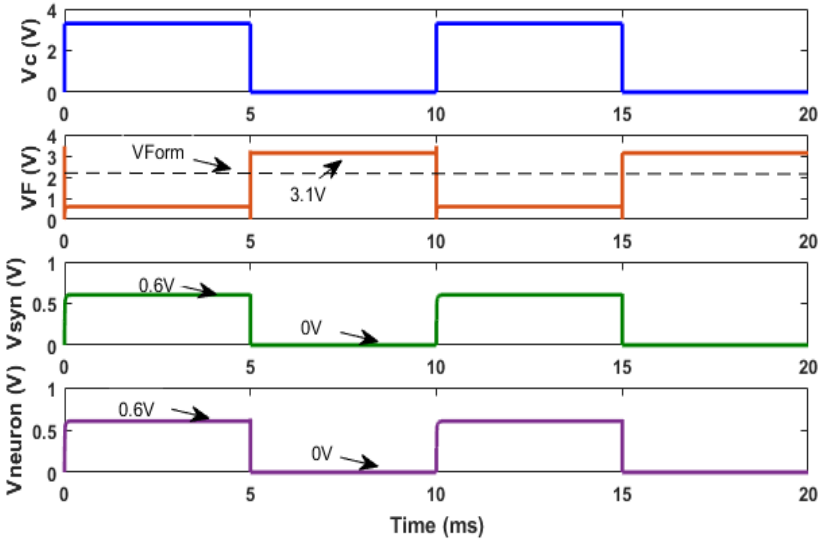

Figure D.3: Critical Voltage nodes during forming

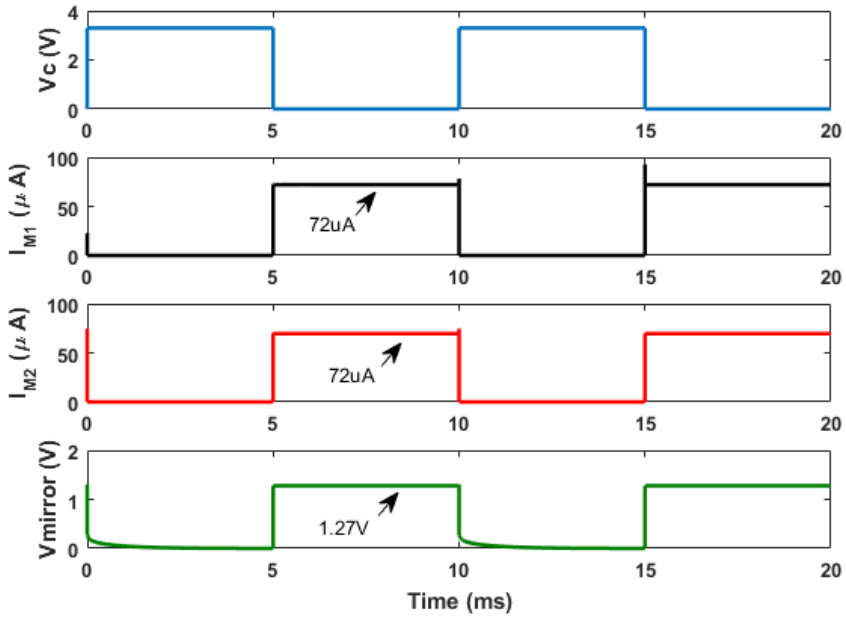

Figure D.4: Current Compliance

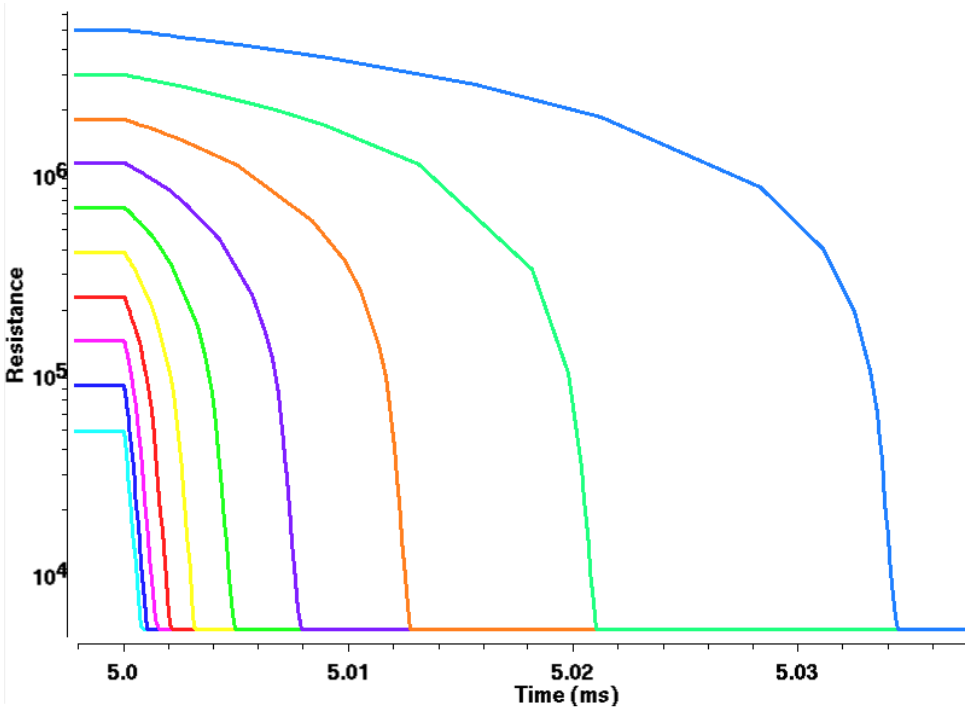

Figure D.5: Parametric analysis run over a wide range of pre-forming resistance values

B Forming Circuit II

Two versions of the forming/programming circuit were fabricated. Fig D.6 depicts the pin-out of the 12X2 probe pads. Table D.1 describes the various signals applied to the test structures. Table D.2 depicts the pin assignment. Note that duplicates of each version are present on the 12X2 probe pads. Hence, signal X(2) refers to signal X , replica (2). Versions I and II of the forming/programming circuits are discussed in subsections B.1 and B.2, respectively.

B.1 Version I

This forming/programming circuit executes both forming and programming of the memristor device. Table D.3 depicts the truth table for the proposed forming/programming circuit. Fig D.7 and Fig D.8 depict the schematic and simulation of the forming/programming circuit, respectively.

B.2 Version II

This circuit is another variation of the forming and programming circuit. Fig D.9 and Fig D.10 depict the circuit schematic and simulation, respectively. $S = 1$ corresponds to the Forming/SET path while $S = 0$ corresponds to the RESET path. $Vbias$ can be set to control the current through the device. $Vpad$ and $Vpad2$ are used to SET and RESET the device, respectively.

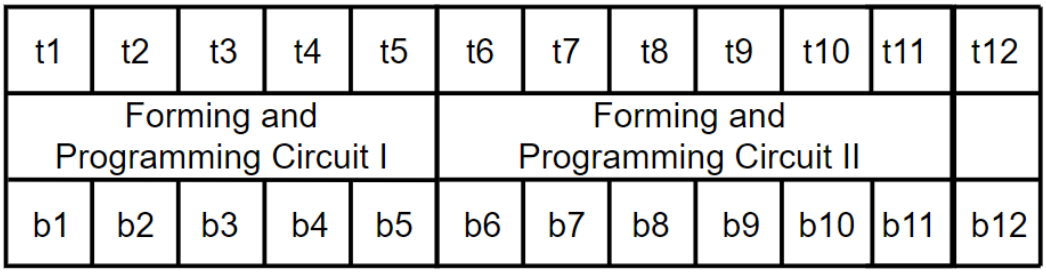

Figure D.6: 12X2 probe pad structure

Table D.1: Signal description for the test structures

Pin Attribute	Pin Type	Signal Description
In	DC pulse	learning path
In	DC pulse (inout)	learning path
Out	DC pulse (inout)	learning path
VF	DC (in)	Forming signal, 3.3V
VP	DC (in)	Programming signal, 3.3V
VPin	DC pulse (in)	Programming pin
S	DC (in)	MUX selector, 3.3V
Vpad	DC pulse (inout)	learning path
Vpad2	DC pulse (inout)	learning path
Vbias	DC (input)	current compliance

Table D.2: Pin assignment of the test structures

Pin name	Pin Connection	Pin Direction	Pin name	Pin Connection	Pin Direction
t1	In	inout	b1	Out	inout
t2	Vpin	input	b2	VP	input
t3	VF	input	b3	Out(2)	inout
t4	VP(2)	input	b4	In(2)	inout
t5	VF(2)	input	b5	Vpin(2)	input
t6	Vpad	inout	b6	S	input
t7	Vbias	input	b7	Vpad2	inout
t8	Vpad(2)	inout	b8	S(2)	input
t9	Vbias(2)	input	b9	Vpad2(2)	inout
t10	Vpad(3)	inout	b10	S(3)	input
t11	Vbias(3)	input	b11	Vpad2(3)	inout
t12	VDD	-	b11	GND	-

Table D.3: Forming and programming scheme

V_F	V_P	F
0	X	*learning*
1	0	$Forming/SET$
1	1	$RESET$

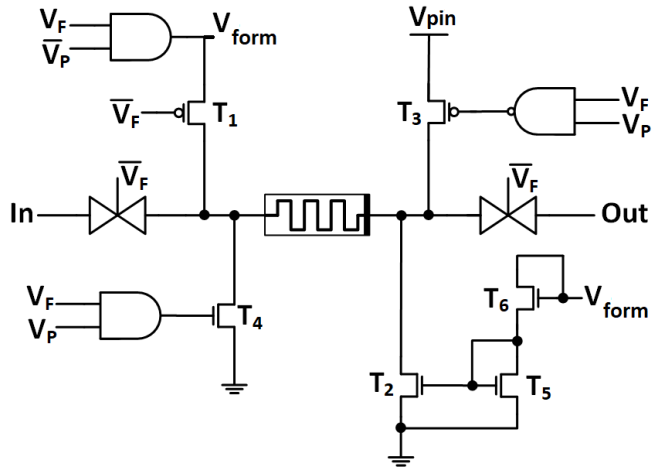

Figure D.7: Forming and Programming circuit I

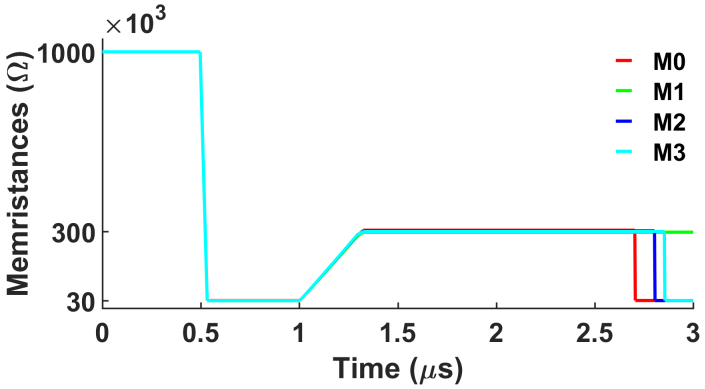

Figure D.8: Forming and programming of 4 ReRAM devices. M0, M2, M3 are programmed to LRS while M1 is kept at HRS

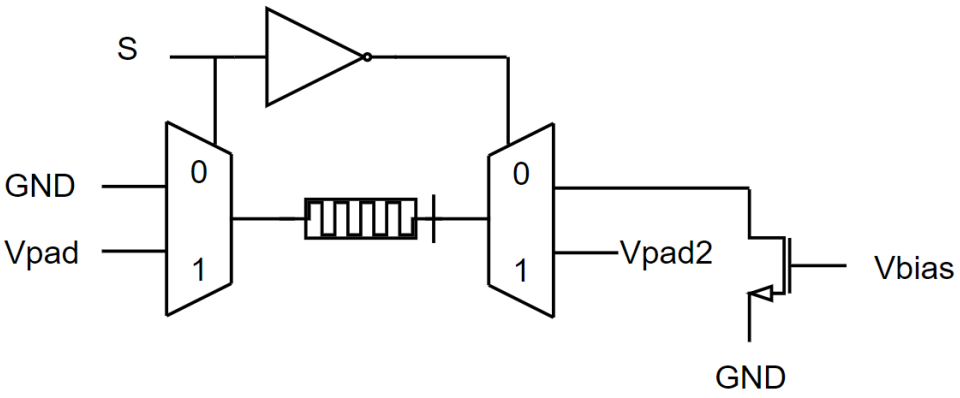

Figure D.9: Forming and Programming circuit II.

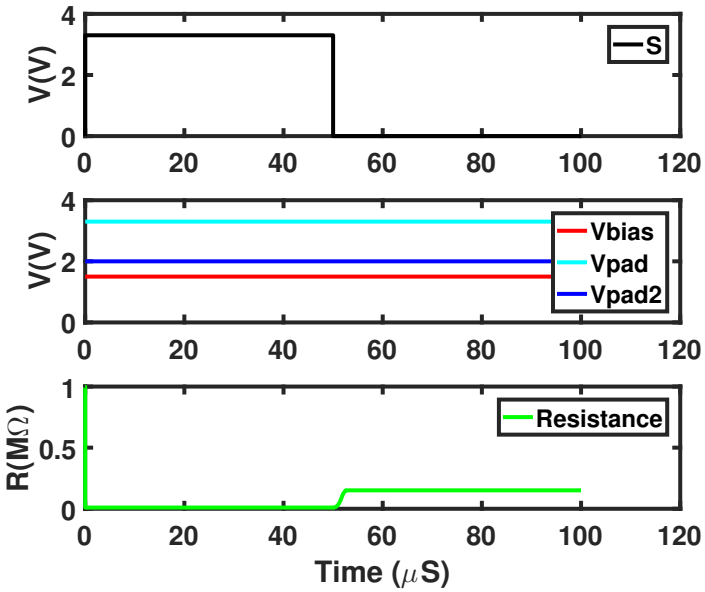

Figure D.10: Forming the memristor device to LRS then switching it to HRS.

Vita

Sherif Amer was born in Cairo, Egypt. He attended the American University in Cairo (AUC) where he received his BSc and MSc in Electronics Engineering in 2014 and 2016, respectively. Since August 2016, he has been a Research assistant and a Chancellor's fellow in the department of Electrical Engineering and Computer Science (EECS) at the University of Tennessee working towards his PhD. His work focuses on device modeling and circuit design of memristive neuromorphic systems. Starting August 2019, he will join Schweitzer Engineering Labs inc. as an Associate Electronics Design Engineer.