

Copyright  
by  
Cheng-Chih Hsieh  
2017

The Dissertation Committee for Cheng-Chih Hsieh  
certifies that this is the approved version of the following dissertation:

**Cerium Oxide Based Resistive Random Access Memory  
Devices**

Committee:

---

Sanjay K. Banerjee, Supervisor

---

Davood Shahrjerdi, Co-Supervisor

---

Jack Lee

---

Edward Yu

---

Gary Gibson

**Cerium Oxide Based Resistive Random Access Memory  
Devices**

by

**Cheng-Chih Hsieh,**

**DISSERTATION**

Presented to the Faculty of the Graduate School of  
The University of Texas at Austin  
in Partial Fulfillment  
of the Requirements  
for the Degree of

**DOCTOR OF PHILOSOPHY**

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2017

Dedicated to my family.



## Acknowledgments

I wish to thank the multitudes of people who helped me through these years. First and foremost, I would like to begin by thanking my advisor at University of Texas at Austin, Dr. Sanjay K. Banerjee, for his support during my graduate studies at U.T. Austin. His intelligence and kindness, along with his academic council and membership truly enabled the completion of my degree. I am very grateful for his guidance and support. Secondly, I would like to thank my co-advisor at New York University, Dr. Davood Shahrjerdi, for his insightful advice and passion in research. He continually and convincingly conveyed a spirit of adventure in regard to research and scholarship. Without his guidance and persistent help this dissertation would not have been possible. I would also like to thank Dr. Jack Lee, Dr. Edward Yu, and Dr. Gary Gibson for taking the time to serve on my dissertation committee.

I would like to thank administrative and technical staffs of the Microelectronics Research Center for their unflinching and continuous support. I sincerely thank Jean Toll for her assistance and support when dealing with University purchases, tuition issues, academic deadlines. The front office staff, led by Dr. James Hitzfelder, always provides professional and instant help when I have problems with logistics and safety concerns.

A from-the-bottom-of-my-heart thanks is well deserved for the technical

and cleanroom staffs of the Microelectronics Research Center. Garcia Ricardo, Jesse James, Bill Ostler, Johnny Johnson, and Marylene Pallard. It's impossible to keep cleanroom tools functional without their support and maintenance. They also trained me on several tools which are important for my research and kindly shared with me about tips for operating on those tools. Their relentless dedication are pillars for supporting the Microelectronics Research Center.

Next I would like to express my sincere gratitude to graduate students in the Microelectronics Research Center for their camaraderie. William Hsu, my best friend in my life, and also my colleague in the group. The days we studied together and chats in the cleanroom are unforgettable. Tanuj Travedi trained me essential skills for being a good teaching assistant of Integrated Circuits Nanomanufacturing Techniques. The memories we worked together for the course will last forever in my mind. My studies on cerium oxide would not have been possible without highest quality film grown by Dr. Anupam Roy. He is also a great mentor for life. Hema Chandra Prakesh and Amritesh Rai provided a lot of help for performing electrical characterization and logistics of lab inventory. Stephen Szczepaniak offered helps for correct my awkward grammar for the thesis. Rik Dey and Tanmoy Pramanik helped me grow cerium oxide thin films.

Last but not the least, tremendous gratitude goes to my family for all their endless love and encouragement. I am forever indebted to my parents for giving me the opportunities and experiences that have made me who I am. Smiles of my beloved daughter, Emma, always propel me forward to accom-

plish my dreams. At the end I would like express the deepest appreciation to my dear wife, Yoshie Mitsui, who spent sleepless nights with and was always my support in the moments when I need her. This journey would not have been possible if not for them, and I dedicate this milestone to them.

# Cerium Oxide Based Resistive Random Access Memory Devices

Publication No. \_\_\_\_\_

Cheng-Chih Hsieh, Ph.D.

The University of Texas at Austin, 2017

Supervisors: Sanjay K. Banerjee  
Davood Shahrjerdi

Resistive Random Access Memory (RRAM) is an emerging technology of non-volatile memory (NVM). Although the observation of metal oxide that can undergo an abrupt insulator-metal transition into a conductive state has been known for over 40 years, researchers started investigating those materials for memory applications in late 1990s. It has been considered as the next generation memory technology to replace current flash memory because RRAM has demonstrated feasible switching characteristics and potential to build high density arrays and also RRAM is also compatible with contemporary CMOS processes, which means RRAM can be integrated into current CMOS chips. While the structure of RRAM is a simple metal-insulator-metal (MIM) device, there are numerous materials that exhibit resistive switching. The switching behavior is not only dependent on the switching layer materials but also dependent on the choice of metal electrodes and their interfacial properties. Many

metal oxides such as hafnium oxide ( $\text{HfO}_2$ ), titanium oxide ( $\text{TiO}_2$ ), aluminum oxide ( $\text{Al}_2\text{O}_3$ ), nickel oxide ( $\text{NiO}$ ), tantalum oxide ( $\text{TaO}_2$ ) and etc. have been studied in details; however, some materials are unexplored such as cerium oxide. In addition to nonvolatile storage applications, RRAM is considered as one of essential elements for advancing neuromorphic computing because of its analog switching and retention characteristics. This thesis investigated  $\text{CeO}_x$ -based RRAMs, from its fundamental device characteristics to neuromorphic applications.

# Table of Contents

<b>Acknowledgments</b>	<b>v</b>
<b>Abstract</b>	<b>viii</b>
<b>List of Figures</b>	<b>xii</b>
<b>Chapter 1. Introduction</b>	<b>1</b>
1.1 Memory Technology . . . . .	1
1.2 Resistive Random Access Memory . . . . .	4
1.3 Current state of Research . . . . .	6
1.4 Chapter Overview . . . . .	9
<b>Chapter 2. Characterization of Cerium Oxide RRAM</b>	<b>13</b>
2.1 Introduction . . . . .	13
2.2 Fabrication . . . . .	14
2.3 Electrical and Physical Characterization . . . . .	15
2.4 Conclusion . . . . .	23
<b>Chapter 3. Bilayer Cerium Oxide RRAM</b>	<b>27</b>
3.1 Introduction . . . . .	27
3.2 Hafnium Oxide and Cerium Oxide Stacked Resistive Random Access Memory . . . . .	30
3.3 Conclusion . . . . .	42
<b>Chapter 4. Selector Device for RRAM</b>	<b>45</b>
4.1 Introduction . . . . .	45
4.2 S-type NDR Niobium Oxide . . . . .	46
4.3 Back-to-back Schottky Diode . . . . .	54
4.4 Conclusion . . . . .	62

<b>Chapter 5. Short Term Relaxation of RRAM</b>	<b>65</b>
5.1 Introduction . . . . .	65
5.2 Device Fabrication and Measurement . . . . .	66
5.3 Results and Discussion . . . . .	67
5.4 Conclusion . . . . .	74
<b>Chapter 6. Neuromorphic Applications</b>	<b>75</b>
6.1 Introduction . . . . .	75
6.2 Supervised and Unsupervised Learning . . . . .	76
6.3 Neural Networks . . . . .	79
6.4 Device Fabrication . . . . .	80
6.5 Single Hidden Layer ANN for Simple Image Recognition . . . . .	81
6.6 Multiclass Image Recognition . . . . .	86
6.7 Simple Convolutional Neural Network for Image Recognition . . . . .	88
6.8 Conclusion . . . . .	96
<b>Chapter 7. Conclusions</b>	<b>97</b>
7.1 Summary . . . . .	97
7.2 Recommendations for Future Work . . . . .	99
<b>Index</b>	<b>101</b>
<b>Bibliography</b>	<b>102</b>
<b>Vita</b>	<b>121</b>

# List of Figures

1.1	The current memory technology spectrum and emerging memory technologies. Note that register, cache and DRAM are considered as volatile memory, whereas SSD and HDD are considered as NVM. . . . .	2
1.2	(a) Schematic of MIM structure for metaloxide RRAM. (b) (c) Schematic of IV characteristics, showing two modes of operation: (b) unipolar and (c) bipolar . . . . .	6



- 1.3 Schematic of the possible electron conduction paths through a MIM stack. (1) Schottky emission: thermally activated electrons injected over the barrier into the conduction band. (2) FowlerNordheim (FN) tunneling: electrons tunnel from the cathode into the conduction band; usually occurs at high field. (3) Direct tunneling: electron tunnel from cathode to anode directly; usually occur when the oxide is thin enough. If the oxide has substantial number of traps (e.g., oxygen vacancies), trap-assisted tunneling contributes to additional conduction, including the following steps: (4) tunneling from cathode to traps; (5) emission from trap to conduction band, which is essentially the PooleFrenkel emission; (6) FN-like tunneling from trap to conduction band; (7) trap to trap hopping or tunneling, maybe in the form of Mott hopping when the electrons are in the localized states or maybe in the form of metallic conduction when the electrons are in the extended states depending on the overlap of the electron wave function; and (8) tunneling from traps to anode. Adapted from [1]. . . . . 11
- 1.4 (a) shows the analogy of biological a neuron/synapse connection and electrical representation of neuron/synapse. (b) shows a typical RRAM array representing an artificial neural network.[2] 12

2.1	Upper right panel shows that the etchant before etching reaction is white while lower right panel shows the etchant becomes Persian blue after etching. Left panels show color change of bulk $\text{CeO}_x$ film after etching. . . . .	16
2.2	(a) The typical dc sweep I-V characteristics of Al/ $\text{CeO}_x$ /Au structure. Arrows and numbers indicate the sweeping direction and order. The size of each device is $100 \mu\text{m}$ dot in diameter and sandwiched structure of 30 nm Au, 13nm $\text{CeO}_x$ , and 30nm Au. (b) Cross-sectional TEM image with Fast Fourier Transform (FFT) graphs in inset show thickness of each layer. Lattice constant is $5.4\text{\AA}$ at pristine state and $3.93\text{\AA}$ at LRS. . . . .	18

2.3 (a) Ce 3d core-level XPS spectra of MBE and PVD samples. The three-lobed envelopes structure of Ce 3d XPS spectra is due to multiple final states from  $Ce^{4+}$  and  $Ce^{3+}$ . Arrows and symbols are referred to different final states and oxidation state of Ce. Note that symbols with blue color such as  $u'''$ ,  $u''$ ,  $u$ ,  $v'''$ ,  $v''$ , and  $v$  are final states from  $Ce^{4+}$ ; symbols with green color such as  $u_o$  and  $v_o$  are final states from  $Ce^{3+}$ . Other two final states like  $l_o$  and  $t_o$  are not shown in this figure due to relatively low amplitude and little contribution to the three-lobed envelope structure. (b) O 1s XPS spectra of MBE and PVD samples. The binding energy value of peak of MBE sample is higher than PVD sample, which means MBE sample has more  $Ce^{3+}$  component than PVD sample. . . . . 21

2.4	<p>(a) HRS and LRS resistance along with switch cycles. The structure of tested device is 30 nm Al/13 nm PVD grown <math>\text{CeO}_x</math>/30 nm Au at room temperature. (b) HRS and LRS resistance under a continuous 300-mV read voltage. The structure of tested device is 30 nm Al/13nm PVD grown <math>\text{CeO}_x</math>/ 30 nm Au at room temperature and 150°C. The voltage sweep followed the numerical order in Fig. 2.2(a) for each cycle. Step 1: voltage ramps up to 2.5 V. Step 2: voltage ramps down to 0V, after this half cycle, voltage ramps up to 300mV to read and records the resistance. Step 3: voltage ramps up to -2.2 V. Step 4: voltage ramps down to 0V. Then, voltage ramps up to read and records the resistance. . . . .</p>	24
2.5	<p>(a) and (b) Diameter of device versus set voltage and reset voltage of MBE grown <math>\text{CeO}_x</math> RRAMs. (c) and (d) Diameter versus set voltage and reset voltage of PVD grown <math>\text{CeO}_x</math> RRAMs. For each area value, 20 devices have been tested. Inset graphs in (a) and (b) represent the LRS and HRS resistance dependence of MBE devices and inset graphs in (c) and (d) represent the LRS and HRS resistance dependence of PVD devices. . . . .</p>	25

2.6	(a) and (b) Switching layer thickness dependence of forming voltage with MBE and PVD grown thin film. (c) and (d) Set voltage is independent of $\text{CeO}_x$ film thickness for both MBE and PVD $\text{CeO}_x$ film. For each thickness value, 15 PVD and MBE grown $\text{CeO}_x$ RRAM devices have been tested. . . . .	26
3.1	Improving memristor device characteristics using an engineered sub-stoichiometric $\text{HfO}_x$ capping layer. (a) Schematic structure of two memristors with and without the engineered $\text{HfO}_x$ , conceptually illustrating the increase of the oxygen vacancy density in the $\text{CeO}_x$ switching layer. This attribute of the bilayer memristor results in the forming-free operation and the reduction of the Set voltage. XPS spectra of the (b) engineered $\text{HfO}_x$ , and (c) $\text{CeO}_x$ films with and without the $\text{HfO}_x$ capping layer. The XPS studies indicate the increase of the oxygen vacancy concentration in the $\text{CeO}_x$ film capped with the oxygen-deficient $\text{HfO}_x$ layer. (d) Representative current-voltage characteristics of two memristors, indicating the sub-1V operation of the bilayer memristive device. (e) Heat transfer simulations illustrate enhanced Joule heating in the bilayer structure, causing the marked reduction of the Reset voltage (scale bars are 2nm). The observed increase in Joule heating arises from the high thermal resistivity of $\text{HfO}_x$ at nanoscale. The thickness of $\text{HfO}_x$ is 2nm and the total thickness is 20nm. . . . .	34

3.2 Effect of  $\text{HfO}_x$  thickness ratio on the memristor device behavior. The data indicates that the optimal device characteristics ((a) forming voltage, (b) Set voltage, and (c) Reset voltage) occurs at the thickness ratio of about 0.1. Moreover, the device-to-device variation is reduced at this optimal thickness ratio. The equivalency of the forming and Set voltages at the optimal thickness ratio confirms the forming-free operation of the device. Low device variability is critical for implementation of large neural networks with high density of memristive synaptic connections. Therefore, we statistically examined the effect of the  $\text{HfO}_x$  thickness on the important device parameters: Set, Reset, and forming voltages. In these experiments, the  $\text{HfO}_x$  thickness was varied, while keeping the total thickness of the bilayer stack fixed at 20nm. The thickness ratio defined in this work is  $\text{HfO}_x$  thickness to total thickness in bilayer. The data in Figure 3.2 indicates that the insertion of an  $\text{HfO}_x$  capping layer with the optimal thickness ratio of about 0.1 significantly improves the uniformity of the key device parameters. Interestingly, this optimal thickness ratio also coincides with the minimum operating voltages of the bilayer structure. We surmise that the  $\text{HfO}_x$  film begins to act as an independent switching layer beyond this optimal thickness ratio, resulting in significant increase in both the device operating voltages and the device variability. Moreover, the Reset voltage begins to increase as the  $\text{HfO}_x$  film becomes thicker. This observation is in agreement with our heat transfer simulation results in Figure 3.3. In (c), the Reset voltage at 0.4 was too large compared to other ratios so it wasnt included. . . . .

3.3	Effect of HfO <sub>x</sub> film thickness on Joule heating. Numerical heat transfer simulation results for several bilayer HfO <sub>x</sub> /CeO <sub>x</sub> structures with varying HfO <sub>x</sub> to total thickness ratio at the bias voltage of -0.6V. The total thickness of the HfO <sub>x</sub> /CeO <sub>x</sub> stack was kept at 20nm. The Joule heating begins to diminish as the thickness of the HfO <sub>x</sub> was increased, which arises from the thickness dependence of the HfO <sub>x</sub> thermal conductivity. . . . .	37
3.4	Device reliability studies. (a) The endurance test results for the CeO <sub>x</sub> and the optimal HfO <sub>x</sub> /CeO <sub>x</sub> devices. In addition to the improved endurance properties, the bilayer device exhibits larger HRS and LRS values compared to the device with no HfO <sub>x</sub> . The increase of the LRS and HRS values is favorable for reducing the switching power consumption of the bilayer device. Besides, when bilayer device were tested in close loop with adaptive programming and relaxed on-off ratio, it survives over 5×10 <sup>7</sup> cycles without any degradation. (b) The accelerated retention test for the CeO <sub>x</sub> and the HfO <sub>x</sub> /CeO <sub>x</sub> devices measured at 150C at constant stress voltage of +0.2V. The results indicate projected data retention of 10 years for both devices. (c) Representative CDF plot of the cycle-to-cycle programing characteristics for two devices with and without the engineered HfO <sub>x</sub> layer. . . . .	39

3.5	<p>Analog memory characteristic of the bilayer memristor. The normalized conductance of (a) bilayer memristor is plotted as a function of pulse widths and amplitudes when the device switches from a, fully Off state to fully On state, and (b), fully On state to fully Off state. The dashed lines are guide to the eye and the hatched regions denote unmeasured points. The data in (a), and (b) reveal the gradual change in the conductance of the device between the fully Off and On states. Full On/Off switching energy consumptions of 2.6 and 2.1pJ were calculated from the transient (c), Set and (d), Reset voltage and current waveforms, respectively. . . . .</p>	41
3.6	<p>Switching cycles of intermediate states. Cycle test is conducted in both positive polarity and negative polarity to examine if intermediate states are separable and stable. The pulse voltage increased after every 100 cycles in (a) and (b). The resistance variation in each resistance state is very small, which means it is very stable. The separation of each resistance state is not linear with applied voltage. But it has potential to generate more resistance states if more voltage interval are added between each applied voltage properly. This also shows that the switching energy between each intermediate state is at femto joule scale . . . . .</p>	42



3.7	Implementation of STDP learning using the $\text{HfO}_x/\text{CeO}_x$ memristive device. (a) Schematic representation of the learning experiment. Two waveforms with identical shapes were applied to the top and bottom electrodes. In the learning experiments, the time intervals between the pre- and post-synaptic spikes were varied in order to probe the synaptic depression ( $\delta t < 0$ ) and potentiation ( $\delta t > 0$ ). The positive (negative) time difference indicates that the pre-synaptic spike occurs before (after) the post-synaptic one. (b) The plot clearly indicates the marked change in the synaptic strength as a function of different pre/post spike intervals. . . . .	43
3.8	Generation of waveforms for Spike-Timing-Dependent Plasticity (STDP) measurements. Multiple square waves with various pulse heights were tailored for emulating exponential decay pulses used in STDP studies. The time interval of the pulses on the top and bottom electrodes (that are pre- and post-neural spikes) was varied relative to each other while monitoring the conductance of the device. The pulses were created using the Keysight B1500 pulse generator unit. Depending on the time difference between the pulses illustrated in Figure 3.8(a) and (b), the device demonstrates long-term depression ( $\delta t < 0$ ) and potentiation ( $\delta t > 0$ ) . . . . .	44

4.1	Bright field cross-sectional TEM image of a representative NbO <sub>x</sub> selector. Active area of NbO <sub>x</sub> is assumed to be at a uniform temperature T <sub>N</sub> that is higher than the surrounding ambient temperature, T <sub>amb</sub> , due to Joule heating. This heated region is thermally connected to T <sub>amb</sub> through the effective thermal resistance, R <sub>th</sub> , and thermal capacitance, C <sub>th</sub> , of the surrounding device structures. . . . .	49
4.2	I-V curves of two different electroforming processes. Numbers indicate order of sweeps; arrows indicate time evolution. (a) Type I forming. This results in increasing currents as the initially amorphous Nb <sub>2</sub> O <sub>5</sub> is reduced through interaction with the TiN electrodes. (b) Type II forming. This includes crystallization to a more resistive tetragonal NbO <sub>2</sub> state after the initial reduction. The slope of curve 5 in (b) is positive at high currents due to a ~100 Ω resistance in series with the selector. . . . .	51
4.3	(a) Measured (solid) and calculated (dashed) V-I curves for TiN/NbO <sub>x</sub> /TiN selector with 52 nm diameter bottom electrode and 8 nm thick NbO <sub>x</sub> layer for T <sub>amb</sub> = 275–450 K. (b) Arrhenius plot of the conductance measured at biases low enough for it to be ohmic. . . . .	54

4.4	(a) DC I-V characteristics from two-dimensional numerical simulation by Sentaurus from Synopsys Inc.. Two orange dashed lines represent voltage values calculated for NL ratio. The choice of $1\text{MA}/\text{cm}^2$ is based on matching current density of RRAM devices. (b) I-V characteristics on Schottky barrier height dependence. Note that non-linear step size of simulation caused I-V curves to show a small hysteresis at low bias during positive polarity sweep, which is an artifact. (c) NL ratio extracted from Figure 4.4(a)(b). . . . .	58
4.5	(a) Band diagrams at equilibrium and forward bias on D1. The equivalent circuit diagram was shown at forward bias on D1. Series resistance is labeled as $R_S$ . $\phi_{b1}$ and $\phi_{b2}$ represents Schottky barrier height at D1 and D2. (b) Cartoon illustration of cross-sectional view of device. The green arrow indicates the effective area of device. . . . .	59

4.6	(a) Semilog plot of I-V characteristics of different thickness Ti–aSi–Ti diodes. The dashed circle is where series resistance effect becomes prominent. (b) I-V characteristics of Ti–aSi–Ti and Ni–aSi–Ni with 10 nm amorphous silicon. (c) Comparison of the NL ratio between Ti–aSi–Ti and Ni–aSi–Ni at different thickness. The inset of (c) is linear scale I-V characteristics of (b). (d) I-V curve fitting by using (7). (7) fits well at intermediate bias where series resistance effect was low and thermionic emission is dominant. . . . .	63
4.7	(a) Pulse measurement setup and impedance at each stage. The pulse generator has an output impedance of 50 $\Omega$ , and input channel impedance of oscilloscope is 50 $\Omega$ as well. (b) shows the transient current response of DUT. Black curve and arrow is input pulse generated by source and blue curve and arrow is current response of MSM diode. The device is a 300 nm $\times$ 300 nm crossbar. . . . .	64
4.8	(a) Test scheme of cycle test. (b) Endurance test up to $10^8$ cycles. (c) DC stress test at 0.9 V for 1000 seconds. . . . .	64

5.1	Green and black curves in Figure 5.1 are initial resistance of HRS and LRS after programming to RRAM cell. Blue triangles and red dots are resistance after given time delay. From left to right, the time delay between initial read and delay read is 100 $\mu$ s, 1ms ,and 1s, respectively. . . . .	66
5.2	Schematic illustration of the adaptive programming algorithm. Read pulses are 200ns at 0.2V. Write pulses start at 0.2V with 0.1V increment. The interval between each pulse is 10 $\mu$ s. At each attempt to switch the resistive state of the device, a write pulse is applied to the RRAM device followed by a read pulse to check whether the value of the read current is higher(lower) than the target value during the SET(RESET) cycles. If this condition is met, we consider this attempt as a successful SET (RESET). . . . .	68
5.3	Testing setup for fast sampling measurement. LabView software was used for programming and controlling FPGA and Python was used for data processing after data collecting by FPGA and LabView. . . . .	69
5.4	Distribution of HRS and LRS read currents for (a) 1R and (b) 1S1R structures. . . . .	70
5.5	RTN noise at different current amplitude of MSM selector. (a) operating at 12.8 $\mu$ A, (b) operating at 1 $\mu$ A. . . . .	70

5.6	CDF plots using probit units for (a) 1S1R, (b) 1R with 100ns forming pulse, and (c) 1R with $5\mu s$ forming pulse. The arrows labeled "time" in (a) indicate the temporal progress of the experiment. The horizontal arrows indicate the gap between tail bits of the HRS and the LRS. Probit of the HRS curves are plotted in a decreasing fashion versus current, while the LRS curves are plotted in an increasing fashion versus current. The read pulse voltage in (a) is 1.7V because of voltage drop at selector device. . . . .	71
5.7	Effect of the operating window and the forming pulse duration on the long-term reliability endurance of RRAMs. Devices were subject to the adaptive programming with a read pulse width of 200 ns. . . . .	73
6.1	The I-V characteristics of voltage controlled generic memristor can be derived from state dependent Ohm's law and state equation. The I-V curve shown here is from a sinusoidal voltage source.[3] . . . . .	77
6.2	Yellow circle represents supervised learning and blue circle represents unsupervised learning. The intersection of two circles is the "intermediate" learning. . . . .	78

6.3	The black crosses and smiley faces represent the input data for classification. The red dashed line is the SVM classifier. The gray triangles in the right diagram represent the input data for clustering. . . . .	79
6.4	Schematics representation of ANN. The activation function of synapse in the hidden layer is a delta function. The total number of input neurons is R. Total number of synapse in the hidden layer is M. . . . .	81
6.5	Optical microscope images of a $10 \times 10$ RRAM array are at left and 30 degree tilted and top view SEM images are at right. . .	82
6.6	The input training data was encoded to input signals for training neural networks. Red pixels in training data were converted to hot for non-inverted neural network and blue pixels were converted to hot for inverted neural network. . . . .	83
6.7	Testing data was encoded into input signals for each column, then passing to all columns in each neural network. The result is based on comparison of current between non-inverted array and inverted array. . . . .	84

6.8	(a) represent a RRAM device with nonlinear weight change function versus in a crossbar array. The dashed line in the middle represents the ideal linear weight change. (b) shows the device-to-device variation from 10 RRAM devices in the same array. (c) The prediction accuracy at different pulse voltage during the training. Note that there are two color groups in (c). The red one and blue one are results from different RRAM arrays. Note that the prediction accuracy is precision. Precision in statistics is defined by the ratio of true positives over true positives plus false positives. . . . .	85
6.9	(a) represents different magnitude of non-linearity used in TensorFlow simulation for (c). (b) shows the result by TensorFlow simulation of device-to-device variation on prediction accuracy. Moderate nonlinear weight change is 2 in (a) and strong nonlinear weight change is 4 in (a). (c) TensorFlow simulation of non-linearity on prediction accuracy. Note that the simulation assumes no device variation on RRAM synapses. . . . .	85
6.10	Training images used for multiclass image recognition. The left images are ideal patterns for each class. . . . .	87
6.11	Training scheme for multiclass image recognition. The input image encoded into signals and pass to the assigned columns in RRAM array. One column is trained with non-inverted signals, and the other is trained with inverted signals. . . . .	87



6.12	Testing scheme for multiclass image recognition. The non-inverted input passed to all columns. The prediction result is based on "One-vs-All" method, <i>i.e.</i> , each binary classifier provides its prediction and the final result is the one with maximum current. .	88
6.13	Prediction accuracy versus number of training data set. Note that the prediction accuracy of each letter is based on definition of precision. And averaged accuracy is based on marco-average of precision. . . . .	89
6.14	A Simple CNN architecture with depth of three. The input data is mapped to the feature layer with convolution by filter matrix. Then the feature layer is mapped to subsample layer by pooling. Finally, the classification is performed by a fully connected layer. The depth in CNN means the number of feature layers used for convolution. Each feature layer might have different filter matrix. . . . .	90
6.15	The input data is convoluted with the filter matrix and mapped to feature layer. Here, a $3 \times 3$ filter matrix is used for edge detection. The $3 \times 3$ pink matrix represents the feature layer. .	90

6.16	The synapses in the feature layer are divided into several pools, and pooling process chooses the maximum weight in each pool then mapping all to a subsample layer. The pool size is $2 \times 2$ and the size of feature layer is $4 \times 4$ , therefore, the subsample layer is $2 \times 2$ . Synapses in the subsample layer connects to a fully connected layer for classification. Synapses in fully connected layer sums up weights with blue arrows minus weights with red arrows. The fully connected layer provides classification result.	91
6.17	The filter matrix and connection between the subsample layer and the fully connected layer is adapted from the result by TensorFlow simulation. . . . .	93
6.18	Training scheme of the CNN. Note that during the training, the write voltage is always 1.2V and no connection to the subsample layer. . . . .	94
6.19	Testing scheme of the CNN. The read voltage is 0.2V and the output is based on the comparison of current from each subsample layer. . . . .	94
6.20	(a) Prediction accuracy of regular testing data set between single hidden layer ANN and CNN. (b) Prediction accuracy of 50% ambiguous testing data set between single hidden layer ANN and CNN. . . . .	95

6.21 (a) Prediction accuracy of CNN with different filter matrix for ambiguous testing data set. The filter matrix is also shown in (a), the one at left is identity filter and the one at right is sharpen filter. (b) Prediction accuracy of each class for CNN with sharpen filter. . . . . 95

# Chapter 1

## Introduction

### 1.1 Memory Technology

In the "Big Data" era, massive amount of data is stored in data centers and computed at the server end, which is also referred to as cloud computing. In future, the majority of data mining and machine learning problems will be solved in the "cloud" instead of local machines because data set is too large to be handled efficiently on a local machine. The performance of cloud computing will strongly depend on the performance of the memory architecture and the memory technologies. Figure 1.1 shows the spectrum of the current memory technologies and emerging memory technologies: spin-transfer torque magnetic random access memory (STT-MRAM), resistive random access memory (RRAM) and phase change random access memory (PCRAM). The left end of the spectrum represents fast, small capacity and volatile memory technologies, on the other hand, the right end of the spectrum represents slow, large capacity and non-volatile technologies. Current L1/L2/L3 cache, dynamic random access memory (DRAM), and solid state drive (SSD) are based on complementary metal-oxide-semiconductor (CMOS) technology. These conventional memory technologies might not be able to meet the requirement with device scaling. Thus, these emerging memory technologies have been investigated

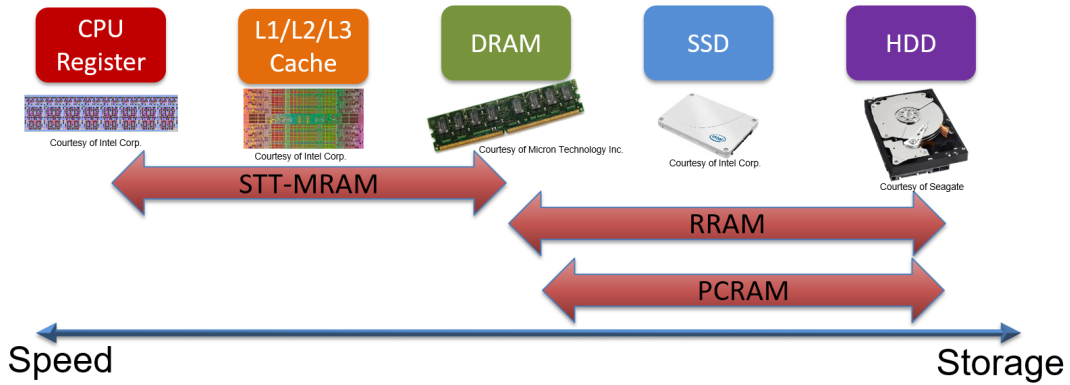


Figure 1.1: The current memory technology spectrum and emerging memory technologies. Note that register, cache and DRAM are considered as volatile memory, whereas SSD and HDD are considered as NVM.

wildly for volatile and non-volatile memory (NVM) applications.

STT-MRAM, PCRAM and RRAM are most popular candidates among all potential technologies and share some common properties such as: two-terminal device structure, non-volatile memory, device switches between a high resistance state (HRS) and a low resistance states (LRS). Two different resistance states can be used to store "0" and "1" for digital memory device. Nevertheless, the physics behind resistance state change is quite different: STT-MRAM utilizes the parallel configuration and anti-parallel configuration of two ferromagnetic layers separated by a tunneling insulator layer to generate LRS and HRS. PCRAM uses chalcogenide compound for switching between the crystalline phase and the amorphous phase to provide LRS and HRS. RRAM switches between LRS and HRS by modulating conductive filament length or interracial reaction on electrode to exchange oxygen de-

fects. These three technologies have different I-V characteristics because of their switching mechanisms. Table 1.1 compares the device performance of conventional CMOS based memory technologies and emerging memory technologies. In Figure 1.1 the arrow of STT-MRAM has overlap with register and L1/L2/L3 cache, and the reason is well explained in Table 1.1. STT-MRAM has smaller area, lower switching energy, similar endurance and comparable read/write time compared to SRAM. Next, RRAM and PCRAM have similar overlap with SSD, HDD and DRAM in Figure 1.1, and it can be seen that both RRAM and PCRAM are better in terms of area, switching energy, read/write time and endurance. These make RRAM and PCRAM very attractive to replace NAND Flash for storage. Recently Intel Inc. and Micron Technology Inc. announced transistor-less NVM product, which is called 3D XPoint™. The product could use PCRAM or RRAM technology based on patents filed by both companies[4–6] and the array architecture revealed. Intel Inc. claimed that the 3D XPoint NVM can be part of main memory like DRAM[7]. This indicates that RRAM and PCRAM have potentials to replace DRAM. Although in Table 1.1 DRAM has faster read/write time and lower switching energy, RRAM and PCRAM do not require refresh to retain data; this can significantly improve the efficiency of the computing system. Furthermore, this will blur the borderline between volatile and non-volatile memory. These emerging NVM technologies can be revolutionary for memory subsystem design.

Table 1.1: Performance metrics of current and emerging memory technologies

	SRAM	DRAM	NAND Flash	STT-MRAM	PCRAM	RRAM
Area (F <sup>2</sup> )	140	6-12	1-4	20	4-16	<4
Switching energy (pJ)	0.0005	0.005	10-100	2-25	2-25	2.6
Read Time (ns)	0.1-0.3	10	25000	<2	10-50	<10
Write time (ns)	0.1-0.3	10	220000	<10	50-500	25
Retention	N/A	<1s	yrs	yrs	yrs	yrs
Endurance	>10 <sup>16</sup>	>10 <sup>16</sup>	10 <sup>4-6</sup>	>10 <sup>15</sup>	10 <sup>9</sup>	10 <sup>8</sup>

## 1.2 Resistive Random Access Memory

As mention in the previous section, RRAM has been considered as one of the promising candidates for the next generation of NVM because of the fast write and read rates, long data retention time, and feasible scalability[8–10]. The resistive switching phenomenon was observed by Hickmott[11] in 1960s. The structure of RRAM is a simple metal-insulator-metal (MIM) device, and there are numerous materials that exhibit resistive switching. Figure 1.2 (a) provides the schematic of a common metal oxide RRAM cell.[1] The basic operation of RRAM is as follows: For the pristine samples in its initial resistance state, a larger voltage is needed to trigger the resistive switching behaviors for the subsequent cycles. This is called the electroforming or forming process. Forming is the process which forms certain phase or configuration so that the switching layer becomes conductive. After forming process RRAM is at high resistance state (HRS). The switching event from HRS to low resistance stae (LRS) is called the set process, which usually requires lower voltage. Conversely, the switching event from LRS to HRS is called the reset process. The switching modes of metal–oxide RRAM can be broadly classified into two switching modes: unipolar and bipolar, which depend on the choice of

materials stack for the RRAM cell. Unipolar switching means the switching polarity of set and reset processes is the same, and set or reset happens at certain voltage amplitude. If the unipolar switching can symmetrically occur at both positive and negative polarity, then it is also called nonpolar switching. Bipolar switching means the switching polarity of set and reset depends on the polarity of the applied voltage. So set can only occur at one polarity and reset can only occur at the other polarity. For either switching modes, to avoid a permanent dielectric breakdown in the set process, the common practice is to apply current compliance if voltage sweep is selected (voltage compliance if current sweep is selected). This usually can be done by the semiconductor parameter analyzer, or transistor or rectifying diode. To read the data from the cell, similar to a flash memory, a small read voltage is applied to determine the memory state (HRS or LRS), without disturbing the memory state. Figure 1.2 (b)(c) demonstrates the concept of unipolar and bipolar switching. Note that even the same switching layer can demonstrate both unipolar and bipolar with different combinations of metal electrodes or different stoichiometric ratio of switching layer.

Many reports have studied the current transport mechanisms in LRS and HRS. Most reports show a linear or ohmic relationship in the LRS. However, the conduction characteristics in HRS are quite different: Poole–Frenkel emission, [12, 13] Schottky emission [14, 15], the space charge limited current (SCLC) characteristic[16, 17] were observed in various metaloxide RRAMs. Figure 1.3 summarizes possible transport mechanisms in RRAM.



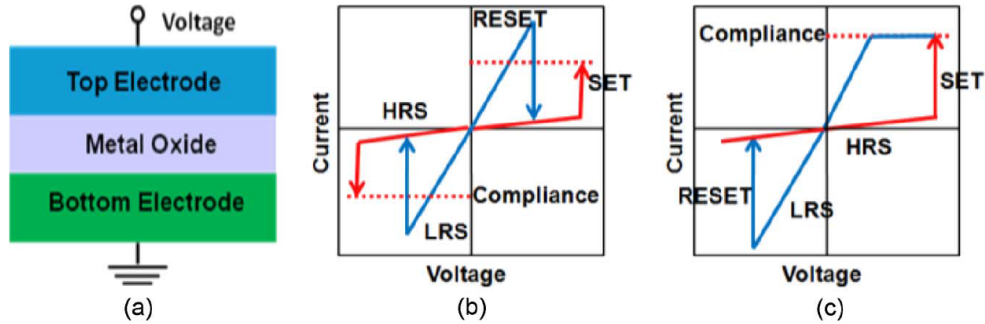


Figure 1.2: (a) Schematic of MIM structure for metaloxide RRAM. (b) (c) Schematic of IV characteristics, showing two modes of operation: (b) unipolar and (c) bipolar

### 1.3 Current state of Research

High dielectric constant binary transition metal oxide RRAMs have attracted a lot of interest because of operating voltage compared with silicon dioxide RRAMs and  $\text{Pr}_{0.7}\text{Ca}_{0.3}\text{MnO}_3$  (PCMO). Besides, the fabrication is compatible with current CMOS process[18–20]. There are some materials also demonstrate decent resistive switching characteristics, for example, cerium oxide.

RRAM is a memory technology that can be integrated with conventional CMOS in a simple way, using a material set compatible with the conventional CMOS fabrication environment and process temperatures that allow its integration at back-end-of-line (BEOL). Because of its low-temperature process, RRAM is often envisioned to be stacked in 3-D in a crossbar architecture with an effective memory cell area of  $4F^2=n$ , where  $n$  is the number of 3-D-stacked memory layers[21]. At the system level, it is envisioned that

a revolution in memory hierarchy and system architecture will be realized by this low-cost, BEOL-compatible, nonvolatile memory with tens of nanosecond bit-alterable read/write speed, over  $10^6$  endurance cycles, and potentially low power/energy consumption.

To achieve implementing RRAM, many efforts have been made on memory cell design for optimal memory array design, 3D crossbar architecture. A current limiter, which can optimally constrain the forming/set current, is necessary for the filamentary switching device to prevent the degradation of HRS and even the failure of the memory device. A serial transistor with RRAM cell is a good candidate of current limiter than an external electrical measurement instrument because of faster response and very large resistance at saturation region. Special consideration must be taken for design of memory cell of one-transistor-one-RRAM (1T1R) structure to avoid having a large parasitic capacitance between the transistor and the RRAM. The parasitic capacitance causes overshoot current during the forming/set process which in turn increases the reset current. Specifically, during the forming/set process, the RRAM resistance changes instantly while the voltage drop across the RRAM cannot drop instantly due to the presence of parasitic capacitance. Therefore, during the overshoot period that the voltage across the RRAM gradually decreases, excessive oxygen vacancies form and the conductive filaments (CFs) tend to grow laterally and increase in diameter or multiple CFs can be generated. Another popular candidate as current limiter is a bidirectional diode with non-linear I-V characteristics for a bipolar switching device

or a p-n diode for unipolar/nonpolar switching device. The advantage of this approach is that this has no area overhead compared to 1T1R scheme and it is easy to fabricate for 3D memory arrays. Such bidirectional diode should have high current density when it turns on and matching operating voltage with RRAM cell. It should not break down while forming process takes place (or RRAM cell is forming-free). There are still no conclusive results on such a bidirectional diode although many promising devices have been proposed, such as mixed ionic electronic conduction (MIEC)[22], VO<sub>2</sub> utilizing metal-insulator transition[23], Ovonic threshold switch (OTS) [24] and Schottky diode[25].

What makes RRAM such an exciting technology is not only NVM applications, but also great potential for realizing neuromorphic computing paradigm. RRAM device behaves similarly to what memristor[26] should be like: RRAM retains its resistance state during switching. Some RRAM devices share similar properties with biological synapses, which is analog switching in conductance. This raises a lot of interest in RRAM for building neuromorphic computing system. Traditional Von-Neumann architecture has bottleneck in parallel computing due to the bandwidth limit between central processing unit (CPU) and memory. On the other hand, the brain-like architecture which emulates human brain functions overcome the bandwidth limit. Neuromorphic architecture provides several benefits in machine learning problems. One recent progress in machine learning is the development of deep learning. Deep learning utilizes the architecture of human brain neural networks, which comprises pre-neurons, synapses and post-neurons. Pre-neurons and post-neurons

propagate or generate electrical signals to synapses and synapses are the key elements for learning. They received signals and update their weight accordingly. RRAM with analog switching is the ideal device to perform weight updating because of the similarity to biological synapse. Figure 1.4 shows a RRAM crossbar array for neuromorphic computing. Top and bottom electrode connect to CMOS neuron circuits; RRAM, which is called memristor in (a) is sandwiched by bottom electrode and top electrode, represents an electrical synapse.

## 1.4 Chapter Overview

This thesis investigates the fundamental electrical and physical characterization of cerium oxide based RRAM devices, bilayer RRAM devices, selector devices, short-term relaxation issue, and lastly, neuromorphic applications. This chapter introduced background and current key issues related to NVM and RRAM technology. Chapter 2 studies the basic  $\text{CeO}_x$  RRAM device, starting from device fabrication, materials characterization to electrical characterization. Performance metrics are discussed in Chapter 2 as well. Chapter 3 addresses the issue of non-analog switching in reset operation for simple  $\text{CeO}_x$  RRAM. The novel bilayer  $\text{HfO}_x/\text{CeO}_x$  is proposed in Chapter 3, and the bilayer device provides several benefits compared to single layer  $\text{CeO}_x$  RRAM. Chapter 4 proposes metal-semiconductor-metal (MSM) diode as selector device to eliminate leakage current in RRAM array. Chapter 5 discusses the origin of short-term relaxation issue in RRAM and combined RRAM/selector

structure. Chapter 5 demonstrates a potential solution to short-term relaxation. Chapter 6 demonstrates the fabrication of RRAM arrays for neural network based pattern classifiers. Various configurations of neural network are proposed and demonstrated. A conclusion is given in Chapter 7.

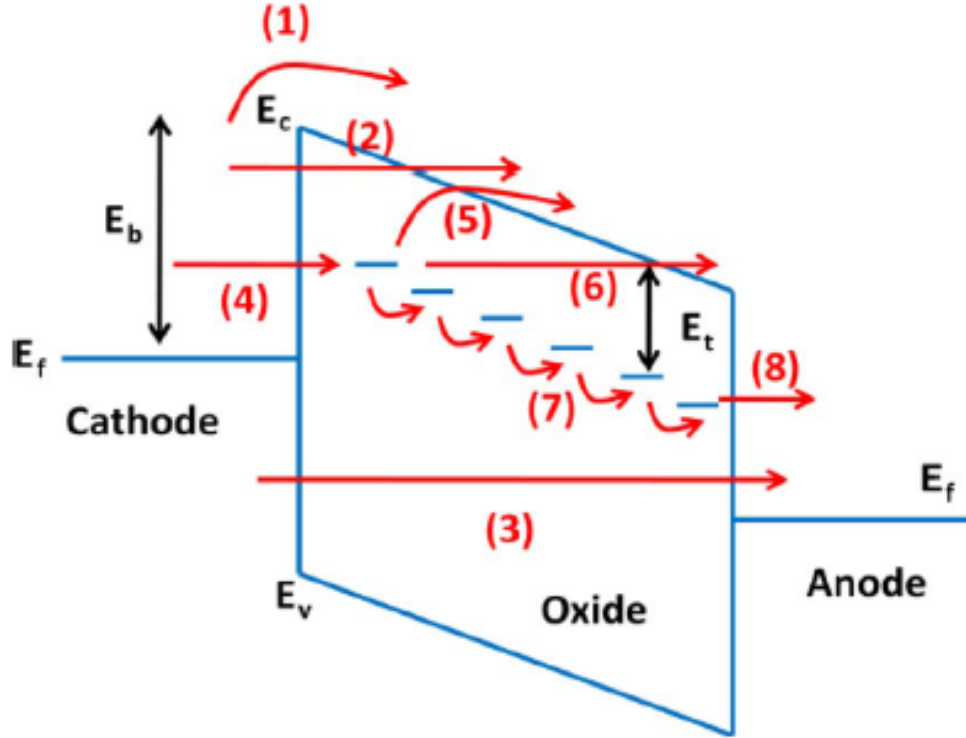


Figure 1.3: Schematic of the possible electron conduction paths through a MIM stack. (1) Schottky emission: thermally activated electrons injected over the barrier into the conduction band. (2) Fowler-Nordheim (FN) tunneling: electrons tunnel from the cathode into the conduction band; usually occurs at high field. (3) Direct tunneling: electron tunnel from cathode to anode directly; usually occur when the oxide is thin enough. If the oxide has substantial number of traps (e.g., oxygen vacancies), trap-assisted tunneling contributes to additional conduction, including the following steps: (4) tunneling from cathode to traps; (5) emission from trap to conduction band, which is essentially the Poole-Frenkel emission; (6) FN-like tunneling from trap to conduction band; (7) trap to trap hopping or tunneling, maybe in the form of Mott hopping when the electrons are in the localized states or maybe in the form of metallic conduction when the electrons are in the extended states depending on the overlap of the electron wave function; and (8) tunneling from traps to anode. Adapted from [1].

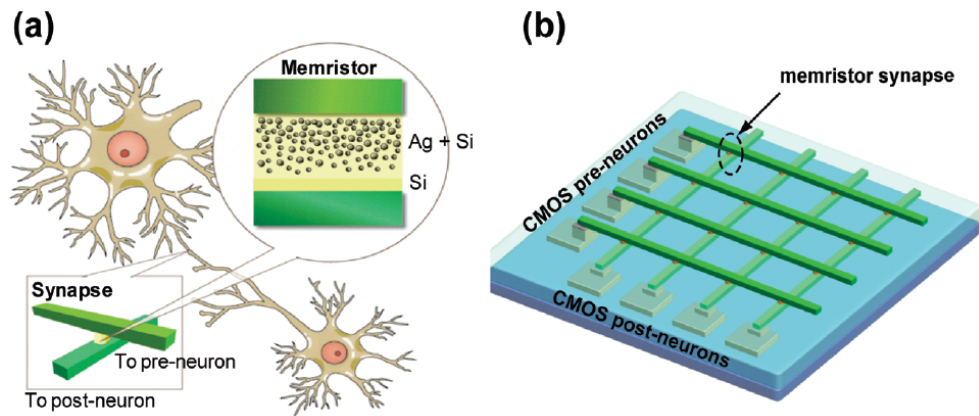


Figure 1.4: (a) shows the analogy of biological a neuron/synapse connection and electrical representation of neuron/synapse. (b) shows a typical RRAM array representing an artificial neural network.[2]

## Chapter 2

# Characterization of Cerium Oxide RRAM

### 2.1 Introduction

Similar to other high dielectric constant binary transition metal oxides, such as hafnium oxide and titanium oxide, cerium oxide has high dielectric constant and several valance states, making cerium oxide a potential material for RRAM application. Nevertheless, fundamental characterization of  $\text{CeO}_x$  based RRAMs, i.e., the scalability, reliability, and mechanism, has been only partially reported[27–29]. In this chapter, we demonstrate key characteristics of  $\text{CeO}_x$  RRAMs. One of the prevailing explanations of resistive switching is the formation of filament path in the switching layer. The typical RRAM device can be described as a sandwich MIM structure: an insulator thin film stacking on the first metal layer followed by second metal layer above the insulator film. The first metal layer is called bottom electrode (BE), and the second metal layer is called top electrode (TE). When the bias is applied on the insulator, oxygen vacancies build up filament paths, which change the resistance of MIM. This process is reversible. For bipolar RRAMs, applying the opposite polarity can reverse the process; the oxygen vacancies move back to bulk region of insulator. Certain types of MIM structure show unipolar switching. Most binary transition metal oxide based RRAMs show bipolar switching.



Although some groups have already reported resistive switching of  $\text{CeO}_x$  based RRAMs, there are still many unanswered questions and interesting behaviors of the system needed to be addressed, since cerium oxide is a relatively new material in the field. In this work, we report scalability, thickness dependence, endurance test, and mechanism discussion.

## 2.2 Fabrication

Devices were fabricated on a n-type Si (111) substrate with 300 nm plasma-enhanced chemical vapor deposition (PECVD) grown silicon dioxide on the top. Gold was used as the bottom electrode and deposited above the silicon dioxide layer by electron beam evaporation. Cerium oxide thin film was then deposited at room temperature (RT) by physical vapor deposition (PVD) in *in situ* oxygen plasma and molecular beam epitaxy (MBE) at 500 degree in Celsius. Then, devices were patterned by photolithography using AZ-5209 photoresist, followed by developing in AZ 726 MIF developer. The top electrode (500 nm - 300  $\mu\text{m}$  in diameter) was deposited by electron beam evaporation above the cerium oxide layer, and put in an acetone bath at room temperature for a day to lift off photoresist, followed by washing the device with acetone and isopropyl alcohol. The final step is to wet etch  $\text{CeO}_x$  by using the top electrode as hard mask; the etching solution is the mixture of hydrochloric acid, potassium hexacyanoferrate,  $\text{K}_4[\text{Fe}(\text{CN})_6]\cdot 3\text{H}_2\text{O}$ , and water. The mechanism of etching is based on cyanide ions forming coordination compounds with cerium ions. The detailed chemical reactions and solution

preparation were illustrated in the article by A.Kossoy *et al.*[30] Figure 2.1 shows color change of etchant and sample before and after reaction.

### 2.3 Electrical and Physical Characterization

Electric characterization of devices were taken by Agilent Semiconductor Parameter Analyzer B1500 and Lakeshore CRX-VF Probe Station. Devices were measured by applying voltage to the top electrode while the bottom electrode is grounded. For most RRAMs, there are two basic operation modes called set and reset. When applying bias upon certain point, the current through RRAM suddenly rises; this process is called set. After set process, RRAM is at LRS. If RRAM is bipolar when the bias is applied with the opposite polarity relative to the bias during the set process, the current will suddenly drop after a certain point. This process is called reset. After reset, RRAM is at HRS. The HRS appeared after reset process can be changed again to LRS by set. The resistance change between LRS and HRS is repeatable and nonvolatile. For the first voltage application on RRAM, usually it requires higher voltage to switch from HRS to LRS, this process is called forming and it is usually considered as a soft breakdown of the MIM structure. Dimension of each device is the same and MBE grown  $\text{CeO}_x$  RRAM shows lower set and reset voltage and larger operation window than PVD  $\text{CeO}_x$  counterpart. The operation window is defined as the ratio of resistance between HRS and LRS. These phenomena can be explained by difference in oxygen point defect mobility in  $\text{CeO}_x$  film.[31] The decrease of the mobility of oxygen vacancy may result in

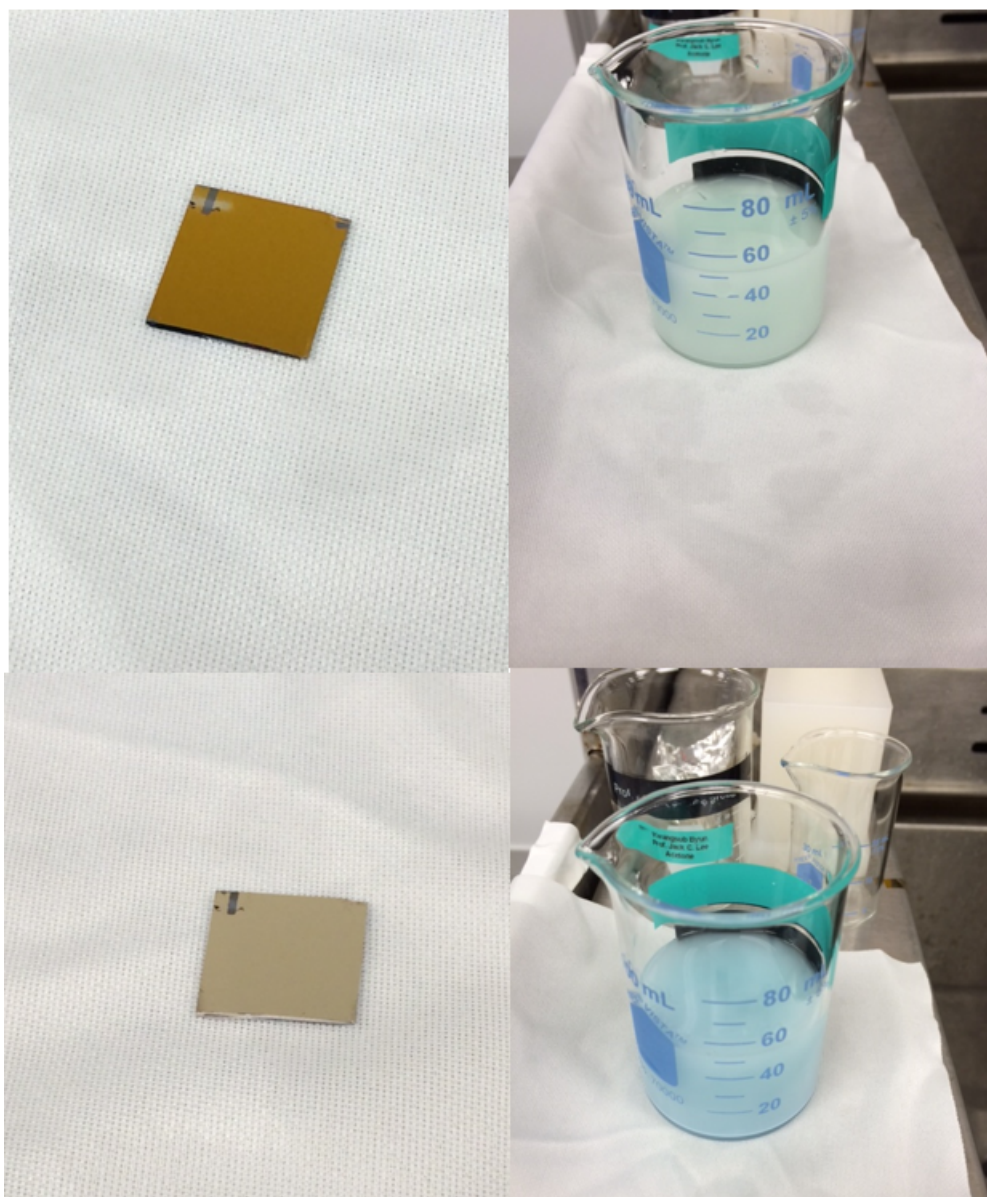


Figure 2.1: Upper right panel shows that the etchant before etching reaction is white while lower right panel shows the etchant becomes Persian blue after etching. Left panels show color change of bulk  $\text{CeO}_x$  film after etching.

higher voltages needed to drive oxygen vacancies in the dielectric layer. MBE  $\text{CeO}_x$  film has better interfacial properties at  $\text{CeO}_x/\text{Al}$  interface, and thus less pinning of oxygen vacancies. Superior operation window observed in MBE  $\text{CeO}_x$  can be explained by difference in the effective defect levels. In the work by Goux *et al.*[32],  $\text{CeO}_x$  film with high temperature process ( $750^\circ\text{C}$ ) tends to have higher activation energy ( $E_a = 320 \text{ meV}$ ), which means point defects are more trapped. A similar mechanism may result in higher resistance values for the HRS in MBE film. Compared with other high- $k$  dielectric based RRAM devices, such as  $\text{TiO}_2$ ,  $\text{AlO}_3$ , and  $\text{HfO}_x$ , both MBE and PVD  $\text{CeO}_x$  based RRAMs have competitive value of the operation window ( $10^4$  for PVD  $\text{CeO}_x$  and  $10^6$  for MBE  $\text{CeO}_x$ ). MBE  $\text{CeO}_x$  RRAM shows low set and reset voltage (1.1 V for set, 0.9 V for reset, and 2.8 V for forming), making  $\text{CeO}_x$  RRAM a suitable candidate for low power consumption applications. Figure 2.2(a) represents a typical dc-sweep of  $\text{CeO}_x$  RRAM device. The physical dimension of each device is the same and MBE grown  $\text{CeO}_x$  RRAM shows lower set and reset voltage and larger operation window than PVD  $\text{CeO}_x$  counterpart. Figure 2.2(b) shows cross-sectional transmission electron microscope (TEM) image of MBE grown  $\text{CeO}_x$  RRAM at pristine state and LRS state. Insets of Figure 2.2(b) show crystalline structure at each state. Lattice constant of crystalline structure near  $\text{Al}/\text{CeO}_x$  interface calculated from analytic software indicates that  $\text{CeO}_2$  is dominant phase at pristine state and more  $\text{Ce}_2\text{O}_3$  at LRS. In addition,  $\text{CeO}_x$  based RRAM devices can switch between HRS and LRS at 50ns pulse width, which indicates that the switching speed can go up

to 20 MHz.

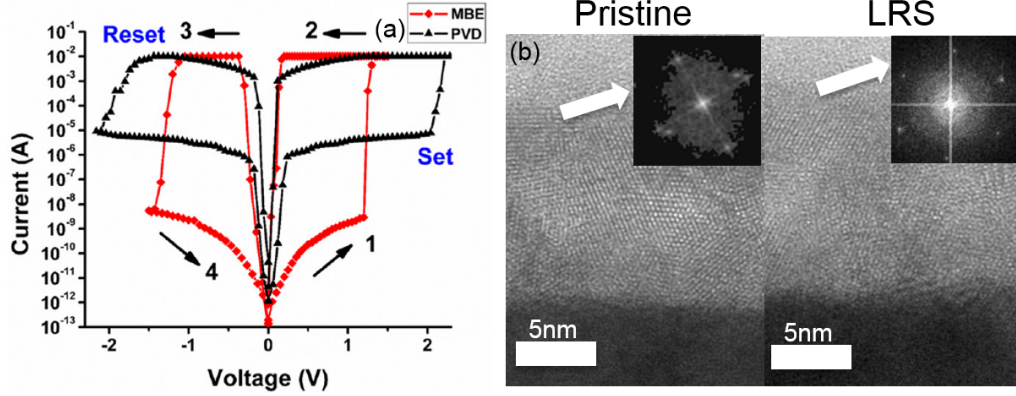


Figure 2.2: (a) The typical dc sweep I-V characteristics of Al/CeO<sub>x</sub>/Au structure. Arrows and numbers indicate the sweeping direction and order. The size of each device is 100 μm dot in diameter and sandwiched structure of 30 nm Au, 13nm CeO<sub>x</sub>, and 30nm Au. (b) Cross-sectional TEM image with Fast Fourier Transform (FFT) graphs in inset show thickness of each layer. Lattice constant is 5.4Å at pristine state and 3.93Å at LRS.

There are two prevailing hypothesis for the resistive switching mechanism in metal oxide based RRAMs. The first one is filamentary conduction in metal oxide. In this model, oxygen vacancies generated by applying a bias form filamentary conduction paths in dielectric. The formation and rupture of filament enable resistive switching in metal oxide based RRAM. The second one is called interface model. The change of oxidation state in memory materials causes resistive switching. In NiO system by Kinoshita *et al.*[33], the resistance switching happened on the anodic side of the conductive filaments in NiO, which implied that the electrochemical reaction involved in resistive switching. The change in the oxidation state is analogous to anodic oxidation.

To understand the conduction mechanism and reason performance discrepancy between MBE  $\text{CeO}_x$  and PVD  $\text{CeO}_x$ , the chemical composition and chemical bond condition of dielectric layer have been investigated by X-ray photoelectron spectroscopy (XPS). All XPS spectra were acquired at room temperature by Vacuum Generator Scientific SCALAB Mark II system and monochromatic Al  $K_A$  ( $h\nu = 1486.7$  eV) X-ray radiation source. The background pressure was kept below  $7.5 \times 10^{-8}$  Torr.[34] And the pass energy for high resolution spectra of Ce 3d and O 1s was 50 eV and 20 eV. Figure 2.3(a) shows Ce 3d XPS spectra of both MBE and PVD grown  $\text{CeO}_x$ . The typical Ce 3d XPS core-level spectra have three-lobed envelopes (around 882–890 eV, 895–910 eV, and 916 eV) due to different final states of a mixed valency.[35, 36] Note that  $u'''$  ( $v'''$ ),  $u''$  ( $v''$ ), and  $u$  ( $v$ ) represent  $\text{Ce}^{4+}$  final states:  $\text{Ce } 3d^9 4f^0 \text{ O } 2p^6$ ,  $\text{Ce } 3d^9 4f^1 \text{ O } 2p^5$ , and  $\text{Ce } 3d^9 4f^4 \text{ O } 2p^4$ , respectively, for  $\text{Ce} 3d_{3/2}$  and  $\text{Ce} 3d_{5/2}$ . Besides  $\text{Ce}^{4+}$  has three final states,  $\text{Ce}^{3+}$  final states also appear:  $3d^9 4f^1 \text{ O } 2p^5$  and  $3d^9 4f^2 \text{ O } 2p^4$ , expressed as  $u'$  ( $v'$ ) and  $u_o$   $v_o$ . Final states configuration of Ce 3d XPS spectra have been studied previously[37, 38], and the focus on Ce 3d XPS spectra in this work is to distinguish and analyze the difference between MBE and PVD grown  $\text{CeO}_x$  thus discuss the mechanism of resistive switching in Al/ $\text{CeO}_x$ /Au system. Arrows in Figure 2.3(a) indicate different final states of Ce 3d XPS spectra and  $u_o$  and  $v_o$  have stronger influence on Ce 3d core-level XPS spectra than  $u_o$  and  $v_o$  do. The deconvoluted components of Ce 3d XPS spectra structure have been discussed by Adnot and Bernis and Hasegawa et al.[35] In Figure 2.3(a), MBE Ce 3d XPS spectra

have higher amplitude of peaks than PVD Ce 3d spectra at 904 eV and 885 eV, which corresponds to  $\mu'$  and  $\nu'$  components. In addition, O 1s spectra also show similar trend between MBE and PVD samples. Figure 2.3(b) shows O 1s spectra of both MBE and PVD grown  $\text{CeO}_x$  thin film. O 1s peak of PVD  $\text{CeO}_x$  is less steep at lower binding energy side than O 1s peak of MBE  $\text{CeO}_x$  at lower binding energy side. The binding energy of O 1s in  $\text{CeO}_x$  and  $\text{Ce}_2\text{O}_3$  has been reported as 529.2 eV and 530.3eV, respectively. The difference of the binding energy implies that O 1s peak of MBE grown  $\text{CeO}_x$  has more  $\text{Ce}_2\text{O}_3$  than PVD grown  $\text{CeO}_x$ . From XPS spectra results of Ce 3d and O 1s, it can be concluded that MBE sample includes more  $\text{Ce}^{3+}$  than PVD sample. A report by Yoshitake et al. about cerium oxide based RRAM claims that the reduced cerium would inhibit the resistive switching in cerium oxide; however, in Al/ $\text{CeO}_x$ /Au system, the influence of  $\text{Ce}^{3+}$  is not obvious. MBE sample shows even lower operation voltage and better operation window than PVD sample. Although  $\text{Ce}^{3+}$  is presented in dielectric layer, the majority oxidation state of cerium is still  $\text{Ce}^{4+}$ . So the presence of  $\text{Ce}^{3+}$  would not deteriorate the resistive switching behavior of  $\text{CeO}_x$ .

One of the requirements for potential RRAM device is switching reliability. Two basic tests to demonstrate device reliability are the cycling test and the data retention test. 13 nm PVD grown  $\text{CeO}_x$  film 10  $\mu\text{m}$ -size cell is tested for cycling test at room temperature. The endurance test was carried out by automated program of 1 $\mu\text{s}$  pulses with set voltage at 2.5V and reset voltage at -2.1V by Agilent Semiconductor Parameter Analyzer B1500. Re-

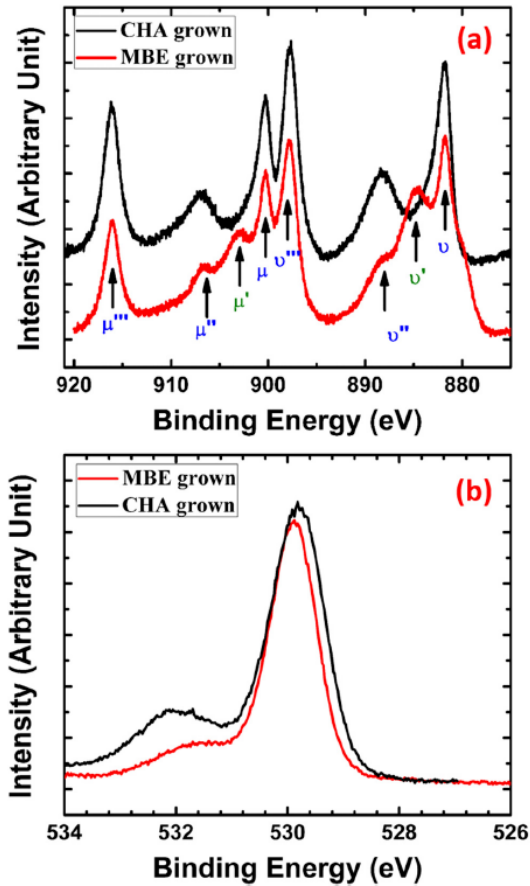


Figure 2.3: (a) Ce 3d core-level XPS spectra of MBE and PVD samples. The three-lobed envelopes structure of Ce 3d XPS spectra is due to multiple final states from  $\text{Ce}^{4+}$  and  $\text{Ce}^{3+}$ . Arrows and symbols are referred to different final states and oxidation state of Ce. Note that symbols with blue color such as  $u'''$ ,  $u''$ ,  $u$ ,  $v'''$ ,  $v''$ , and  $v$  are final states from  $\text{Ce}^{4+}$ ; symbols with green color such as  $u_o$  and  $v_o$  are final states from  $\text{Ce}^{3+}$ . Other two final states like  $l_o$  and  $t_o$  are not shown in this figure due to relatively low amplitude and little contribution to the three-lobed envelop structure. (b) O 1s XPS spectra of MBE and PVD samples. The binding energy value of peak of MBE sample is higher than PVD sample, which means MBE sample has more  $\text{Ce}^{3+}$  component than PVD sample.



sistance is determined when voltage is at 0.3V during dc read. Figure 2.4(a) shows the result that the device could endure at least  $10^5$  cycles, with fair accuracy of resistance for both HRS and LRS, and so far no report on  $\text{CeO}_x$  RRAM has tested reliability over  $10^4$  cycles. The device shows a feasible operating window and also good repeatability of switching. After  $10^5$  cycles, the resistance value at HRS suddenly drops at certain point, and it requires either a strong set or reset to recondition the filament; however, this process increases the electrical stress in the dielectric and gradually damages the device. Figure 2.4(b) shows retention capability for both resistance states of the device. The device was sampled at every 50s with read voltage at 0.3 V at both room temperature and 150 °C. Both HRS and LRS are stable till  $10^5$  s. These two tests indicate that the device can be potentially applied for future RAM device. To study Al/ $\text{CeO}_x$ /Au system for NVM devices in more details, the device is scaled horizontally and vertically. Figures 2.5(a) and 2.5(b) show that PVD  $\text{CeO}_x$  RRAMs have area-independent set and reset voltages as device scaled down from 250  $\mu$  to 500 nm. Figures 2.5(c) and 2.5(d) also show that MBE  $\text{CeO}_x$  RRAMs have the similar area-independent set and reset voltages. Although further scaling of  $\text{CeO}_x$  RRAM needs to be done for building arrays, Figures 2.5(a)2.5(d) show that Al/ $\text{CeO}_x$ /Au system has decent scalability in size. The set and reset voltage do not scale down while the size of device shrinks and resistance values at the HRS and LRS can retain at the same order while the device scales down. In addition, horizontal scaling results also indicate that the interfacial oxidation and reduction in

our Al/CeO<sub>x</sub>/Au system are excluded from possible mechanism since the set and reset voltage are independent to the area of device. For vertical scaling, Figures 2.5(a) and 2.5(b) show that the forming voltage increases when thickness of CeO<sub>x</sub> increases in both MBE and PVD films; this thickness dependent behavior has been observed in other filamentary path type RRAM devices. Chen[39] proposed a first-order model based on probability analysis showing that thickness scaling would change the forming voltage since the forming is a random process in filamentary path type RRAMs. On the other hand, set and reset voltage are independent to thickness of CeO<sub>x</sub> film because set and reset process are local formation and rupture of the filament in switching materials, it requires less energy to eliminate or generate oxygen vacancies compared to forming process.[18] Figures 2.5(c) and 2.5(d) show that Al/CeO<sub>x</sub>/Au system with both MBE and PVD films has thickness independent behaviors and this makes the system promising for building extremely scaled devices.

## 2.4 Conclusion

Resistive switching random access memory devices based on Al/CeO<sub>x</sub>/Au sandwiched structure are fabricated by molecular beam epitaxy and electron beam evaporation and demonstrate low set and reset voltage, large operation window (MBE grown CeO<sub>x</sub> RRAM is larger than 10<sup>6</sup>), stable scalability, and good reliability. Cerium oxide is as favorable as other high-k dielectric materials and should receive more attention for further research and advanced applications such as neuromorphic computing. Besides, based on experiments

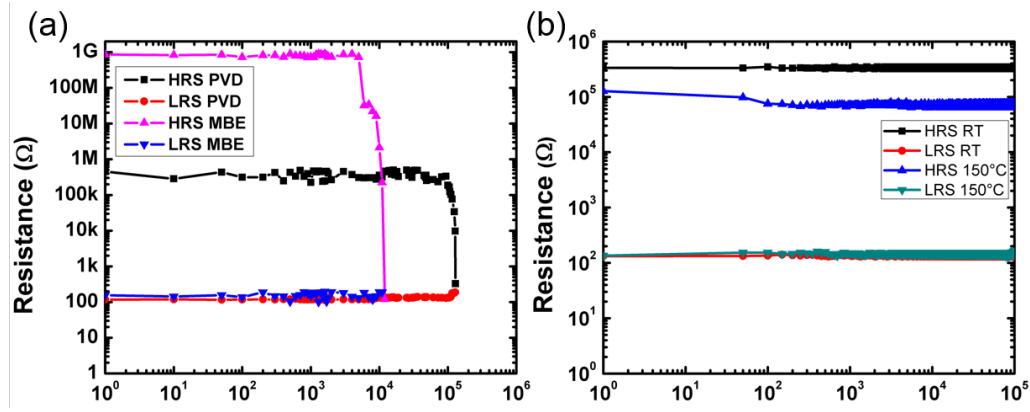


Figure 2.4: (a) HRS and LRS resistance along with switch cycles. The structure of tested device is 30 nm Al/13 nm PVD grown  $\text{CeO}_x$ /30 nm Au at room temperature. (b) HRS and LRS resistance under a continuous 300-mV read voltage. The structure of tested device is 30 nm Al/13nm PVD grown  $\text{CeO}_x$ /30 nm Au at room temperature and 150°C. The voltage sweep followed the numerical order in Fig. 2.2(a) for each cycle. Step 1: voltage ramps up to 2.5 V. Step 2: voltage ramps down to 0V, after this half cycle, voltage ramps up to 300mV to read and records the resistance. Step 3: voltage ramps up to -2.2 V. Step 4: voltage ramps down to 0V. Then, voltage ramps up to read and records the resistance.

performed in the chapter, conductive filament path formation and rupture seem more reasonable than interfacial oxidation/reduction.

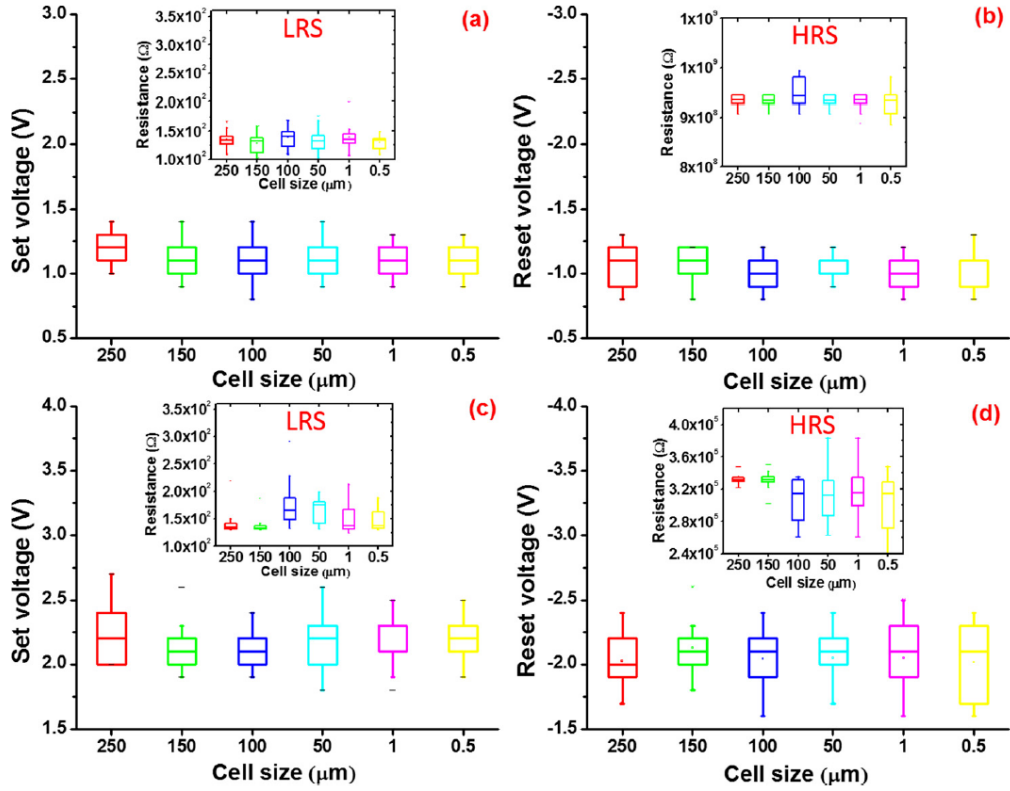


Figure 2.5: (a) and (b) Diameter of device versus set voltage and reset voltage of MBE grown  $\text{CeO}_x$  RRAMs. (c) and (d) Diameter versus set voltage and reset voltage of PVD grown  $\text{CeO}_x$  RRAMs. For each area value, 20 devices have been tested. Inset graphs in (a) and (b) represent the LRS and HRS resistance dependence of MBE devices and inset graphs in (c) and (d) represent the LRS and HRS resistance dependence of PVD devices.

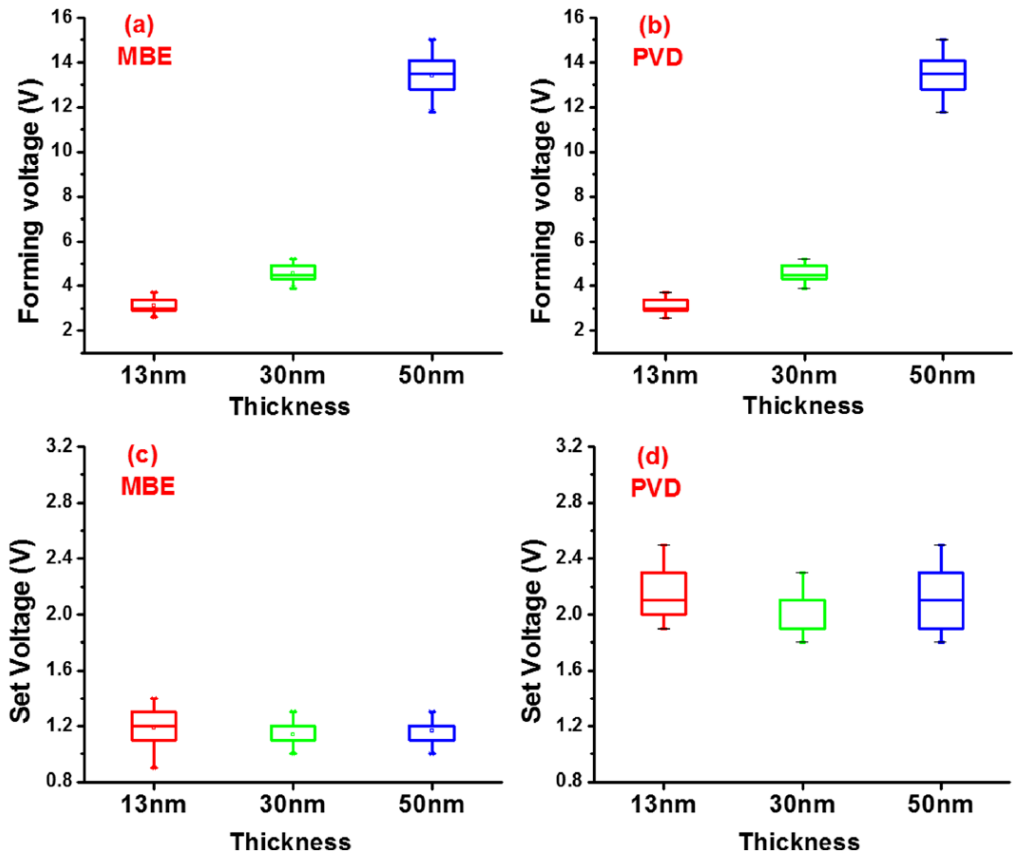


Figure 2.6: (a) and (b) Switching layer thickness dependence of forming voltage with MBE and PVD grown thin film. (c) and (d) Set voltage is independent of  $\text{CeO}_x$  film thickness for both MBE and PVD  $\text{CeO}_x$  film. For each thickness value, 15 PVD and MBE grown  $\text{CeO}_x$  RRAM devices have been tested.

## Chapter 3

# Bilayer Cerium Oxide RRAM

### 3.1 Introduction

Nanoscale metal oxide RRAM have potential in the development of brain-inspired computing systems that are scalable and efficient. In the field of neuromorphic computing, RRAM is also referred as memristor, the circuit element predicted by Chua [26] and found by Hewlett Packard Labs. In Chua's work [3], biological synapses can be treated as memristor because those synapses have dc current-voltage characteristics matching the definition of generic memristors. That is, dc current-voltage characteristics with hysteresis and it passes the origin. In neuromorphic computing systems, memristors represent the native electronic analogues of the biological synapses. This demonstration is an important step towards the physical construction of high density and high connectivity neural networks. A memristor is a two-terminal memory resistor electronic device, in which a metal oxide switching layer is sandwiched between two metal electrodes [40–43]. In general, memristors offer non-linear switching characteristics, and materials and process compatibility with advanced silicon manufacturing. These attributes have spurred the exploration of memristors as synaptic devices for realizing spike-based hardware learning systems that are capable of processing unstructured, temporal data

[2, 44–47]. However, for memristor-based technologies to be viable, the device should exhibit several key characteristics. It should have a compact nanoscale footprint, operate at a voltage close to 1V that is compatible with complementary metal oxide semiconductor (CMOS) technology, have reproducible electrical characteristics, and possess high switching speed to minimize the energy consumption [42]. Furthermore, the hardware integration of synaptic connections in advanced neural networks requires memristors with multiple resistive states [48, 49]. These are challenging requirements and are difficult to implement without significant innovations. The phenomenological principle of memristor device operation is based on the change in the physical properties of a conductive filament (associated with the presence of oxygen vacancies) by applying an electric field across the metal oxide switching layer [50, 51]. The resulting motion of the oxygen vacancies alters the device resistance between low (Set) and high (Reset) states, depending on the direction and the amplitude of the electric field. So far, a variety of structures from a large set of materials (various combinations of metal oxide switching layers and metal electrodes) have been studied in the literature [9]. Several key findings can be drawn from those studies regarding the performance, energy and scalability of this type of devices. The most important finding reveals the trade-off between the switching energy and the data retention time, that is often referred to as voltage-time dilemma[52]. This trade-off is associated with the energy barrier of the device structure.

For example, devices made of metal oxides with small energy bandgap

( $E_g$ ), such as titanium oxide ( $\text{TiO}_x$ ,  $E_g \sim 3.4\text{eV}$ ), generally exhibit low operating voltage and compromised data retention, while those with large bandgap, such as hafnium oxide ( $\text{HfO}_x$ ,  $E_g \sim 5.4\text{eV}$ ) demonstrate the opposite [53]. However, the fabrication of devices with bilayer switching stacks has shown to be effective in mitigating this trade-off. In particular, the improvement in data retention was obtained by the incorporation of an ultra-thin wide bandgap metal oxide capping layer (for example aluminum oxide) [54]. On the other hand, the addition of a reactive capping metal (for example titanium, hafnium, etc.) as an oxygen scavenging layer provided a pathway for reducing the operating voltage of the devices [39, 55]. Despite significant advances, a sub-1V memristive device that simultaneously affords built-in analog behavior, energy efficiency on par with a biological synapse, forming-free operation and low device-to-device variations is still elusive.

The cerium oxide RRAM device shows decent operating voltage, fast speed, reliable switching and high operating window; however, it doesn't demonstrate prominent analog behaviors when it resets. In order to improve analog switching and realize multiple states, an ultra-thin, non-stoichiometric  $\text{HfO}_x$  is added into device stack. In this chapter cerium oxide based bilayer memristors that is forming-free, low-voltage ( $|0.8\text{V}|$ ), energy-efficient (full On/Off switching at  $2\text{pJ}$ , intermediate states switching at  $\text{fJ}$ ), and reliable are shown. Furthermore, pulse measurements reveal the analog nature of the memristive device; that is it can be directly programmed to intermediate resistance states. Leveraging this finding, the device stack demonstrate spike-



timing-dependent plasticity (STDP), a spike-based Hebbian learning rule. In those experiments, the memristor exhibits a marked change in the normalized synaptic strength ( $>30$  times) when the pre- and post-synaptic neural spikes overlap.

### **3.2 Hafnium Oxide and Cerium Oxide Stacked Resistive Random Access Memory**

With memristive-based synaptic device by engineering the material properties of an  $\text{HfO}_x$  capping layer in a bilayer structure with a cerium oxide ( $\text{CeO}_x$ ) switching layer, the combination of sub-stoichiometric structural properties of the  $\text{HfO}_x$  capping layer and its enhanced thermal resistivity at nanoscale dimensions leads to the significant improvement in switching behavior of the devices in terms of the operating voltages, device performance uniformity, reproducibility, and reliability . Furthermore, this structure yields forming-free devices with an analog resistance state that is inherent to the device itself. This key attribute of our  $\text{HfO}_x/\text{CeO}_x$  devices enables the implementation of Hebbian learning [56], validating the plasticity of the synaptic connection. Memristor devices ( $150\mu\text{m}$  diameter) were fabricated on silicon substrates capped with 300nm silicon dioxide. The device structure consists of gold bottom electrode,  $\text{HfO}_x/\text{CeO}_x$  switching layer, and aluminum top electrode. The total thickness of the bilayer switching layer in all experiments was kept at 20nm, while varying the thickness of individual the  $\text{HfO}_x$  and  $\text{CeO}_x$  layers. The metal electrodes were deposited using electron-beam evaporation.

Figure 3.1(a) conceptually illustrates the effect of the engineered  $\text{HfO}_x$  capping layer on the concentration of oxygen vacancies in the  $\text{CeO}_x$  switching layer. X-ray photoelectron spectroscopy (XPS) was performed to guide the development of the bilayer structure. Figure 3.1(b) shows the Hf 4f spectrum of the engineered  $\text{HfO}_x$  capping layer, revealing the sub-stoichiometric nature of the film. The data indicates the presence of Hf 4f<sub>7/2</sub> and Hf 4f<sub>5/2</sub> peaks at 16.32eV and 18.03eV, respectively, which is consistent with the previous reports in the literature [57, 58]. Metallic Hf was also found in the engineered  $\text{HfO}_x$  layer, evident from the peak at 15.02 eV. The chemical composition of the  $\text{HfO}_x$  was quantified using the CasaXPS software, in which  $x$  was found to be about 1.75. Figure 3.1(c) shows the Ce 3d XPS spectra of the  $\text{CeO}_x$  switching layer with and without the engineered  $\text{HfO}_x$  capping layer. In these experiments, the  $\text{CeO}_x$  and  $\text{HfO}_x$  layers were 20nm and 0.8nm, respectively. The thickness of  $\text{HfO}_x$  was chosen so that it allowed X-ray beam to penetrate into underneath  $\text{CeO}_x$  layer and received decent signal strength. As a result, the adequately small thickness of the  $\text{HfO}_x$  capping layer allowed the XPS analyzer to receive signal from the  $\text{CeO}_x$  layer. As can be seen in Figure 3.1(c), the bilayer structure exhibits discernable u' and v' peaks at 904 eV and 885 eV [35, 36, 59] that are absent in the spectrum of the  $\text{CeO}_x$  layer with no  $\text{HfO}_x$  capping layer. The u' and v' peaks signals the reduction of the  $\text{Ce}^{4+}$  to  $\text{Ce}^{3+}$  states, which can be translated to the formation of excess oxygen vacancies at regions near the  $\text{HfO}_x/\text{CeO}_x$  interface. The marked increase of the oxygen vacancy concentration in the bilayer structure permits the forma-

tion of the conductive filament using a smaller electric field, thereby enabling the low-voltage operation of the bilayer structure. Figure 3.1(d) shows the representative dc current-voltage characteristics of two  $\text{CeO}_x$ -based devices with and without the engineered  $\text{HfO}_x$  capping layer, demonstrating significant reduction of the Set voltage to below 0.8V. In a memristive device, the transition from low to high resistance states occurs as the polarity of the electric field across the device is reversed. As the reverse electric field increases, the oxygen anions in the conductive filament begin to disperse through drift and diffusion processes [47]. Considering the similar thickness of the switching layer in Figure 3.1(d), the improved Reset voltage of the bilayer device may be explained by locally enhanced diffusion of oxygen vacancies. We infer that the enhanced thermal resistivity of  $\text{HfO}_x$  at nanoscale dimensions amplifies Joule heating in the  $\text{CeO}_x$  switching layer, thereby accelerating the dispersion of oxygen anions at a lower electric field. To elucidate this concept, we performed numerical heat transfer analysis using COMSOL simulator for two devices in Figure 3.1(e) at the bias of -0.6V. The simulation results indicate significant enhancement of Joule heating in the bilayer structure. For these simulations, we used the measured electrical parameters of the layers, while the thermal parameters were obtained from the literature[60–63]. The fabrication process is as follows:  $\text{CeO}_x$  layer was reactively evaporated in oxygen plasma ambient at 0.2mTorr and an average deposition rate of 0.06nm/s. The  $\text{HfO}_x$  layer was formed by plasma-assisted atomic layer deposition (PE-ALD) using water and tetrakis (dimethylamido) hafnium ( $\text{Hf}(\text{NMe}_2)_4$ ) precursors.

The film optimization involved varying a wide range of deposition conditions. The optimal  $\text{HfO}_x$  capping layer was deposited at 200°C. The pulse width of the hafnium precursor was 0.25s and the hold time between each pulse was 5s. The optimal oxygen plasma power was found to be 300W. The devices were isolated using a wet etching process by first patterning the  $\text{HfO}_x$  film in buffered oxide etch followed by removing the  $\text{CeO}_x$  layer in a mixture of hydrochloric acid, potassium hexacyanoferrate, and de-ionized water. Devices were measured under vacuum in a Lakeshore CRX-VF Probe Station using Agilent semiconductor parameter analyzer B1500 equipped with B1525 Semiconductor Pulse Generator Unit (SPGU). Care was taken to minimize the impact of parasitic elements, for example, the capacitance.

Low device variability is critical for the implementation of large neural networks with high density of memristive synaptic connections. Therefore, we statistically examined the effect of the  $\text{HfO}_x$  thickness on the important device parameters: Set, Reset, and forming voltages. In these experiments, the  $\text{HfO}_x$  thickness was varied, while keeping the total thickness of the bilayer stack fixed at 20nm. The thickness ratio defined here is the  $\text{HfO}_x$  thickness to the total thickness in bilayer. The data in Figure 3.2 indicates that the insertion of an  $\text{HfO}_x$  capping layer with the optimal thickness ratio of about 0.1 significantly improves the uniformity of the key device parameters. Interestingly, this optimal thickness ratio also coincides with the minimum operating voltages of the bilayer structure. We surmise that the  $\text{HfO}_x$  film begins to act as an independent switching layer beyond this optimal thickness ratio, re-

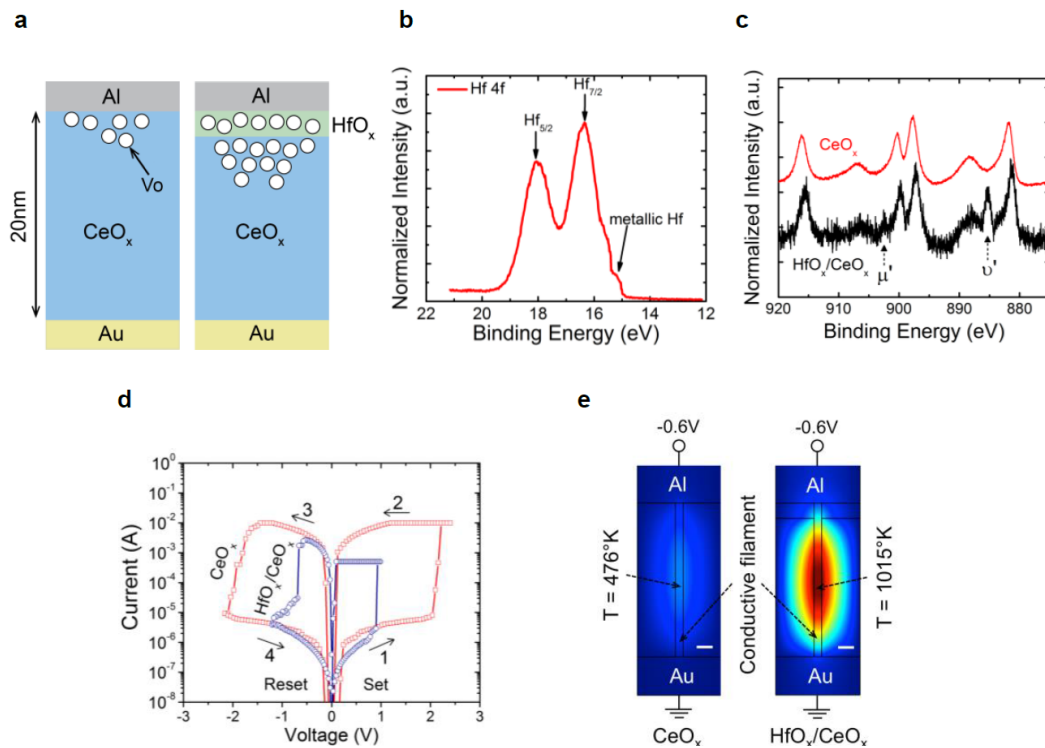


Figure 3.1: Improving memristor device characteristics using an engineered sub-stoichiometric HfO<sub>x</sub> capping layer. (a) Schematic structure of two memristors with and without the engineered HfO<sub>x</sub>, conceptually illustrating the increase of the oxygen vacancy density in the CeO<sub>x</sub> switching layer. This attribute of the bilayer memristor results in the forming-free operation and the reduction of the Set voltage. XPS spectra of the (b) engineered HfO<sub>x</sub>, and (c) CeO<sub>x</sub> films with and without the HfO<sub>x</sub> capping layer. The XPS studies indicate the increase of the oxygen vacancy concentration in the CeO<sub>x</sub> film capped with the oxygen-deficient HfO<sub>x</sub> layer. (d) Representative current-voltage characteristics of two memristors, indicating the sub-1V operation of the bilayer memristive device. (e) Heat transfer simulations illustrate enhanced Joule heating in the bilayer structure, causing the marked reduction of the Reset voltage (scale bars are 2nm). The observed increase in Joule heating arises from the high thermal resistivity of HfO<sub>x</sub> at nanoscale. The thickness of HfO<sub>x</sub> is 2nm and the total thickness is 20nm.

sulting in significant increase in both the device operating voltages and the device variability. Moreover, the Reset voltage begins to increase as the  $\text{HfO}_x$  film becomes thicker. This observation is in agreement with our heat transfer simulation results in Figure 3.3. In Figure 3.2(c), the Reset voltage at 0.4 was too large compared to other ratios so it wasn't included in Figure 3.2(c).

A fresh memristive device generally requires an initial electroforming step; that is the formation of a conductive filament using a relatively large electric potential (known as the forming voltage) before the device can operate at normal Set and Reset voltages. The disparity between the Set and the forming voltages necessitates that the devices in a crossbar array are isolated and accessed individually for electroforming [64, 65] in order to avoid the breakdown of the neighboring formed devices in the array. However, the physical constraints of these strategies limit the implementation of high-density crossbar arrays. Our bilayer  $\text{HfO}_x/\text{CeO}_x$  device is free from such a limitation, exhibiting forming-free behavior; *i.e.* the Set voltage is adequate to form the conductive filament in a fresh memristive device (See Figure 3.2(a) and (b)). This characteristic is attributed to the efficacy of the  $\text{HfO}_x$  capping layer in creating sufficiently high concentration of excess oxygen vacancies in the  $\text{CeO}_x$  switching layer. Figure 3.3 shows the heat transfer simulations by COMSOL for devices with varying  $\text{HfO}_x/\text{CeO}_x$  thickness ratio, with the total thickness of the  $\text{HfO}_x/\text{CeO}_x$  stack was 20 nm. The peak temperature value was found to be the highest when the thickness ratio was about 0.1. The simulation results suggest that capping the  $\text{CeO}_x$  switching layer with a sufficiently thin layer

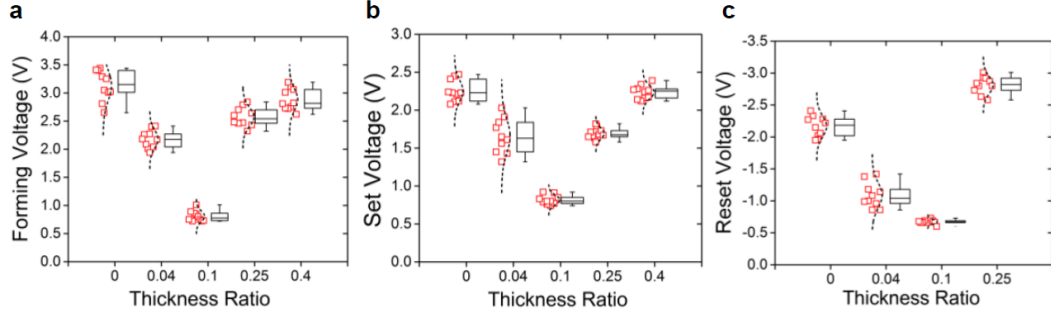


Figure 3.2: Effect of  $\text{HfO}_x$  thickness ratio on the memristor device behavior. The data indicates that the optimal device characteristics ((a) forming voltage, (b) Set voltage, and (c) Reset voltage) occurs at the thickness ratio of about 0.1. Moreover, the device-to-device variation is reduced at this optimal thickness ratio. The equivalency of the forming and Set voltages at the optimal thickness ratio confirms the forming-free operation of the device. Low device variability is critical for implementation of large neural networks with high density of memristive synaptic connections. Therefore, we statistically examined the effect of the  $\text{HfO}_x$  thickness on the important device parameters: Set, Reset, and forming voltages. In these experiments, the  $\text{HfO}_x$  thickness was varied, while keeping the total thickness of the bilayer stack fixed at 20nm. The thickness ratio defined in this work is  $\text{HfO}_x$  thickness to total thickness in bilayer. The data in Figure 3.2 indicates that the insertion of an  $\text{HfO}_x$  capping layer with the optimal thickness ratio of about 0.1 significantly improves the uniformity of the key device parameters. Interestingly, this optimal thickness ratio also coincides with the minimum operating voltages of the bilayer structure. We surmise that the  $\text{HfO}_x$  film begins to act as an independent switching layer beyond this optimal thickness ratio, resulting in significant increase in both the device operating voltages and the device variability. Moreover, the Reset voltage begins to increase as the  $\text{HfO}_x$  film becomes thicker. This observation is in agreement with our heat transfer simulation results in Figure 3.3. In (c), the Reset voltage at 0.4 was too large compared to other ratios so it wasnt included.

of  $\text{HfO}_x$  enhances the Joule heating, owing to the pronounced thermal resistivity of  $\text{HfO}_x$  at nanoscale. However, as the thickness of the  $\text{HfO}_x$  increases,

the Joule heating begins to diminish, which is consistent with the thickness dependence of the  $\text{HfO}_x$  thermal conductivity. The enhanced Joule heating effect in the optimal structure is therefore expected to enhance the diffusion of the oxygen vacancies during the Reset process, thereby reducing the Reset voltage.

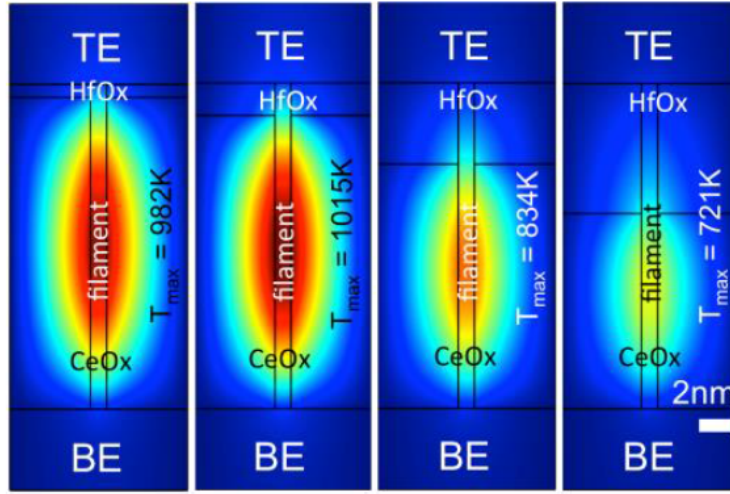


Figure 3.3: Effect of  $\text{HfO}_x$  film thickness on Joule heating. Numerical heat transfer simulation results for several bilayer  $\text{HfO}_x/\text{CeO}_x$  structures with varying  $\text{HfO}_x$  to total thickness ratio at the bias voltage of  $-0.6\text{V}$ . The total thickness of the  $\text{HfO}_x/\text{CeO}_x$  stack was kept at  $20\text{nm}$ . The Joule heating begins to diminish as the thickness of the  $\text{HfO}_x$  was increased, which arises from the thickness dependence of the  $\text{HfO}_x$  thermal conductivity.

The bilayer structure exhibits excellent switching reliability at the thickness ratio of 0.1, which conceivably stems from the reduced operating voltage of the device. In Figure 3.4(a), the optimal memristor bilayer structure survives more than  $2 \times 10^5$  cycles of programming (endurance test). The device with closed-loop programming algorithm can survive over  $5 \times 10^7$  cycles. The



adaptive programming works as below: the system set up targeted Set current and Reset current, in our case, we set  $10\mu\text{A}$  as Set threshold and  $1\mu\text{A}$  as Reset threshold so the operating window we target is 10. The adaptive programming algorithm tries to switch device starting from 0V then read once to check if the read current is above or below the threshold current, if yes, then the device is considered to have switched, if not, the algorithm increases the write voltage by steps of 0.1V until the device switches. The accelerated retention test in Figure 3.4(b) indicates projected data retention of 10 years for the bilayer devices. The cumulative distribution function (CDF) in Figure 3.4(c) illustrates the representative switching characteristics between different programming cycles for two  $\text{CeO}_x$  devices with and without the engineered  $\text{HfO}_x$  capping layer and also bilayer device with adaptive programming, also referred as close-loop testing. The CDF plot indicates the on-off ratio of bilayer device is larger than single  $\text{CeO}_x$  device. The Off-state characteristic of the device appears to have been degraded, perhaps due to the non-uniformity of the Joule heating effect. The bilayer device exhibits average low- and high-resistance states (LRS and HRS) of about  $600\Omega$  and  $2.8\text{M}\Omega$  that are larger than those of the device with no  $\text{HfO}_x$  capping layer by factors of 4 and 10, respectively. The resulting increase in the LRS and HRS values is beneficial for reducing the switching power consumption of the device during the Set and Reset operations.

The series connection of one transistor and one memristor (1T-1R) is a popular approach for implementing multi-state memory function. The use of such configurations, however, limits the memristor integration density because

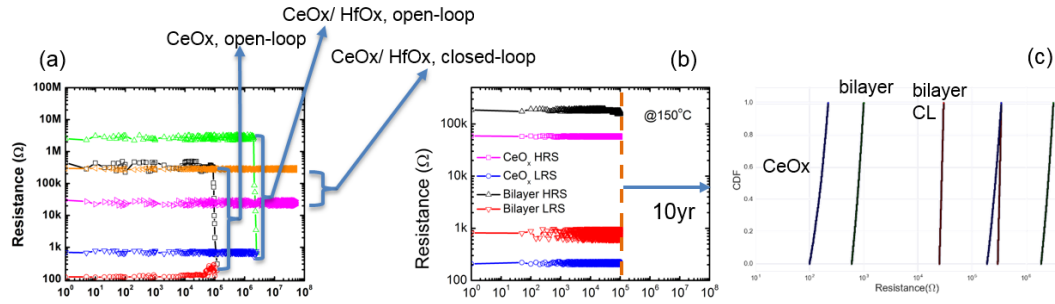


Figure 3.4: Device reliability studies. (a) The endurance test results for the  $\text{CeO}_x$  and the optimal  $\text{HfO}_x/\text{CeO}_x$  devices. In addition to the improved endurance properties, the bilayer device exhibits larger HRS and LRS values compared to the device with no  $\text{HfO}_x$ . The increase of the LRS and HRS values is favorable for reducing the switching power consumption of the bilayer device. Besides, when bilayer device were tested in close loop with adaptive programming and relaxed on-off ratio, it survives over  $5 \times 10^7$  cycles without any degradation. (b) The accelerated retention test for the  $\text{CeO}_x$  and the  $\text{HfO}_x/\text{CeO}_x$  devices measured at 150C at constant stress voltage of +0.2V. The results indicate projected data retention of 10 years for both devices. (c) Representative CDF plot of the cycle-to-cycle programming characteristics for two devices with and without the engineered  $\text{HfO}_x$  layer.

of the physical constraints imposed by the transistor dimensions as well as the need for a complicated driver circuit in order to independently control each transistor. To circumvent these practical issues, the multi-state characteristic must be inherent to the two-terminal memristive device itself. Figures 3.5(a) and (b) illustrate the pulse measurement results for a bilayer device (with the optimal structure), indicating the gradual change in the conductance of the filament between the fully On and Off states. The observed resistive states are inherent to the device because no current compliance limit was used during these measurements. Interestingly, the bilayer device also exhibits weak voltage-time dependence for pulses shorter than a few microseconds, which

could be attributed to the dominant effect of the  $\text{HfO}_x$  capping layer on the device switching behavior. Using the corresponding transient voltage and current waveforms in Figure 3.5(c) and (d), the full On/Off energy consumption during Set and Reset steps was calculated to be 2.6 and 2.1 pJ, respectively. Considering the analog characteristic of the resistive states together with the large HRS to LRS ratio in excess of  $10^3$ , the energy consumption for switching between the intermediate resistance states will be much smaller (about tens of fJ, assuming memory states with an increment of  $100\Omega$ , please see Figure 3.6).

Inspired by the brain, spike-based hardware learning systems have potential to be efficient and compact for processing unstructured data [66]. In such systems, the learning mechanism follows the spike-based form of Hebbian learning, *i.e.* STDP, in which the change in the strength of the synapse depends on the time difference ( $t$ ) between the pre- and post-synaptic neural spikes. Figure 3.7(a) illustrates the synaptic waveforms. The waveforms with exponential decays were emulated with a series of square pulses. For these experiments, we have chosen an average spike rate of about 1MHz, which is  $10^5$  times faster than that of the brain. This corresponds to a time step of  $1\mu\text{s}$  for updating the internal state of neurons and calculating the synaptic currents, assuming the neuron spiking probability of 0.01 as in the brain. Note that the acceleration of the learning rate is beneficial for handling large amount of data, while allowing the reduction of the energy consumption of the memristors. Figure 3.7(b) shows the plot of the normalized conductance change

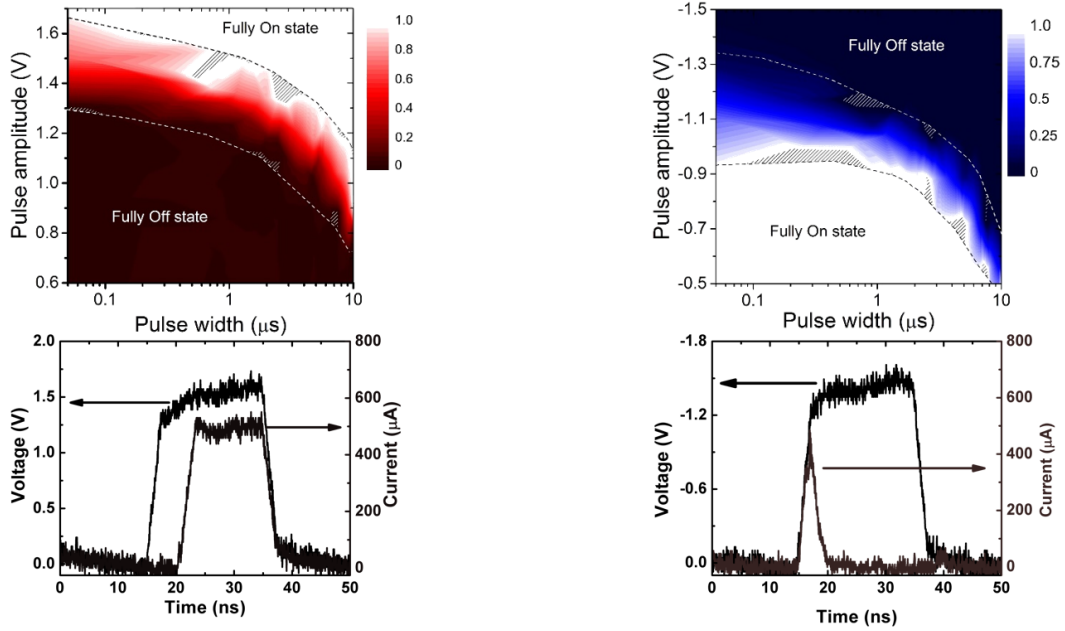


Figure 3.5: Analog memory characteristic of the bilayer memristor. The normalized conductance of (a) bilayer memristor is plotted as a function of pulse widths and amplitudes when the device switches from a, fully Off state to fully On state, and (b), fully On state to fully Off state. The dashed lines are guide to the eye and the hatched regions denote unmeasured points. The data in (a), and (b) reveal the gradual change in the conductance of the device between the fully Off and On states. Full On/Off switching energy consumptions of 2.6 and 2.1pJ were calculated from the transient (c), Set and (d), Reset voltage and current waveforms, respectively.

of the optimal bilayer device as a function of the time difference between the pre- and post-synaptic neural spikes. The data is fitted with exponential decay functions, confirming an STDP behavior similar to that of a biological synapse. Moreover, the data indicates a remarkable change in the normalized conductance of the device ( $>30$  times) when the pre- and post-synaptic spikes overlap.

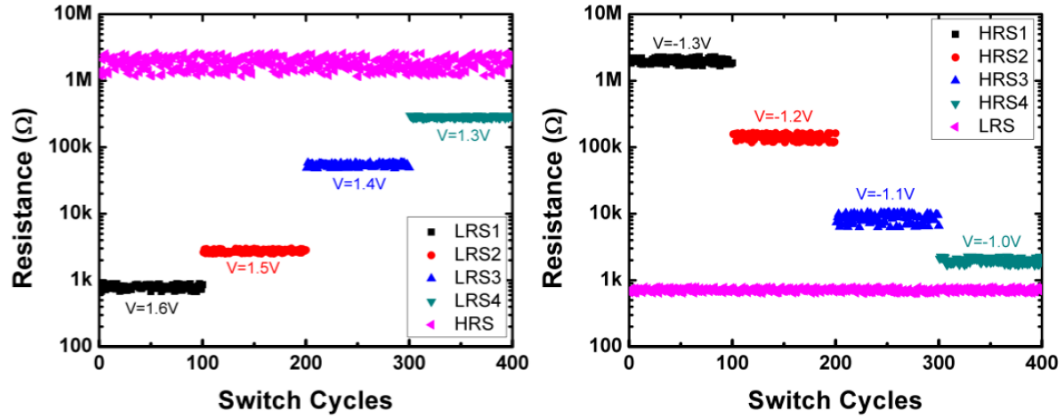


Figure 3.6: Switching cycles of intermediate states. Cycle test is conducted in both positive polarity and negative polarity to examine if intermediate states are separable and stable. The pulse voltage increased after every 100 cycles in (a) and (b). The resistance variation in each resistance state is very small, which means it is very stable. The separation of each resistance state is not linear with applied voltage. But it has potential to generate more resistance states if more voltage interval are added between each applied voltage properly. This also shows that the switching energy between each intermediate state is at femto joule scale

### 3.3 Conclusion

In the chapter summary, a new bilayer  $\text{HfO}_x/\text{CeO}_x$  memristors is demonstrated by tailoring the structural properties of the nanoscale  $\text{HfO}_x$  capping layer. The memristive device was readily implemented using CMOS-compatible materials and processes. The device is forming-free and thus amenable to high-density integration. More importantly, this device also exhibits analog resistance states, sub-1V operating voltages, high conductance change at fast nanosecond pulses, and energy efficient operation. Furthermore, the STDP learning rule was successfully implemented, following the Hebb's rule of learn-

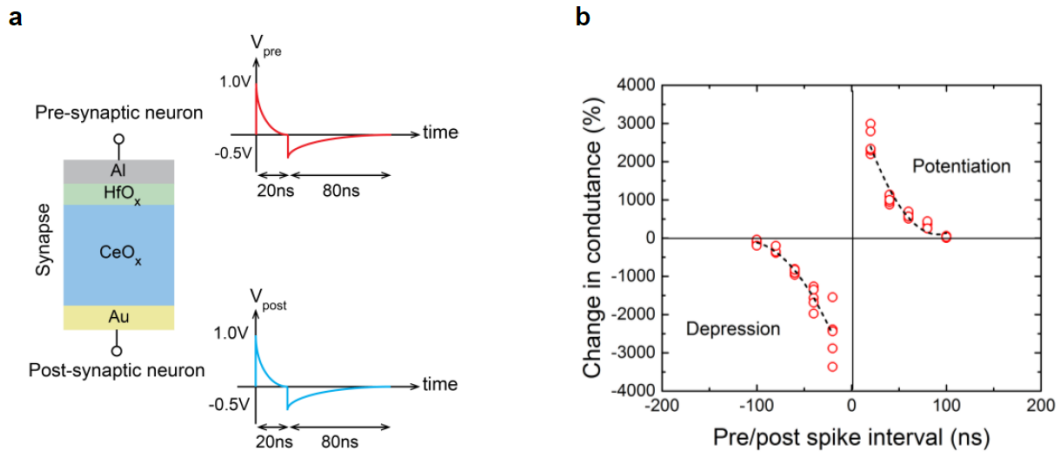


Figure 3.7: Implementation of STDP learning using the HfO<sub>x</sub>/CeO<sub>x</sub> memristive device. (a) Schematic representation of the learning experiment. Two waveforms with identical shapes were applied to the top and bottom electrodes. In the learning experiments, the time intervals between the pre- and post-synaptic spikes were varied in order to probe the synaptic depression ( $\delta t < 0$ ) and potentiation ( $\delta t > 0$ ). The positive (negative) time difference indicates that the pre-synaptic spike occurs before (after) the post-synaptic one. (b) The plot clearly indicates the marked change in the synaptic strength as a function of different pre/post spike intervals.

ing; that is, neurons fire together wire together. The salient features of this new memristor meet the main requirements for a native synaptic device and can be used for hardware implementation of STDP-based learning systems.

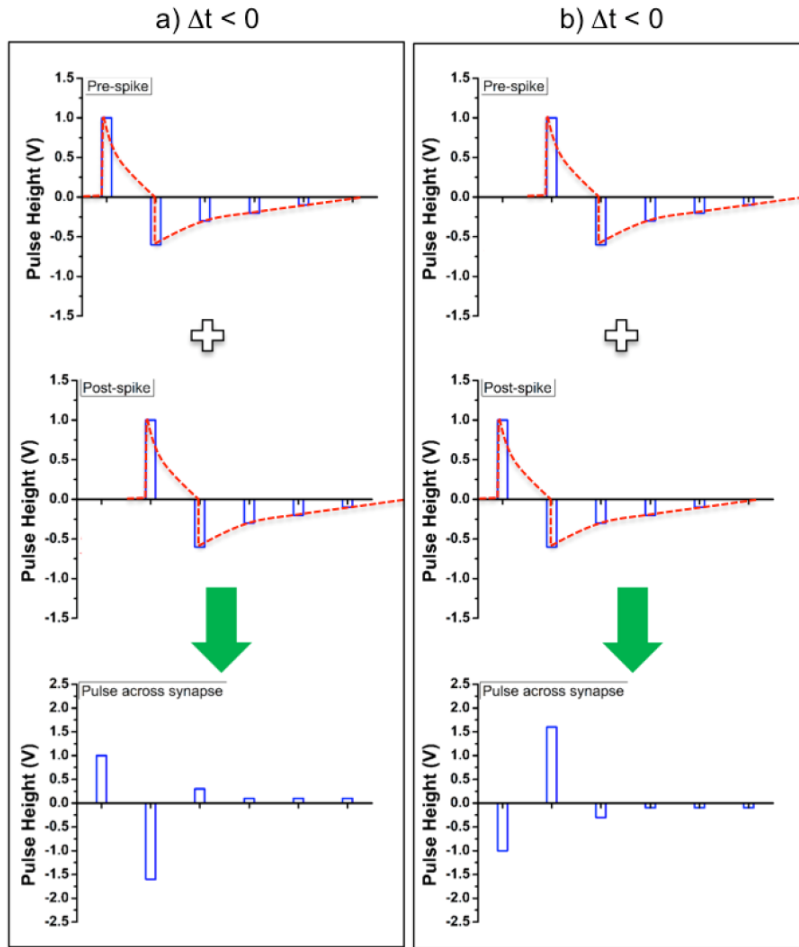


Figure 3.8: Generation of waveforms for Spike-Timing-Dependent Plasticity (STDP) measurements. Multiple square waves with various pulse heights were tailored for emulating exponential decay pulses used in STDP studies. The time interval of the pulses on the top and bottom electrodes (that are pre- and post-neural spikes) was varied relative to each other while monitoring the conductance of the device. The pulses were created using the Keysight B1500 pulse generator unit. Depending on the time difference between the pulses illustrated in Figure 3.8(a) and (b), the device demonstrates long-term depression ( $\delta t < 0$ ) and potentiation ( $\delta t > 0$ )

## Chapter 4

# Selector Device for RRAM

### 4.1 Introduction

One of the serious issues of RRAM technology is leakage current. With continuously scaled nanodevices and electrodes, challenges associated with the sneak path leakage and the interconnect series resistance in such a purely passive crossbar architecture become major concerns as they diminish the read margin and limit the maximum array size. To mitigate these problems, several solutions have been proposed to introduce nonlinear IV characteristics to the memory cells.[67, 68] For unipolar resistive devices, an extra diode can be connected in series to inhibit reverse conduction through unselected cells. For the more preferable bipolar resistive devices, bipolar nonlinear selector devices are needed. The simplest way to address the issue is to add a transistor along with memory cell. Adjusting the transistor's gate, source and drain bias can provide leakage current control and current compliance; however, this is not favorable due to the larger area ( $6F^2$ ), making RRAM unpreferable to one-transistor-one-diode for DRAM applications ( $6F^2$ ) and NAND Flash ( $4F^2$  to  $6F^2$ ) for NVM applications. Researchers have been active to find simple bipolar selector devices which can provide  $4F^2$  area. To date, selectors based on different mechanisms, such as Schottky barriers, tunnel barriers, metalin-



insulator transition, Zener diodes, mixed ionic electronic conduction (MIEC), and punchthrough diodes, have been investigated. Recently, the feasibility of complementary resistive switch and electromechanical diodes has also been explored. Despite the reported progress on selector performance, a series of bottlenecks on large current density, high ON/OFF ratios, fast switching speed, and process reliability remain to be resolved.[69–74]

## 4.2 S-type NDR Niobium Oxide

Devices exhibiting negative differential resistance(NDR) can be classified into two categories: current-controlled [75] (CC-NDR), or S-type, and voltage-controlled (VC-NDR), or N-type. Circuit elements exhibiting N-type NDR are available in the form of Esaki diodes, Gunn diodes and resonant tunnel diodes (RTD). On the other hand, although S-type NDR has been observed in structures exhibiting interband tunneling[76, 77], threshold switching[78], electronic instabilities[79], insulator-metal transitions in metal oxides, and as a precursor to memristive on-switching, S-type discrete or integrated circuit components are not readily available. The advent of easily fabricated S-type NDR devices would be of great commercial interest both as a circuit element in existing technologies and as an enabler of emerging technologies. A prime example of the latter is resistance-based memory technologies that utilize memristive, phase-change, conductive bridge, or spin-torque memory elements in crossbar array architectures. These technologies are under intense development due to their potential for providing fast, low-power, nonvolatile

random-access memory (NVM). Such memory would revolutionize computer architectures by facilitating the consolidation of memory and storage, ultimately replacing hard drives, Flash, and conventional DRAM in both memory and storage roles. One of the prime impediments to utilization of these emerging NV-RAM technologies is that they store information in the form of resistances that depend only weakly on voltage.

Consequently, when used in crossbar arrays, these memory elements must be paired with a highly non-linear two-terminal circuit element that passes current when the full voltage is applied across an addressed memory cell but sharply limits the current leaking through partially biased memory cells in the array. Without such a selector, reading and writing individual linear memory elements in a large array is not possible. The extremely non-linear current-voltage characteristics of NDR devices are ideal for this role. Note that conventional diode technologies do not meet all the requirements of a selector, which include small size, low temperature CMOS compatible back-end-of-line (BEOL) manufacturability, high current density operation, and, in most cases, bipolar operation. We describe here an easily manufactured, bipolar, room temperature S-type NDR circuit element that fulfills the needs of a crossbar memory selector. These devices rely on the fact that any electrical conduction mechanism whose conductivity depends strongly enough on temperature can, in principle, exhibit NDR due to Joule self-heating at sufficiently large biases and currents. In practice, NDR is only observed for a limited set of conduction mechanisms where the onset of NDR occurs at temperatures and

fields low enough for the devices materials to survive. This chapter is focused on an instantiation of a Joule-heating based S-type NDR selector based on niobium oxide ( $\text{NbO}_x$ ).

Figure 4.1 shows a cross-sectional transmission electron microscope (TEM) image of one of these selectors. To fabricate these devices, planarized substrates were prepared that included TiN nanovias through a dielectric bilayer of  $\text{SiO}_2$  and  $\text{Si}_3\text{N}_4$ . These nanovias range from 32 nm to 2  $\mu\text{m}$  across and are connected at the bottom to a common tungsten electrode. Blanket films of  $\text{NbO}_x$ , TiN, Pt, and Cr were deposited on top of these substrates after removing the native oxide from the exposed surface of the TiN nanovias. The  $\text{NbO}_x$  was deposited by reactively sputtering Nb in different partial pressures of oxygen to create samples with values of  $x$  near either 2 or 2.5 as determined from XPS measurements. TEM-based electron diffraction shows that the  $\text{NbO}_x$  films were amorphous as deposited. The Cr was included as a hard etch mask for photolithographically patterning top Pt contacts above each nanovia, which enabled individual testing of the resulting isolated selectors.

After an initial electrical forming process, stable NDR is observed in the devices with starting Nb to O ratios near 2:5 (a- $\text{Nb}_2\text{O}_5$ ), but not in those that start with a 1:2 ratio (a- $\text{NbO}_2$ ). The initially oxygen rich samples are formed by applying slow (1s) logarithmic current ramps with successively greater amplitude using a Keysight B1500 Parameter Analyzer. Depending upon the a- $\text{Nb}_2\text{O}_5$  layer thickness, the forming proceeds in one of two ways. For thinner devices, the conductivity increases and a region of NDR appears in their

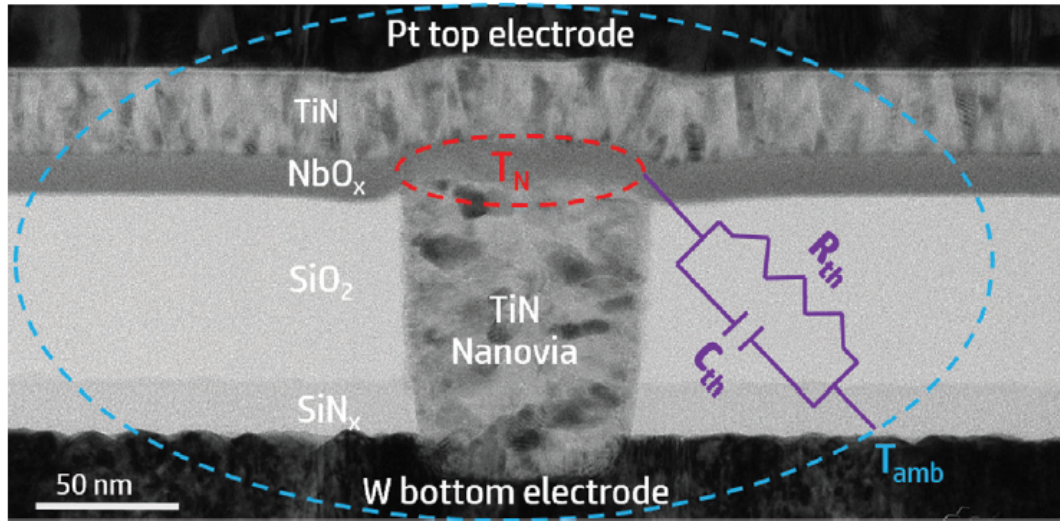


Figure 4.1: Bright field cross-sectional TEM image of a representative  $\text{NbO}_x$  selector. Active area of  $\text{NbO}_x$  is assumed to be at a uniform temperature  $T_N$  that is higher than the surrounding ambient temperature,  $T_{amb}$ , due to Joule heating. This heated region is thermally connected to  $T_{amb}$  through the effective thermal resistance,  $R_{th}$ , and thermal capacitance,  $C_{th}$ , of the surrounding device structures.

voltage - current characteristics (V-I) at higher currents (type I forming). The conductivity of thicker devices also increases initially but then abruptly decreases, again resulting in NDR at higher currents (type II forming, see Figure 4.2 for representative progression of V-I curves for both forming types). In both cases, the process is stopped when increasing the amplitude of the current sweep no longer changes the V-I curves. The sweep currents required to reach a stable state and the resulting low bias conductance both scale with the area of the bottom contact for both types of forming, suggesting that the entire region above the contact is formed. TEM-based electron energy-loss spectroscopy composition maps and electron diffraction measurements on

cross-sections of the formed selectors reveal that the  $a\text{-Nb}_2\text{O}_5$  is reduced by reaction with the TiN electrodes during both type I and type II forming processes. For type II devices, this reduction is followed by crystallization into the tetragonal  $\text{NbO}_2$  structure, which impedes further reduction. Films with a starting composition near  $a\text{-NbO}_2$  on the other hand are rapidly reduced, before crystallization can occur, to a composition that is too close to metallic to exhibit NDR. Empirically, selectors with the thickest tested  $a\text{-Nb}_2\text{O}_5$  layers (42 nm) usually underwent type II forming, although a metastable type I state could sometimes be achieved through careful control of the maximum applied current. Stabilizing a partially reduced type I state became easier for thinner films, with the thinnest tested layers (8 nm) always exhibiting only type I forming.

The NDR in these devices actually results from runaway Joule self-heating governed by a bulk electrical conduction mechanism in the  $\text{NbO}_x$  that is well-described by a modified three-dimensional Poole-Frenkel (3DmodPF) expression. This has important implications for improving the performance of selectors based on this principle. It provides, for example, guidance on how to lower their leakage current and tune their threshold voltage. It also provides insights into the dynamical thermal and electrical interactions between these selectors and their adjacent memory elements, which strongly impact the writing and reading processes. The standard expression for Poole-Frenkel conduction assumes that the carriers hop in just one dimension. Hartke[80] developed a more realistic three-dimensional treatment which Young[81] mod-

## Electroforming Process

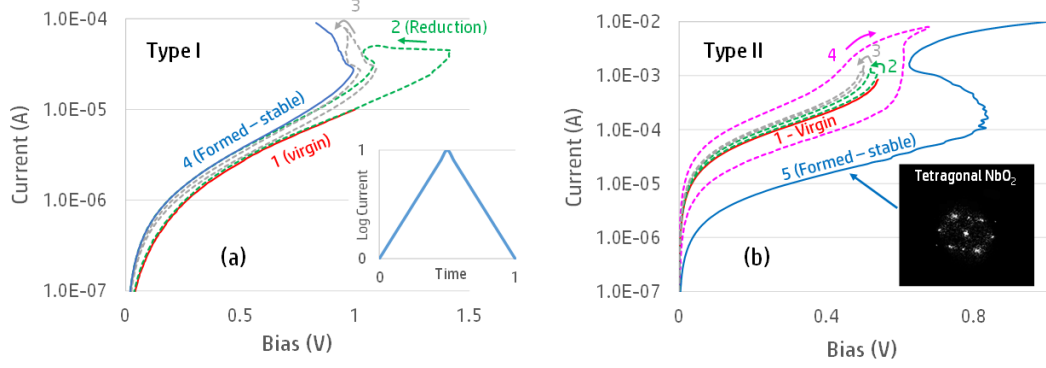


Figure 4.2: I-V curves of two different electroforming processes. Numbers indicate order of sweeps; arrows indicate time evolution. (a) Type I forming. This results in increasing currents as the initially amorphous  $\text{Nb}_2\text{O}_5$  is reduced through interaction with the TiN electrodes. (b) Type II forming. This includes crystallization to a more resistive tetragonal  $\text{NbO}_2$  state after the initial reduction. The slope of curve 5 in (b) is positive at high currents due to a  $\sim 100 \Omega$  resistance in series with the selector.

ified to include the effects of traps and donors in a fashion first applied to the one dimensional expression by earlier groups.[82, 83] The following is for the current density in  $\text{NbO}_x$ :

$$j(F, T) = \sigma F = \sigma_0(T) \left( \frac{k_B}{\beta} \right)^2 \left\{ 1 + \left( \frac{\beta \sqrt{F}}{a k_B T} - 1 \right) e^{\frac{\beta \sqrt{F}}{a k_B T}} \right\} + \frac{\sigma_0(T) F}{2} \quad (4.1)$$

where

$$\sigma_0(T) = e \mu N_c \left( \frac{N_d}{N_t} \right)^2 e^{\frac{E_d + E_t}{2 k_B T}}$$

$F$  = electric field,  $k_B$  = Boltzmanns constant,  $\mu$  = electron mobility,  $N_d$  and  $N_t$  are the volume densities of donors and traps, respectively.  $E_d$  and  $E_t$  are the corresponding energies.  $N_c$  is the effective density of states in the

conduction band, and  $\epsilon_i$  is the high frequency dielectric constant. The quantity  $a$  is unity for standard Poole-Frenkel conduction and two in the modified process. Note that in the latter case the factor  $\beta/a$  mimics the value of  $\beta$  that scales the energy barrier lowering term in the standard expression for Schottky emission. This fact is often used to explain what is referred to as anomalous Poole-Frenkel conduction: bulk conduction that has the exponential dependence on electric field and temperature expected for interface-limited Schottky emission. The electrical conductance described by Eq. (4.1) grows rapidly with increasing temperature. Consequently, as the current driven through the NbO<sub>x</sub> increases the resulting Joule-heating induced temperature rise leads to increased conductivity and, therefore, greater power dissipation. This produces further increases in temperature. At a critical current, this positive feedback results in NDR. A simple but accurate compact model for this behavior is obtained by assuming the temperature,  $T_N$ , of the active region of the NbO<sub>x</sub> is uniform and described by

$$C_{th} \frac{dT_N}{dt} = \frac{T_{amb} - T_N}{R_{th}} + IV \quad (4.2)$$

Here,  $R_{th}$  is the effective thermal impedance between the current-carrying portion of the NbO<sub>x</sub> and the surrounding ambient environment. Similarly,  $C_{th}$  is the effective thermal capacitance of the active region. This is illustrated by an equivalent thermal circuit in Figure 4.1. In this model, the NbO<sub>x</sub> can be viewed as a locally active memristor[26], with the temperature  $T_N$  as the dynamical state variable. Eqs.(4.1) and (4/2) serve, respectively, as the instantaneous conduction and dynamical state equations in this formalism. We

have used Eqs. (4.1) and (4.2) to simultaneously fit sets of quasistatic V-I curves taken over a range of ambient temperatures for type I and type II formed devices with a variety of NbO<sub>x</sub> layer thicknesses and bottom electrode diameters. Representative results for a 52 nm diameter device with an 8 nm thick NbO<sub>x</sub> film that underwent type I forming are displayed in Figure 4.3(a) for temperatures between 275 and 450 K. These data were taken using an MMR Technologies Variable Temperature Microprobe System. Our compact model accurately matches the data. The fitting parameters were determined by first plotting the natural logarithm of the measured low bias conductivity as a function of 1/T as indicated in Figure 4.3(b). The slope and intercept of this Arrhenius plot yield the energy  $E \equiv (E_d + E_t)/2$  and prefactor  $\sigma \equiv e\mu N_C (\frac{N_d}{N_t})^2$  because in the low field limit Eq. (4.1) implies

$$\sigma_{low}(T) = \frac{a^2 + 1}{2a^2} \sigma_o(T) = \sigma_p \left( \frac{a^2 + 1}{2a^2} \right) e^{\frac{-E}{k_B T}} \quad (4.3)$$

In all cases, we assumed  $a=2$ . The full V-I curves were then matched at all the measured ambient temperatures by choosing a single temperature-independent value for each of two parameters:  $R_{th} = 1.27 \times 10^6$  K/W and  $\epsilon_i = 22$ . Similar temperature-dependent sets of V-I curves for devices with diameters between 32 and 165 and NbO<sub>x</sub> thicknesses from 8 to 42 nm were modeled with comparably close agreement between theory and data. The values determined for the activation energy  $E$  ranged from 0.15 to 0.24 eV. In all cases, a value of  $\epsilon_i = 22$  worked well, implying an index of refraction at frequencies near the visible of  $n_i = 4.7$ . This is in reasonable agreement with ellipsometric measurements on NbO<sub>x</sub> films with compositions close to NbO<sub>2</sub>.



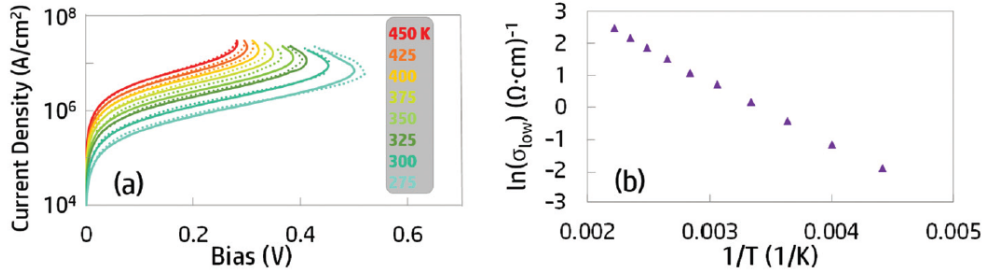


Figure 4.3: (a) Measured (solid) and calculated (dashed) V-I curves for TiN/NbO<sub>x</sub>/TiN selector with 52 nm diameter bottom electrode and 8 nm thick NbO<sub>x</sub> layer for  $T_{amb} = 275\text{--}450$  K. (b) Arrhenius plot of the conductance measured at biases low enough for it to be ohmic.

### 4.3 Back-to-back Schottky Diode

The nonlinear (NL) device must be Back-End-Of-Line (BEOL) compatible because most RRAM processes are currently integrated in BEOL. In this section, we present highly NL devices, which are fast, robust, have low operating voltage, with high current density, and offer good scalability. They use a simple fabrication processes employing symmetric back-to-back Schottky diodes by metal-semiconductor-metal(MSM) structure. We also discuss impacts of material choices and device geometry on performance of these MSM diodes by numerical simulation, Schottky diode current characteristics and experimental results. We started with simulations on Sentaurus to evaluate the feasibility of MSM diode as selector device for future RRAM crossbar arrays. We are interested in the correlation between performance metrics of selector device and physical properties of MSM diode. In our simulation setup, we chose metal work function to match experimental values of Schottky barrier height values reported in literature which means we take Fermi-level pinning

into account. We assume the dominant current transport mechanism in this ultra-thin MSM diode to be thermionic emission, recombination and tunneling. The doping concentration is assumed to be n-type doped ( $10^{13}\text{cm}^{-3}$ ). DC characteristics in Figure 4.4(a) with different thickness for Schottky barrier height of 0.6 eV, which is close to experimental values of titanium of silicon barrier. We see a thickness dependence on current density and NL ratio. The definition of NL ratio in this paper follows half-read scheme which is widely used in many memory systems. The NL ratio is defined as the current density ratio of 1 MA/cm<sup>2</sup> and the current at half of voltage where current density reaches 1 MA/cm<sup>2</sup>. As thickness increases, the NL ratio decreases. This is probably because the voltage region where current grows exponentially is delayed by higher resistance in the diode, and higher series resistance in thicker film also affects effective voltage drop in diode at higher current region. This makes I-V characteristics of thicker diodes deviate from ideal exponential curve at lower current compared to thinner diodes, thus NL ratio in thicker film is lower. The relation between Schottky barrier height and current density in Figure 4.4(b) agrees well with simple Schottky diode model. In our simulation, we found out the high doping concentration adversely affects NL ratio. This is intuitive since CMOS technology has been using highly doped source and drain to achieve ohmic contact, which contradict our purpose of reaching a high NL ratio. We will discuss more details about current characteristics of MSM diode later, and the effect of Schottky barrier height in current characteristics of MSM diode is similar to simple Schottky diode. Higher Schottky

barrier height results in lower current density near zero bias. One important design consideration of selector device is to have both high NL ratio and high current density at low voltage. Hence, based on the above simulation results, the desired MSM diode should have thin thickness of semiconductor layer (less than 14nm), low doping concentration and appropriate Schottky barrier height to give decent current density while keeping good NL ratio.

After optimizing the design by numerical simulation, devices were fabricated on a n-type Si (111) substrate with 300 nm plasma-enhanced chemical vapor deposition (PECVD) grown silicon dioxide on the top as an isolation layer. 80nm bottom electrode (BE) was formed by electron beam evaporation at 273K. 10nm to 20nm semiconductor layer was deposited on the top of BE by PECVD at 250°C without *ex-situ* annealing. The growth condition and parameters were adopted from Moravej *et al.*[84], to aim to grow nanoscale hydrogenated amorphous silicon thin film. Then 80nm top electrode (TE) were formed by electron beam evaporation of titanium at 273K. Devices were patterned as crossbar with variant device perimeters from 300nm to 10 $\mu$ m by electron beam lithography. Electric measurements were taken by Agilent Semiconductor Parameter Analyzer B1500 and Lakeshore CRX-VF Probe Station. Pulse measurement was carried out by Agilent B1525 Pulse Generator Unit. Devices were measured by applying voltage to the TE while the BE is grounded. The equivalent circuit diagram in Figure 4.5(a) illustrates series resistance in MSM diode. The design of the crossbar is targeted to minimize impacts on parasitic components when doing pulse measurements. It can be

seen in Figure 4.5(b) metal lines are tapered to reduce parasitic capacitance. The overlap area of BE and TE defines the area of MSM diode.

Figure 4.6(a) compares experimental DC I-V characteristics of Ti–aSi–Ti for different thickness of silicon layer. The effect of series resistance in 20nm amorphous Si MSM diode can be observed in the region highlighted by dashed circle. Ti–aSi–Ti has higher current density than Ni–aSi–Ni at the same thickness of amorphous silicon in Figure 4.6(b), which implies amorphous silicon is unintentionally doped with n-type impurities. For n-type silicon, nickel tends to pin closer to valence band of silicon while titanium is pinned at midgap. Thus we observe lower current density in Ni–aSi–Ni. NL ratio in Figure 4.6(c) agrees well with Figure 4.6(c), which also indicates series resistance plays an important role when designing high current density selector. The inset of Figure 4.6(c) shows linear scale plot of current density. It is obvious that there is asymmetry of current density between each polarity for titanium and nickel MSM diodes. Amorphous silicon and electrodes weren't deposited in the same instrument or same vacuum environment. So this causes inevitable interface difference between the two Schottky diodes. To extract Schottky diode parameters from I-V characteristics of MSM diode, we need to start with simple Schottky diode current characteristics and combine with some assumptions to derive an insightful current equation for MSM diode device. Later we will point out when ideality factor larger than 1, MSM diode also demonstrates asymmetric I-V curve. Overall, experimental results are in good agreement with simulation prediction except for a lower current den-

sity, probably because in simulation some materials-related parameters such as effective mass, recombination coefficients, and bandgap are not an very accurately known for amorphous silicon.

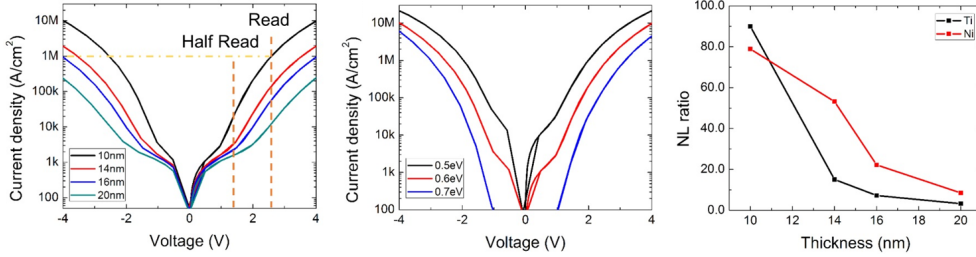


Figure 4.4: (a) DC I-V characteristics from two-dimensional numerical simulation by Sentaurus from Synopsys Inc.. Two orange dashed lines represent voltage values calculated for NL ratio. The choice of  $1\text{MA}/\text{cm}^2$  is based on matching current density of RRAM devices. (b) I-V characteristics on Schottky barrier height dependence. Note that non-linear step size of simulation caused I-V curves to show a small hysteresis at low bias during positive polarity sweep, which is an artifact. (c) NL ratio extracted from Figure 4.4(a)(b).

We want to understand current-voltage characteristics of MSM diode so that we can extract fundamental parameters such as Schottky barrier height and ideality factor by fitting experimental I-V curves. The band diagram in Figure 4.5(a) indicates that MSM diode can be considered to be back-to-back Schottky diodes. The current-voltage equation of a simple Schottky diode is in the following form:

$$J = J_s \exp\left(\frac{qV}{\eta kT}\right) \quad (4.4)$$

$J_s$  is reverse saturation current of Schottky diode:

$$J_s = A * T^2 \exp\left(\frac{q\phi_b^0}{kT}\right) \quad (4.5)$$

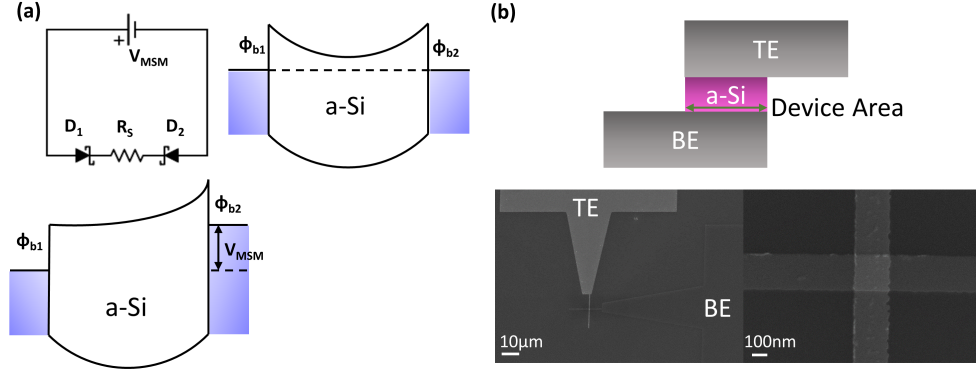


Figure 4.5: (a) Band diagrams at equilibrium and forward bias on D1. The equivalent circuit diagram was shown at forward bias on D1. Series resistance is labeled as  $R_S$ .  $\phi_{b1}$  and  $\phi_{b2}$  represents Schottky barrier height at D1 and D2. (b) Cartoon illustration of cross-sectional view of device. The green arrow indicates the effective area of device.

where  $A$  is Richardson constant,  $T$  is temperature and  $k$  is Boltzmann constant. The voltage drops in two Schottky diodes are  $V_1$  and  $V_2$  and voltage across MSM diode is  $V_{MSM} = V_1 + V_2$ . We can also write  $J_1 = -J_2$  from current continuity. Based on these two conditions we can write current-voltage equation as in Nouchiet *al.*[85]:

$$J_{MSM} = \frac{2J_{S1}J_{S2} \sinh\left(\frac{qV_{MSM}}{2\eta kT}\right)}{J_{S1} \exp\left(\frac{qV_{MSM}}{2\eta kT}\right) + J_{S2} \exp\left(\frac{-qV_{MSM}}{2\eta kT}\right)} \quad (4.6)$$

The equation above can be further simplified if Schottky barrier height of diode 1 and diode 2 are symmetric:

$$J_{MSM} = J_S \tanh\left(\frac{qV_{MSM}}{2\eta kT}\right) \quad (4.7)$$

However, the current obtained from (4.7) reaches saturation at very low bias; this is contradicting to reported experimental results[86, 87] and our experi-

mental data. Note that (4.7) doesn't consider image-force lowering. Image-force lowering is based on image charges in metal layer induced by charges in semiconductor layer near metal-semiconductor interface. Image charges establish an electric field along metal-semiconductor interface and Schottky barrier height becomes voltage dependent:

$$\phi_{b1}(V_1) = \phi_{b1}^0 + qV_1\left(1 - \frac{1}{\eta}\right) \quad (4.8)$$

$$\phi_{b2}(V_2) = \phi_{b2}^0 + qV_2\left(1 - \frac{1}{\eta}\right) \quad (4.9)$$

where  $\phi_b^0$  is Schottky barrier height under zero bias and  $\eta$  is ideality factor. Assuming voltage drop in MSM diode is mostly on diode 2 since diode 1 is forward bias in Fig.2(a), we have  $V_{MSM} \approx V_2$ . We also assume (4.8) equals to (4.9) because of symmetric MSM structure, then we can rewrite (1):

$$J = J_S \sinh\left(\frac{qV_{MSM}}{2kT}\right) \exp\left(\frac{qV_{MSM}}{2kT}\right) \exp\left(\frac{-qV_{MSM}}{\eta kT}\right) \quad (4.10)$$

With Eq.(4.10) we can evaluate Schottky barrier height and ideality factor in MSM diodes and discuss the impacts on performance matrices of selector device. Fig.3(d) shows fitting by Eq. (4.10) on 10nm Ti-aSi-Ti and Ni-aSi-Ni MSM diode. The Schottky barrier height between titanium and amorphous silicon extracted by curve fitting with Eq. (4.10) is  $\approx 0.79eV$ , and the Schottky barrier height between nickel and amorphous silicon is about  $0.86eV$ . Note that the bandgap of hydrogenated amorphous silicon is around  $1.6-1.8eV$ . Titanium is usually pinned at midgap at silicon interface and nickel is usually pinned closer to valence band of silicon, so the barrier height values extracted from

Fig3.(d) are what we expect. The ideality factor extracted from Figure 4.6(d) are 1.18 and 1.36 for titanium and nickel respectively. Eq. (4.10) implies that non-ideal ideality factor can be attributed to unequal current density in MSM diode. The magnitude of current density difference between each polarity is sensitive to ideality factor. As mentioned in the previous section, two Schottky diodes might have slightly different interfaces with silicon. This might cause slightly different Schottky barrier height and thus contribute to asymmetric I-V characteristics. Although there is asymmetric current density for the same voltage when applying different polarity, the ratio of current density asymmetry is less than one order, as seen in Figure 4.6(c). The voltage difference to reach  $1\text{MA}/\text{cm}^2$  is only  $0.1\text{V}$  for different voltage polarity for both Ti–aSi–Ti and Ni–aSi–Ni MSM diode. This will not impact the operation of selector device, but the effect of non-ideal MSM diode and interface properties should be taken into account when applying the selector device in RRAM arrays.

We performed transient analysis to test the speed of Ti–aSi–Ti device. The test setup is shown in Figure 4.7(a). We did impedance matching for the source to ensure the actual voltage drop in device under test (DUT) was correct and we normalized the current scale measured at oscilloscope because input channel impedance was  $50\text{ ohm}$  but the impedance through DUT was a few thousands ohm. It can be seen that in Figure 4.7(b) MSM diode can response to  $60\text{ns}$  pulse without any notable delay or distortion of signal. The current overshoot was only  $25\%$  higher than the mean current level during pulses. The result of transient analysis is reasonable because the dominant



current transport in MSM diode is thermionic emission. This makes MSM diode favorable over pn junction diode which is based on minority carrier injection.

MSM diode also shows excellent reliability as potential selector device. Figure 4.8(a) illustrates cycle test scheme of MSM diode. This is very similar to the test scheme of RRAM device since selector devices need to be integrated with RRAM cells. After each set or reset pulse, DC read at read voltage or half read voltage was performed to record the value of resistance. Figure 4.8(b) shows MSM diodes survived after  $10^8$  cycles without any notable degradation. Besides, MSM diodes demonstrated very good DC stress test, as seen in Figure 4.8(c) at half-read voltage over  $10^3$  seconds. These test results validate that the MSM diode is very robust and a suitable candidate of selector device.

#### 4.4 Conclusion

This chapter introduces two devices with potential for selector device of RRAM :  $\text{NbO}_x$  based NDR device and highly nonlinear MSM diode. The Ti-a-Si-Ti MSM diode with Schottky emission provides very high nonlinear ratio and low turn-on voltage, which can help eliminate sneak path leakage in RRAM array and MSM diode with over  $10^5$  nonlinear ratio is capable for building gigabyte (GB) crossbar array. The asymmetric I-V characteristics can be modeled by different Schottky barriers, which indicates that top electrode interface and bottom electrode interface can cause asymmetry I-V curves.

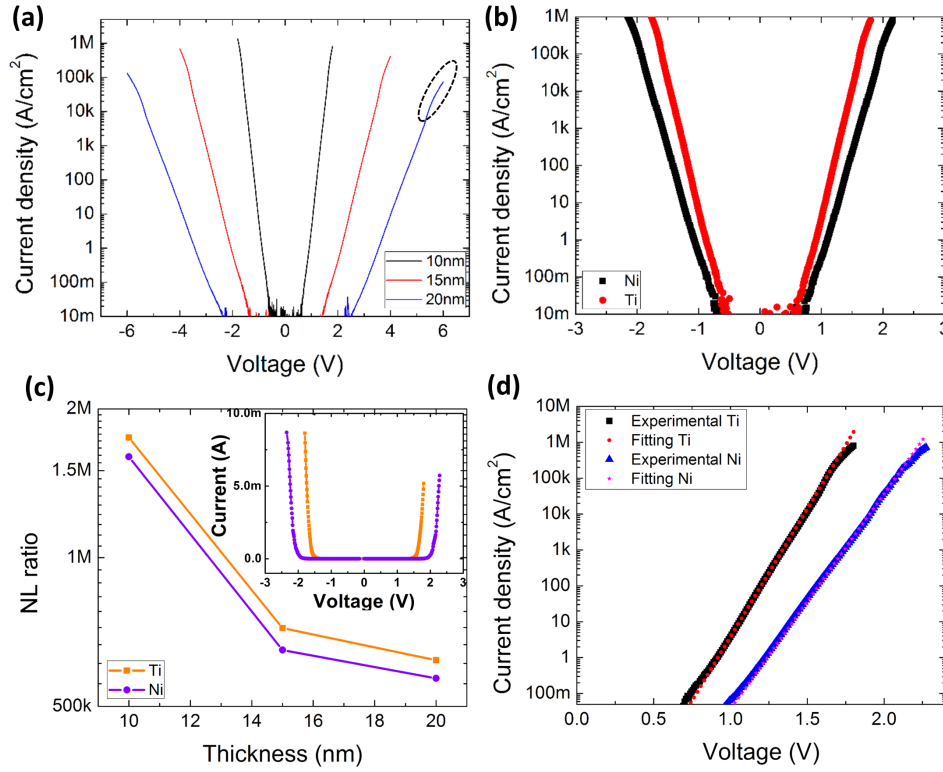


Figure 4.6: (a) Semilog plot of I-V characteristics of different thickness Ti-aSi-Ti diodes. The dashed circle is where series resistance effect becomes prominent. (b) I-V characteristics of Ti-aSi-Ti and Ni-aSi-Ni with 10 nm amorphous silicon. (c) Comparison of the NL ratio between Ti-aSi-Ti and Ni-aSi-Ni at different thickness. The inset of (c) is linear scale I-V characteristics of (b). (d) I-V curve fitting by using (7). (7) fits well at intermediate bias where series resistance effect was low and thermionic emission is dominant.

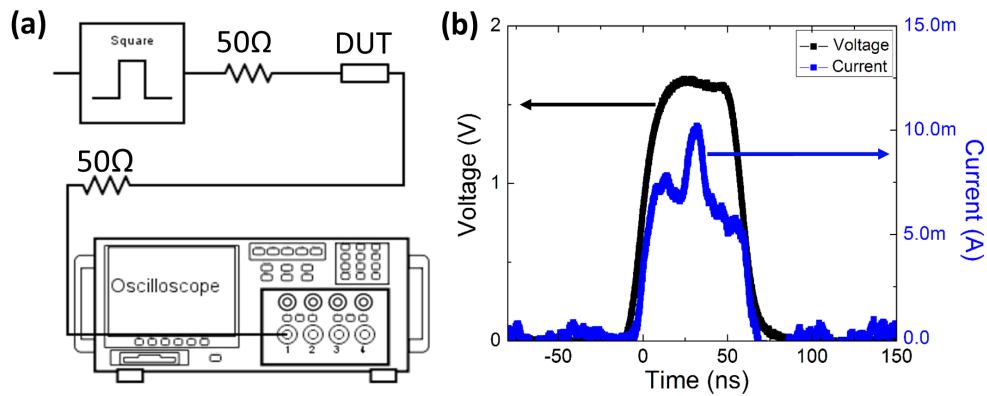


Figure 4.7: (a) Pulse measurement setup and impedance at each stage. The pulse generator has an output impedance of  $50\ \Omega$ , and input channel impedance of oscilloscope is  $50\ \Omega$  as well. (b) shows the transient current response of DUT. Black curve and arrow is input pulse generated by source and blue curve and arrow is current response of MSM diode. The device is a  $300\ \text{nm} \times 300\ \text{nm}$  crossbar.

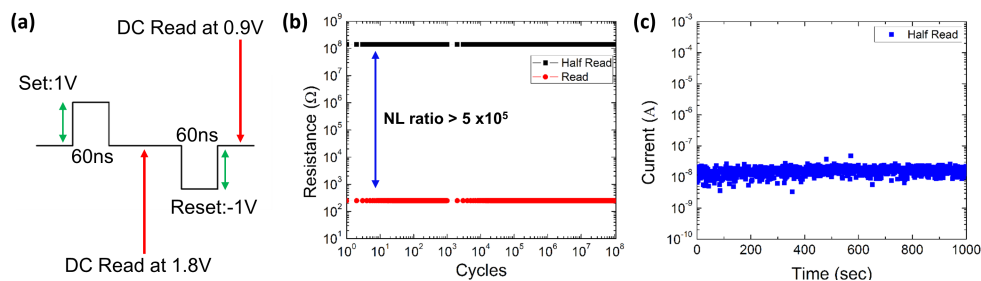


Figure 4.8: (a) Test scheme of cycle test. (b) Endurance test up to  $10^8$  cycles. (c) DC stress test at 0.9 V for 1000 seconds.

## Chapter 5

### Short Term Relaxation of RRAM

#### 5.1 Introduction

Data programming stability is one of key factors for evaluating emerging non-volatile memory technology. The cycle-to-cycle stability, sometimes called repeatability, has been studied in the previous chapters. In this chapter, another programming stability issue will be discussed, which is usually referred as short-term relaxation. This phenomenon relates to the decay of the resistance state right after applying a programming pulse. The short-term relaxation occurs at a timescale on the order of  $\mu\text{s}$  to a fraction of second. As such, this phenomenon is different from the retention issues cycle-to-cycle stability. This problem is particularly important for the implementation of adaptive programming algorithms and artificial neural networks (ANN) because in these implementations the output is fed back in short time intervals (e.g. recursive neural network (RNN)). Fantini[88] et al. has reported the pioneering statistical study of this problem in HfOx-based RRAMs. In this chapter, we systematically examine the short-term program instability and its ensuing reliability issues in our  $\text{CeO}_x$  bilayer RRAM with and without a selector device. Our selector device is a MSM diode, which provides high non-linearity (NL) ratio and a current density in excess of ( $1\text{MA}/\text{cm}^2$ ). The

RRAM device is made of  $\text{CeO}_x$  filament and a sub-stoichiometric  $\text{HfO}_x$  oxygen scavenging layer[89]. Figure 5.1 shows cumulative distribution function(CDF)-like curves of resistance at HRS and LRS evolve as time elapses, and the gap between HRS and LRS gradually shrinks as time elapses.

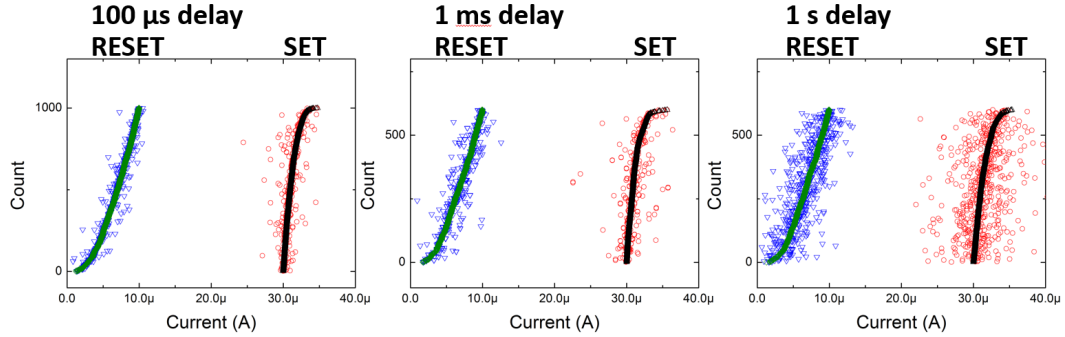


Figure 5.1: Green and black curves in Figure 5.1 are initial resistance of HRS and LRS after programming to RRAM cell. Blue triangles and red dots are resistance after given time delay. From left to right, the time delay between initial read and delay read is  $100 \mu\text{s}$ ,  $1\text{ms}$ , and  $1\text{s}$ , respectively.

## 5.2 Device Fabrication and Measurement

In our process, the MSM selector diode and the RRAM device are stacked vertically. The MSM diode was fabricated first followed by the fabrication of the RRAM device on top. For short-term relaxation studies, we focused on the adaptive programming algorithm because it provides better endurance compared to the open-loop single pulse programming scheme. Figure 5.2 illustrates the procedure for the adaptive programming algorithm and the continuous pulse reads. Figure 5.3 shows the testing setup for adaptive programming algorithm. Field programmable gate array (FPGA) board provides

flexibility to design experiments for fast-sampling tests. Digital-to-analog data converter (DAC) was used to generate test pulses to device under test (DUT). Transimpedance amplifier (TIA) and analog-to-digital data converter (ADC) send the response from DUT back to FPGA. Here we used 200 ns pulses for write and read operation. The read pulses were applied for up to 1s after each successful SET or RESET. Continuous read pulses provide real-time information about the fluctuation and the relaxation of the stored bit in the RRAM device under test, which is important for understanding the short-term relaxation phenomenon.

### 5.3 Results and Discussion

To examine the effect of the MSM selector on the short-term relaxation, we compared the post-programming read current distributions of two structures, namely a one-RRAM (1R) structure and a one-selector-one-RRAM (1S1R) structure. All experiments were performed on virgin devices to eliminate the possible effect of device history on the short-term relaxation behavior. These devices were preconditioned by switching 50 cycles using the adaptive programming algorithm while skipping the continuous read step. The continuous read steps were disabled during device preconditioning to avoid possible degradation of the RRAM operating window. The operating window of an RRAM is defined as the ratio of the resistance at HRS to LRS. After preconditioning, the devices were switched once with the adaptive programming algorithm, followed by a continuous read step that lasts up to 1s. This test

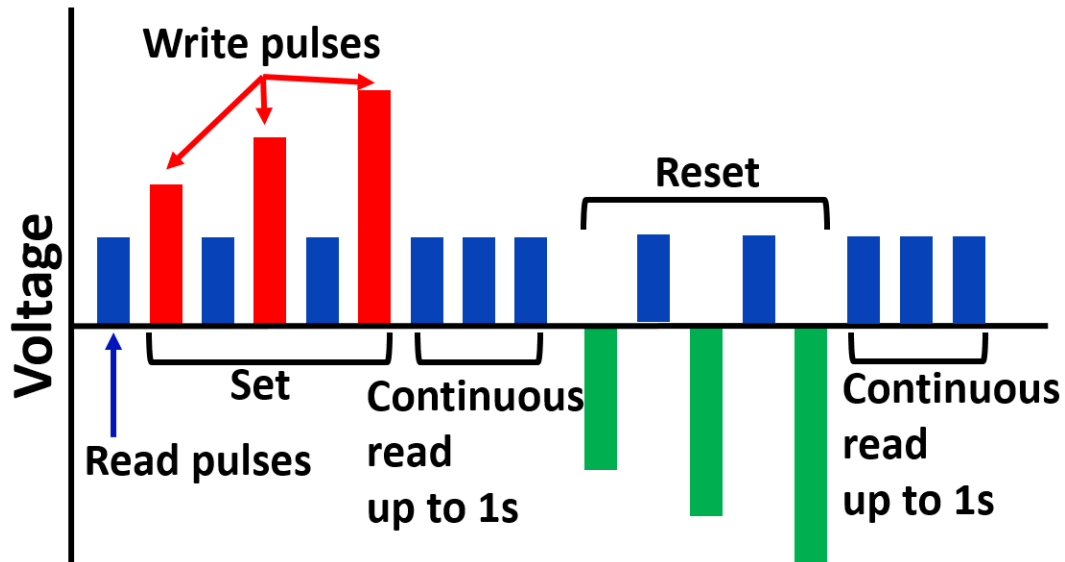


Figure 5.2: Schematic illustration of the adaptive programming algorithm. Read pulses are 200ns at 0.2V. Write pulses start at 0.2V with 0.1V increment. The interval between each pulse is  $10\mu s$ . At each attempt to switch the resistive state of the device, a write pulse is applied to the RRAM device followed by a read pulse to check whether the value of the read current is higher(lower) than the target value during the SET(RESET) cycles. If this condition is met, we consider this attempt as a successful SET (RESET).

procedure enabled us to distinguish the short-term relaxation from the long-term degradation induced by repetitive read cycles. The test results of a 1R device is shown in Figure 5.4(a), revealing the presence of two distinct current states at the LRS and a continuum of random states at the HRS. Figure 5.4(b) shows the distribution of the read current for a 1S1R structure. Compared to the 1R structure, the 1S1R structure demonstrates larger spread in the read currents at both the HRS and the LRS. These results indicate higher degradation of the operating window of the 1S1R structures, which can conse-

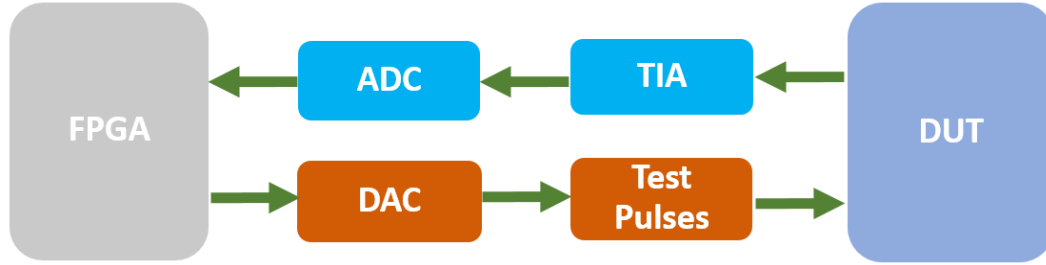


Figure 5.3: Testing setup for fast sampling measurement. LabView software was used for programming and controlling FPGA and Python was used for data processing after data collecting by FPGA and LabView.

quently lead to the increase of the raw read error rate. The degradation of the operating window is particularly important for devices with small operating current. Increasing the resistance difference between the HRS and the LRS is a potential solution for mitigating this problem. However, we observed that increasing the operating window of our bilayer  $\text{CeO}_x$ -based RRAMs degrades the endurance properties of the device. This issue will be discussed in more detail later.

The continuous read procedure affords higher time resolution, which is important for studying the short-term relaxation issue. In Figure 5.5(a)(b), with continuous read, we are able to observe random telegraph noise (RTN) in 1R and 1S1R. And it can be seen that RTN in MSM device has current amplitude dependence. This dependence affects short-term relaxation in 1S1R device. To clearly illustrate the change in the operating window between 1R and 1S1R structure, the HRS curves were plotted in the form of  $1 - p(x)$ . Furthermore, the probit unit [90] was used in Figure 5.6 because it linearizes



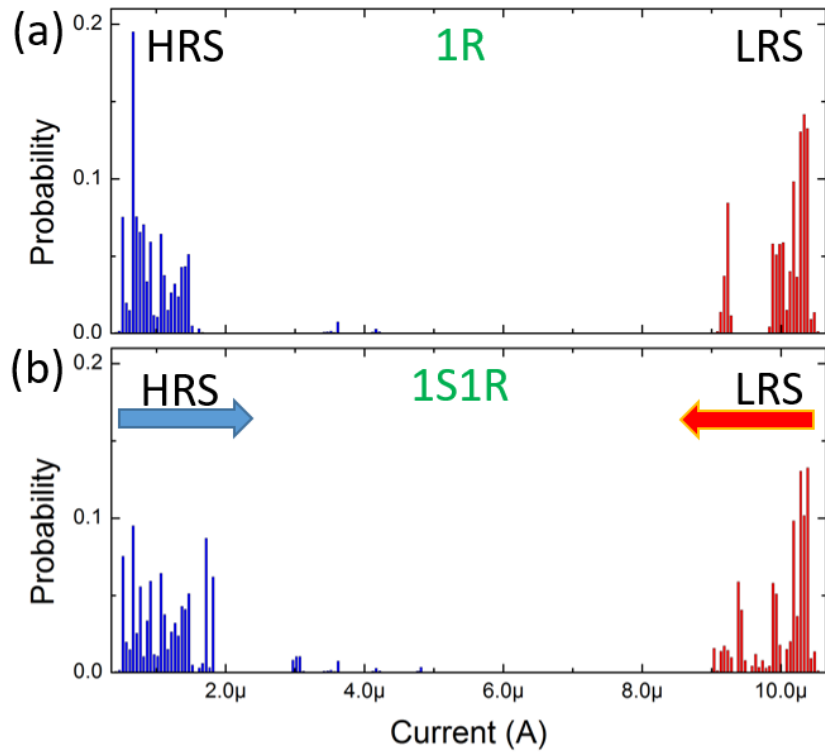


Figure 5.4: Distribution of HRS and LRS read currents for (a) 1R and (b) 1S1R structures.

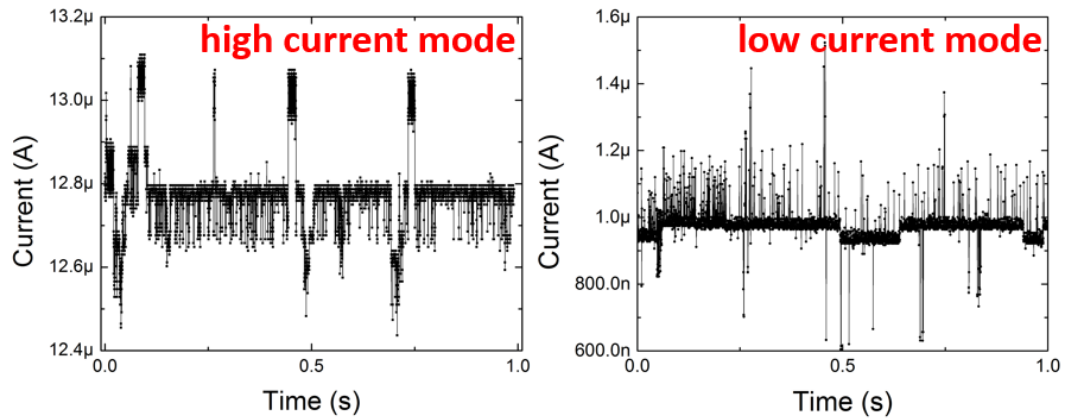


Figure 5.5: RTN noise at different current amplitude of MSM selector. (a) operating at  $12.8\mu\text{A}$ , (b) operating at  $1\mu\text{A}$ .

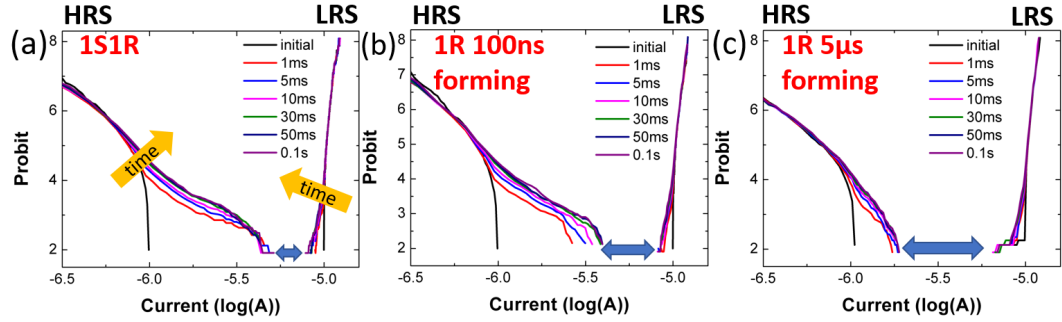


Figure 5.6: CDF plots using probit units for (a) 1S1R, (b) 1R with 100ns forming pulse, and (c) 1R with  $5\mu\text{s}$  forming pulse. The arrows labeled "time" in (a) indicate the temporal progress of the experiment. The horizontal arrows indicate the gap between tail bits of the HRS and the LRS. Probit of the HRS curves are plotted in a decreasing fashion versus current, while the LRS curves are plotted in an increasing fashion versus current. The read pulse voltage in (a) is 1.7V because of voltage drop at selector device.

a lognormal-like distribution. In The measurement results reveal that the tail bits of the 1S1R structure in Figure 5.6(a) exhibit a narrower gap than the 1R device in Figure 5.6(b). This is possibly due to strong random telegraph noise (RTN) from MSM device in HRS. Note that the MSM selector devices typically exhibit RTN behavior, as shown in the Figure 5.5. The RTN in an MSM device primarily originates from the charge trapping at the interfaces between the metal electrodes and the semiconductor [91]. The peak-to-peak variation in Figure 5.5(a) is about  $0.6\mu\text{A}$  and the peak-to-peak variation is about  $0.8\mu\text{A}$  in Figure 5.5(b). The lower current amplitude with larger noise fluctuation indicates worse signal to noise ratio. Thus, in a 1S1R structure, the noise of the MSM diode can give rise to the fluctuation of the voltage drop across the RRAM device in HRS, thereby increasing to the overall inherent

noise of the RRAM in a 1S1R structure compared to that of the 1R structure in HRS. Since the implementation of the 1S1R structures are pursued for most practical applications, it is important to consider their short-term relaxation behavior when designing algorithms for ANN applications.

Next, we examined the effect of the forming pulse duration on the short-term relaxation behavior of the 1R structures, shown in Figure 5.6(b) and (c). The RRAM device in Figure 5.6(b) was formed using a 100ns long pulse, while the device in Figure 5.6(c) was formed using a  $5\mu\text{s}$  long pulse. The comparison of the test results in Figure 5.6 clearly illustrates the improvement of the tail bits distribution for the RRAM device formed using the  $5\mu\text{s}$  long forming pulse. This observation can be explained using the previously proposed model based on the diffusive dynamics [92–95]. The thermodynamic stabilization is the basis for this model, in which the charged defects are assumed to relax to a thermal equilibrium state after biasing the RRAM cell. The use of longer forming pulses is expected to give rise to the formation of more stable states for charge defects, thereby mitigating the short-term relaxation issue. Furthermore, we compared the short-term relaxation behavior of the bilayer  $\text{CeO}_x/\text{HfO}_x$  device in Figure 5.6(b) with the  $\text{HfO}_x$  device in [88], indicating faster relaxation of the bilayer device. We attribute this observation to the increase of the diffusion pre-factors at the interface of the  $\text{CeO}_x$  and the  $\text{HfO}_x$ , confirmed by the hybrid density functional theory and molecular dynamics (DFT-MD) simulations [96].

Lastly, we investigated the effect of the forming pulse duration on the

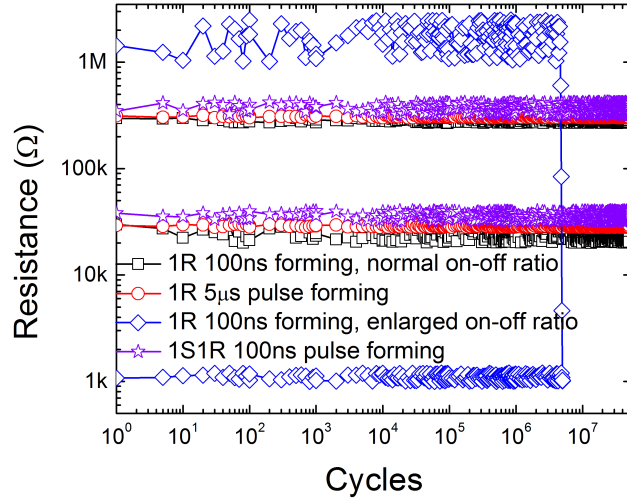


Figure 5.7: Effect of the operating window and the forming pulse duration on the long-term reliability endurance of RRAMs. Devices were subject to the adaptive programming with a read pulse width of 200 ns.

long-term endurance of the device. Figure 5.7 shows the comparison of the endurance behavior for 1R devices with 100ns forming pulse and  $5\mu s$  forming pulse. Furthermore, we examined the effect of the operating window on the endurance characteristics of 1R devices, shown in Figure 5.7. Black square curve in Figure 5.7 is 1R device operated at typical on-off ratio, which is LRS  $20\text{ k}\Omega$ , HRS  $200\text{ k}\Omega$ . The blue diamond curve in Figure 5.7 represents 1R device with intentionally enlarged on-off ratio by setting switching threshold in adaptive programming, and this gave LRS around  $1\text{ k}\Omega$ , HRS around  $1\text{ M}\Omega$ . A 100ns pulse was used for forming the 1S1R structures. As pointed out earlier, increasing the operating window can mitigate the short-term relaxation issue. However, our results indicate the degradation of the endurance charac-

teristics of the bilayer RRAM device upon increasing the operating window. Our findings suggest that the use of longer forming pulses might be a more favorable solution because it does not compromise the device endurance while mitigating the short-term relaxation problem.

## 5.4 Conclusion

In this chapter, we examined the short-term program instability of the 1S1R and the 1R structures using  $\text{CeO}_x$ -based bilayer RRAM devices. Our results indicate that the 1S1R structures are more susceptible to the short-term relaxation. Furthermore, we found out that increasing the forming pulse width can alleviate the short-term relaxation without compromising the long-term endurance.

## Chapter 6

# Neuromorphic Applications

### 6.1 Introduction

In 1971, Chua[26] postulated that there should be a missing circuit element besides resistor, capacitor, and inductor. He named this element as "memristor" because it can retain its resistance state. In 2008, R.S.Williams from Hewlett Packard Labs(HP Labs) claimed that his team discovered the missing memristor[40], which has similar material stack and I-V characteristics as RRAM. L.Chua confirmed that the HP memristor is the missing memristor. Chua's definition of generic memristor[3] is that any two-terminal device that exhibits a pinched hysteresis loop in the voltage-current plane when driven by any periodic voltage or current signal that elicits a periodic response of the same frequency. Figure 6.1 shows a voltage controlled generic memristor I-V characteristics. Based on this relaxed definition of memristor, biological synapses and RRAMs can be considered as generic memristors. In chapter 3, the bilayer  $\text{HfO}_x/\text{CeO}_x$  has demonstrated long-term potentiation and depression behaviors of human synapses. This raises a lot of interest in RRAM for building brain-liked computers. Traditional Von-Neumann architecture has bottleneck in parallel computing due to the bandwidth limit between central processing unit(CPU) and memory. To overcome the bot-

tleneck, a new computing paradigm has been proposed by some research groups[2, 46, 48, 49, 97, 98]. This new architecture emulates human brain functions, especially in Hebbian learning. The brain-inspired computing paradigm sometimes is referred as neuromorphic computing. CMOS based neuromorphic circuits has been reported; however, this approach consumes too much power. Note that the average power consumption of human brain is about 15 mW, which is much lower than the power of personal computer(PC) CPU, which is about 90W[42]. Low-power memristor synapse circuits combines with CMOS neuron circuits can meet low power consumption while fulfilling other performance requirements. Thus, it's worth to study RRAM devices for neuromorphic applications.

## 6.2 Supervised and Unsupervised Learning

One of potential applications of neuromorphic computing is supervised learning. Typically speaking, machine learning tasks can fall into two types: supervised and unsupervised learning. The key difference between supervised learning and unsupervised learning is, during the training stage, whether the training data is labeled or not[99]. The labeled training data is fed to the model as the input; on the other hand, the learning model of unsupervised learning tries to cluster the unlabeled training data into groups. The number of cluster groups to be petitioned is defined by the model, and the petitioning is purely based on the statistical properties of the input data. Figure 6.2 shows various types of learning tasks in each category. Figure 6.3 illustrates

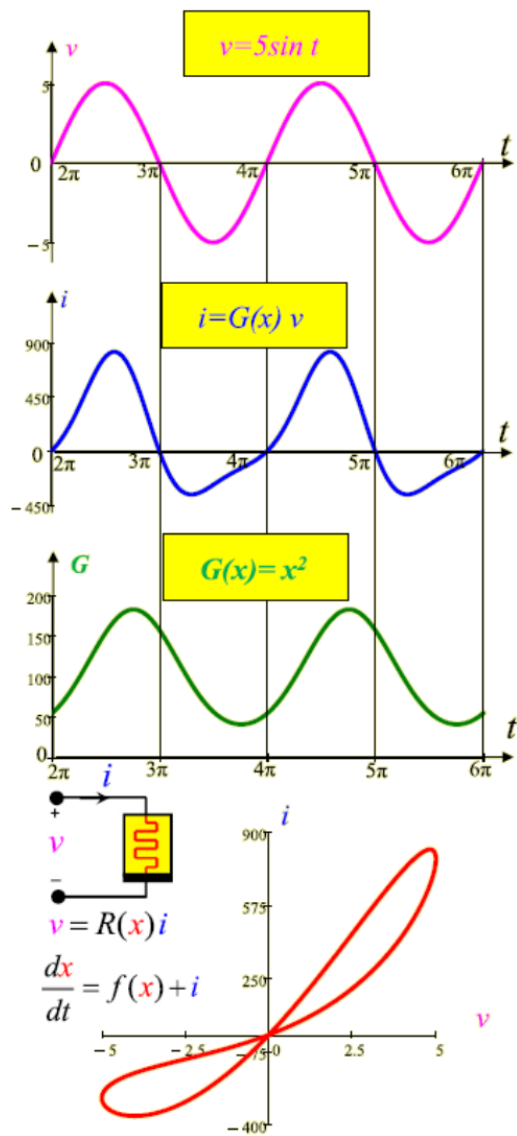


Figure 6.1: The I-V characteristics of voltage controlled generic memristor can be derived from state dependent Ohm's law and state equation. The I-V curve shown here is from a sinusoidal voltage source.[3]

the difference between classification and clustering. In classification, the input data during training process is labeled. In this example, the labels are black



cross and smiley face. The classifier is the model which classify labels. The classifier shown in Figure 6.3 is a support vector machine(SVM). SVM is a linear classifier which provides easy implementation and quick analysis on the input data[100]. However, it's not a very accurate classifier when the input data is not linearly separable. For the clustering model in Figure 6.3, it can be seen that all the data are gray triangles and they are not labeled. There are three groups of data separated by three red dashed circles. In this example, each group of data is a cluster. The clustering model itself defines how many clusters there should be in the model, while in classification, the number of classes in classifier is determined by the number of classes seen in the training data.

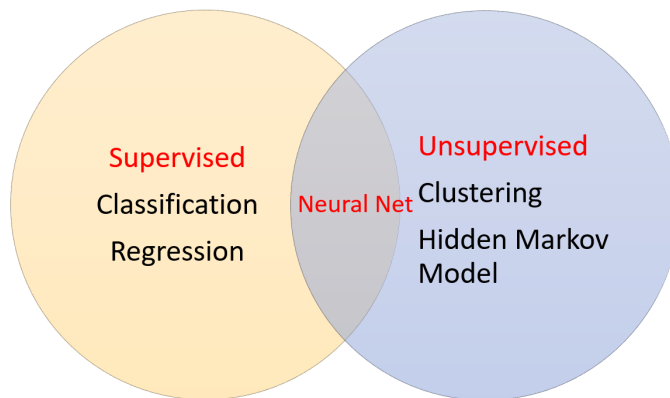


Figure 6.2: Yellow circle represents supervised learning and blue circle represents unsupervised learning. The intersection of two circles is the "intermediate" learning.

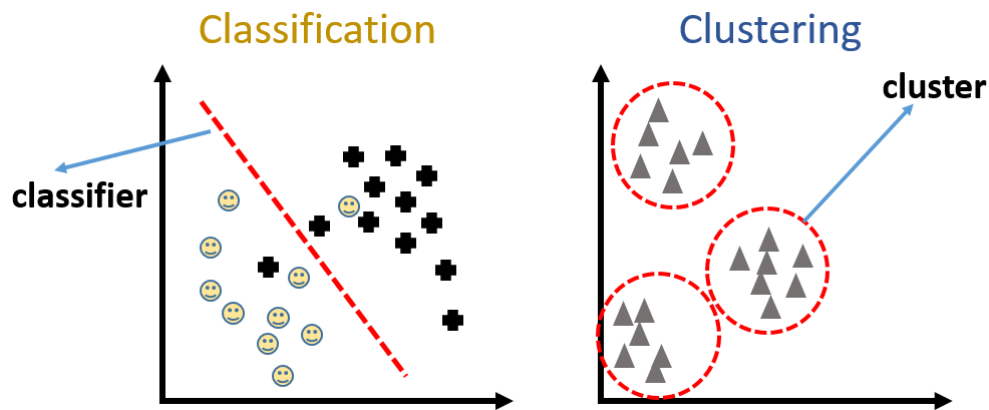


Figure 6.3: The black crosses and smiley faces represent the input data for classification. The red dashed line is the SVM classifier. The gray triangles in the right diagram represent the input data for clustering.

### 6.3 Neural Networks

In recent years, artificial neural network (ANN) has become popular and raised a lot of interest among machine learning researchers. A simple definition of a neural network by M.Caudill[101] is as: *"a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs"*. All types of ANNs have learning rules which determine the way that synapses update their weights. A typical feed-forward ANN with delta learning rule is shown in Figure 6.4. The hidden layer and output layer in Figure 6.4 are fully connected layers. In a fully connected layer each neuron/synapse is connected to every neuron/synapse in the previous layer, and each connection has its own weight. This is a general purpose connection pattern and makes no assumptions about the features in the data. There can be multiple hidden layers

between input layer and the output layer, some ANNs utilize multiple hidden layers such as convolutional neural network(CNN)[102]. The connection between hidden layer in CNN is not fully connected, instead, it's called convolutional connection. In a convolutional layer each synapse is only connected to a few nearby synapses in the previous layer, and the same set of weights is used for every synapse. This connection pattern only makes sense in cases where the data can be interpreted with the features to be extracted being local and equally likely to occur at any input position, such as human face image recognition and human voice recognition. Generally, ANNs are useful for both supervised learning and unsupervised learning tasks. In the following sections, some simple classification tasks with ANN will be presented.

## 6.4 Device Fabrication

Crossbar arrays were used for the implementation of RRAM based ANN. The bottom metal lines and contact pads were patterned by electron beam (E-beam) lithography first then followed by E-beam evaporation. The metal line width is between 100 nm to 300 nm. Contact pad size was 100  $\mu\text{m}$  by 100  $\mu\text{m}$ . After forming bottom metal lines,  $\text{CeO}_x$  and  $\text{HfO}_x$  were deposited by PVD and ALD. The deposition procedure is similar to bilayer RRAM devices in Chapter 3. Lastly, top metal lines and contact pads were patterned by E-beam lithography and followed by E-beam evaporation again. The geometry of top layer was the same as the bottom layer. Figure 6.5 shows a 10  $\times$  10 crossbar array under optical microscope and scanning electron microscope

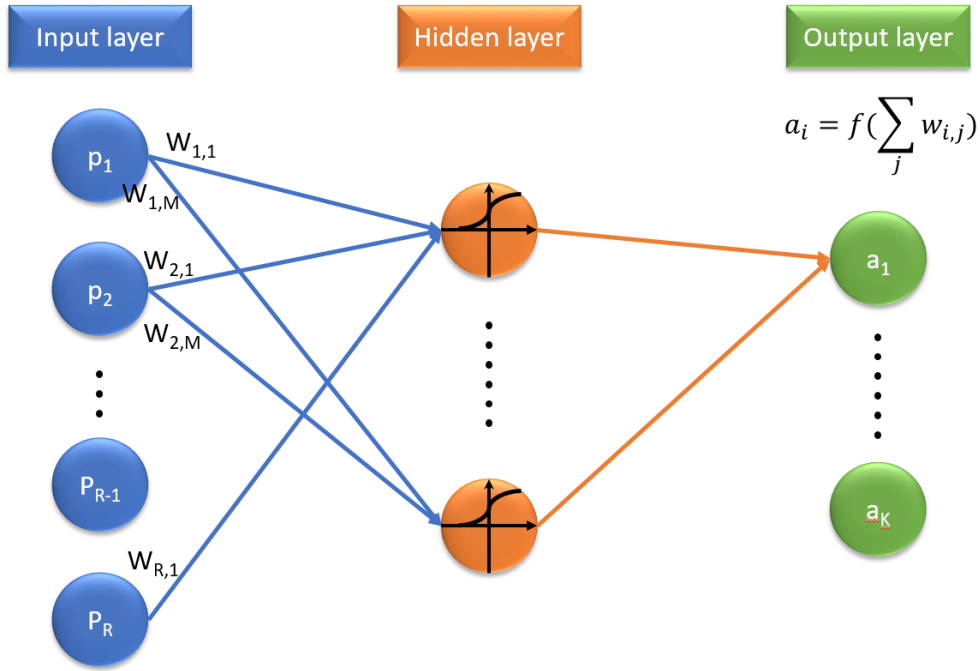


Figure 6.4: Schematics representation of ANN. The activation function of synapse in the hidden layer is a delta function. The total number of input neurons is  $R$ . Total number of synapse in the hidden layer is  $M$ .

(SEM).

## 6.5 Single Hidden Layer ANN for Simple Image Recognition

This section demonstrates a pair of  $4 \times 4$   $\text{CeO}_x/\text{HfO}_x$  bilayer RRAM crossbar arrays for simple image recognition. In this demonstration, the  $4 \times 4$  neural network classifies a  $4 \times 4$  binary image of letter "4". Figure 6.6 shows the training and Figure 6.7 shows testing procedure of the image recognition. All the training and testing images are binary pixel images. The

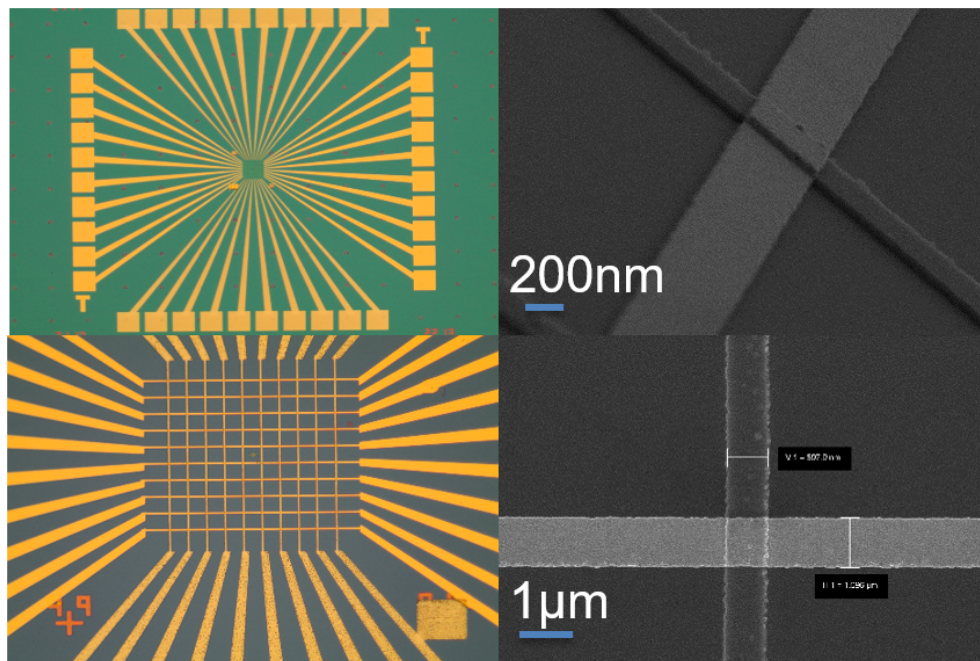


Figure 6.5: Optical microscope images of a  $10 \times 10$  RRAM array are at left and 30 degree tilted and top view SEM images are at right.

classifier implemented here is a binary classifier, which only predicts whether the testing image is "4" or not. The training data set was created by one-bit deviation from the ideal "4" letter. Due to limited probes in probe station, the training data was encoded by one-hot encoding, so that the encoded data can be used for training RRAM synapses. One neural network was trained with red pixels and the complemented one was trained with blue pixels. The voltage applied on each hot synapse was 1.3 V. The signal pulse width was 100 ns. During testing stage, both neural network were fed with the same input signals for each column, which is different from the training stage. The readout current from each corresponding pair was then fed into a comparator.

If any pair has the higher current from inverted array than the current from non-inverted array, then the comparator outputs "No". On the other hand, if the current from non-inverted array is higher than the current from inverted array, then the comparator outputs high, which means "Yes".

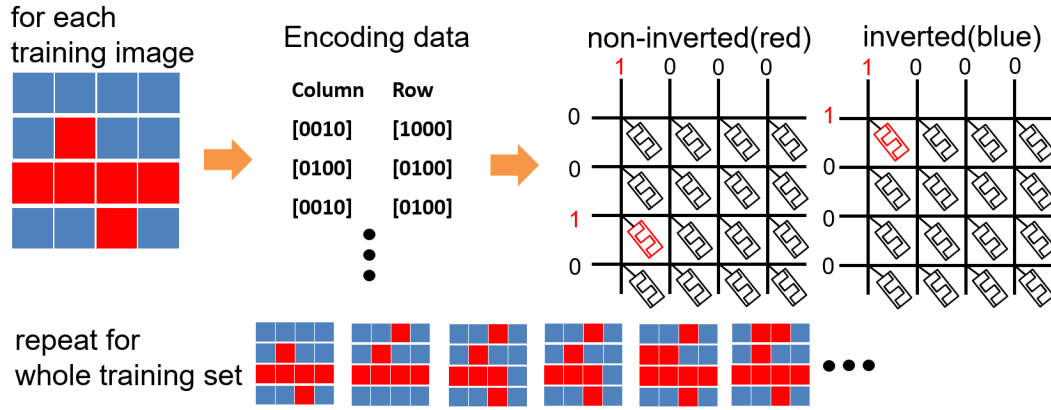


Figure 6.6: The input training data was encoded to input signals for training neural networks. Red pixels in training data were converted to hot for non-inverted neural network and blue pixels were converted to hot for inverted neural network.

Figure 6.8(a) shows the relation between conductance change of the bilayer  $\text{HfO}_x/\text{CeO}_x$  RRAM and number of pulses applied on it. It can be seen that the transition of conductance change is nonlinear, which is common among conductive filament type RRAMs[2, 98, 103]. Ideally the conductance change is linear with number of pulses so there will be maximum number of resistance states. The more resistance states, the better prediction accuracy. Besides, device-to-device variation is also observed in Figure 6.8(b). Figure 6.8(c) shows the prediction accuracy at different pulse voltage during training from two RRAM ANN classifiers. The pulse voltage dependence in Figure

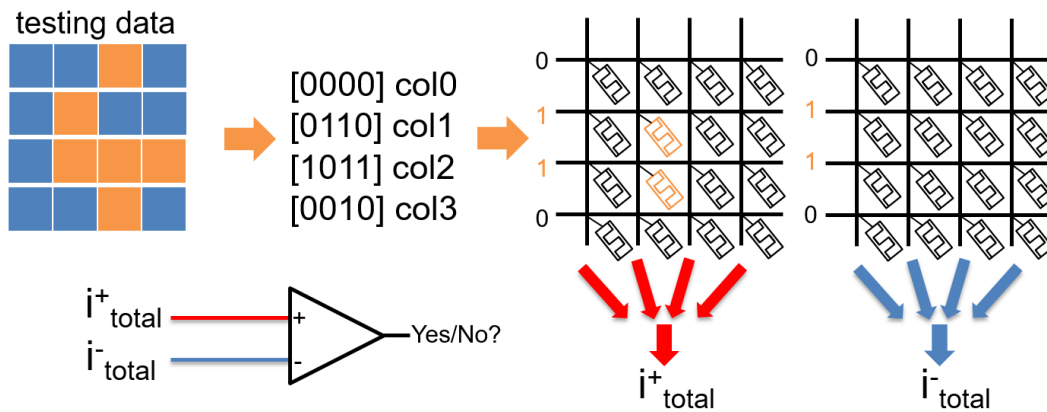


Figure 6.7: Testing data was encoded into input signals for each column, then passing to all columns in each neural network. The result is based on comparison of current between non-inverted array and inverted array.

6.8(c) indicates that larger pulse voltage would shorten the required training period; however, it might result synapse weight update overshoot for multi-level pixel pattern recognition. Although two ANN classifiers have different magnitude of device-to-device variation, the prediction accuracy has similar trend along with the pulse voltage. This is interesting because this process related variation is inevitable, but it doesn't impact the prediction accuracy when classifying binary pixel images. To study deeper into device-to-device variation, TensorFlow simulation was performed to show dependence of device-to-device variation on prediction accuracy. It can be seen that in Figure 6.9(b), the prediction accuracy has low dependence on non-linearity for binary pixel image recognition. Figure 6.9(c) shows relation between the magnitude of non-linearity and prediction accuracy. Both binary pixel and gray level pixel image recognition show similar trend, and gray level pixel images are more

prone to non-linearity of weight change.

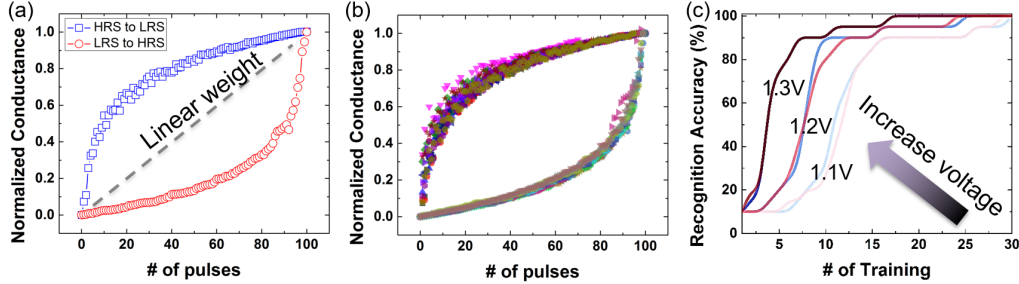


Figure 6.8: (a) represent a RRAM device with nonlinear weight change function versus in a crossbar array. The dashed line in the middle represents the ideal linear weight change. (b) shows the device-to-device variation from 10 RRAM devices in the same array. (c) The prediction accuracy at different pulse voltage during the training. Note that there are two color groups in (c). The red one and blue one are results from different RRAM arrays. Note that the prediction accuracy is precision. Precision in statistics is defined by the ratio of true positives over true positives plus false positives.

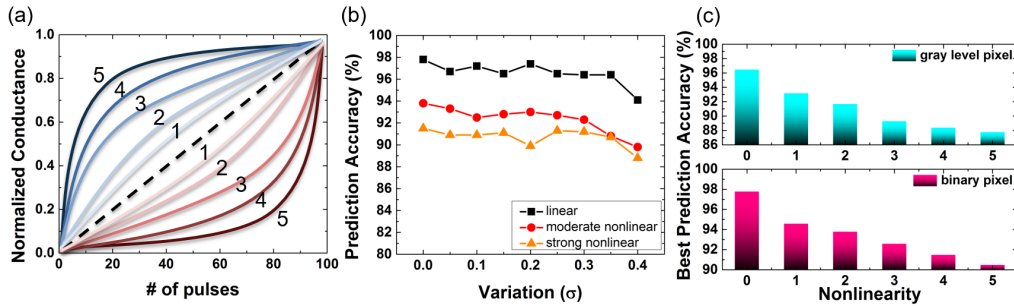


Figure 6.9: (a) represents different magnitude of non-linearity used in TensorFlow simulation for (c). (b) shows the result by TensorFlow simulation of device-to-device variation on prediction accuracy. Moderate nonlinear weight change is 2 in (a) and strong nonlinear weight change is 4 in (a). (c) TensorFlow simulation of non-linearity on prediction accuracy. Note that the simulation assumes no device variation on RRAM synapses.



## 6.6 Multiclass Image Recognition

Binary classification has been demonstrated in the previous section, and this section demonstrates multiclass classification. The ANN classifier learns how to classify three different letters: "T", "H" and "L" and the training data set is shown in Figure 6.10. Overall, the training scheme for multiclass classification is similar to binary classification except each training image only trains the assigned columns instead of the whole array. In other words, this ANN multiclass classifier comprises three binary classifier, and the prediction decision is based on prediction from each binary classifier and compares one to the rest of all. This is called "One-vs-All" method. The RRAM array which can classify  $N/2$  classes of  $M$  pixels image is shown in Figure 6.11. In Figure 6.11, the training image is encoded into signals and both non-inverted and inverted signals pass to assigned column for letter "T". The testing scheme is different from binary classification in the previous section since there are multiple classes can be the answer. The testing scheme is shown in Figure 6.12. The testing image is encoded into non-inverted signals and pass to all columns. The prediction is based on the highest current from each binary classifier.

The experimental result is shown in Figure 6.13. It can be seen that the prediction accuracy of each letter has similar dependence on the number of training set. In the testing data set, two-pixel deviated images are included for "H" and "L", and these ambiguous images causes the prediction error for recognizing "H" and "L" in Figure 6.13. However, there are only ideal and

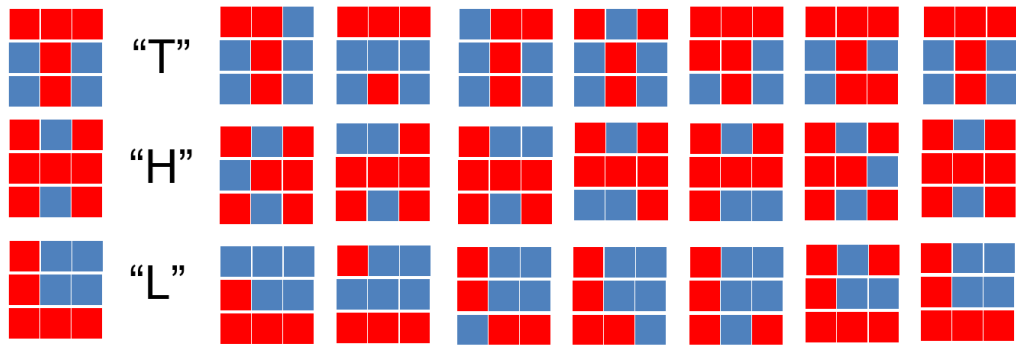


Figure 6.10: Training images used for multiclass image recognition. The left images are ideal patterns for each class.

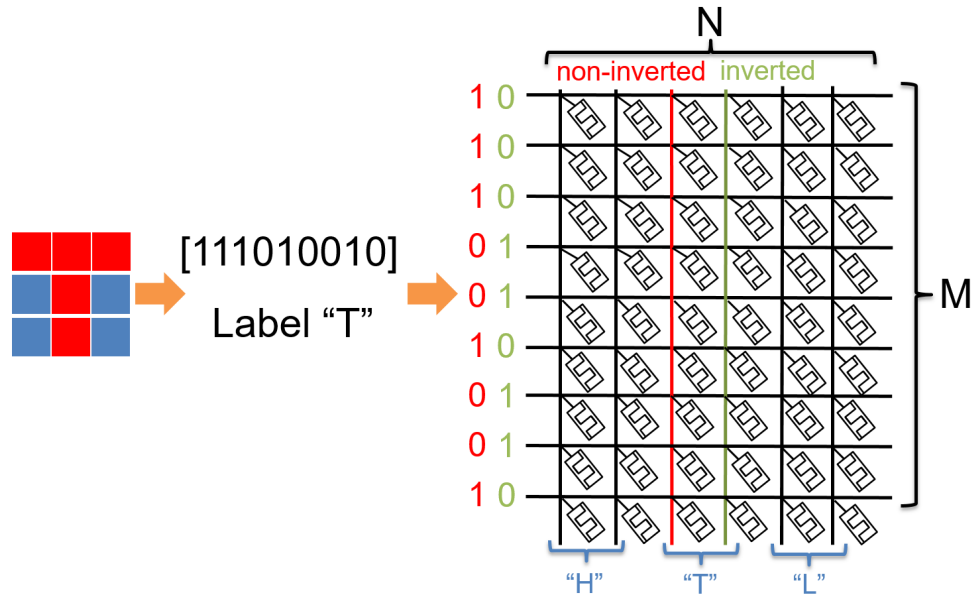


Figure 6.11: Training scheme for multiclass image recognition. The input image encoded into signals and pass to the assigned columns in RRAM array. One column is trained with non-inverted signals, and the other is trained with inverted signals.

one-pixel deviated images for "T", thus the prediction accuracy can reach to 100%. To improve the prediction accuracy, multiple layer ANN is helpful when

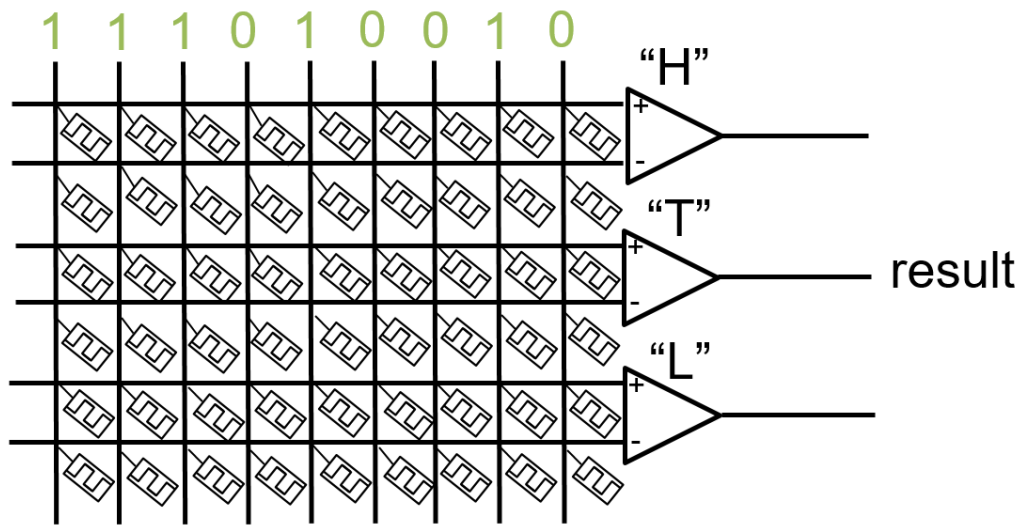


Figure 6.12: Testing scheme for multiclass image recognition. The non-inverted input passed to all columns. The prediction result is based on "One-vs-All" method, *i.e.*, each binary classifier provides its prediction and the final result is the one with maximum current.

the image is complicated or the image has local features.

## 6.7 Simple Convolutional Neural Network for Image Recognition

In the above sections, single hidden layer ANNs have been demonstrated for binary and multiclass classification. In this section, a simple convolutional neural network(CNN) is proposed for binary classification. A CNN comprises one or more convolutional layers, and then followed by one or more fully connected layers as in a conventional multilayer neural network[104]. Figure 6.14 shows the typical architecture of CNN. Key operations of CNN are: convolution, pooling and backpropagation. The purpose of convolution is to

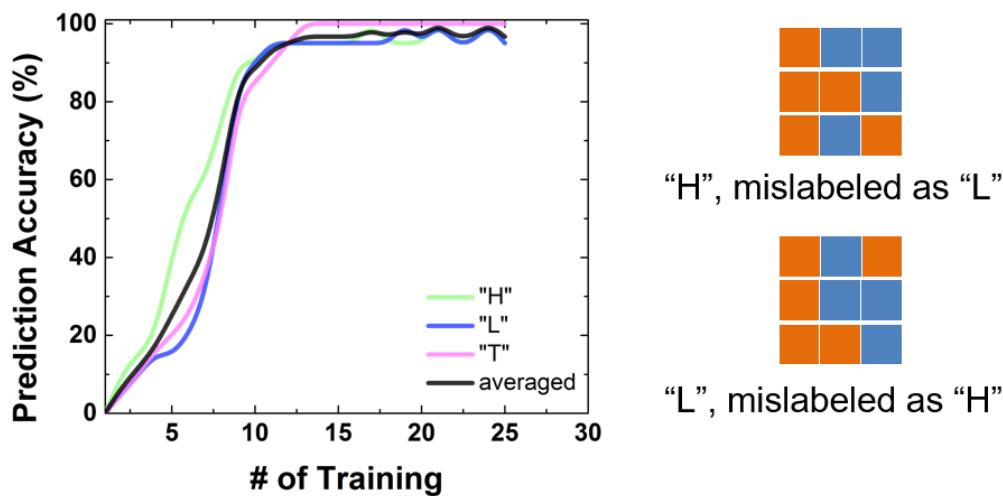


Figure 6.13: Prediction accuracy versus number of training data set. Note that the prediction accuracy of each letter is based on definition of precision. And averaged accuracy is based on marco-average of precision.

extract features presented locally in the image. Figure 6.15 shows the procedure of convolution; a filter matrix is used to extract local features from the image. In Figure 6.15, the filter matrix is for edge detection. The selection of size of filter matrix depends on how local features present in the image. The difference between a fully connected layer and a convolutional layer is that, in a fully connected layer, each synapse receives signals from all synapses/neurons in the previous layer, while in a convolutional layer, each synapse only receives signals from a small portion of synapses/neurons in the previous layer.

Pooling process is illustrated in Figure 6.16. Pooling is important in CNN because it reduces the dimensionality of each feature map but retains the crucial information. Synapses in the feature layer are divided into several pools with size of  $2 \times 2$ . Pooling process chooses either the maximum weight,

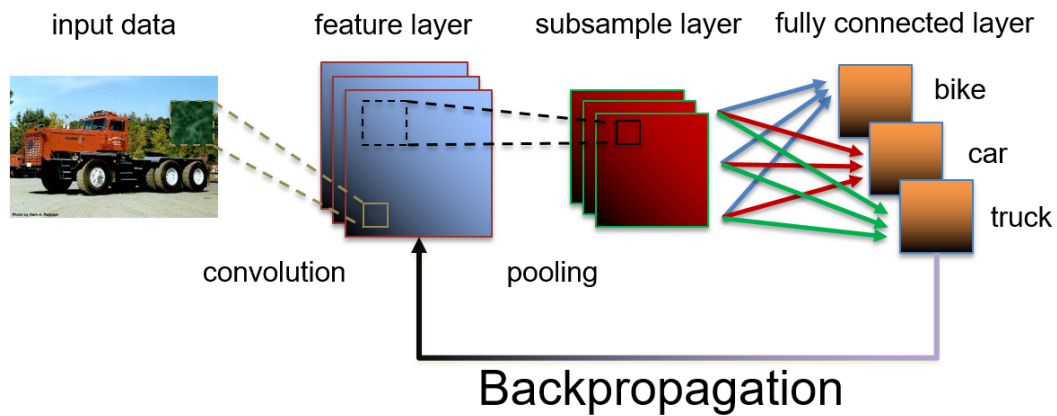


Figure 6.14: A Simple CNN architecture with depth of three. The input data is mapped to the feature layer with convolution by filter matrix. Then the feature layer is mapped to subsample layer by pooling. Finally, the classification is performed by a fully connected layer. The depth in CNN means the number of feature layers used for convolution. Each feature layer might have different filter matrix.

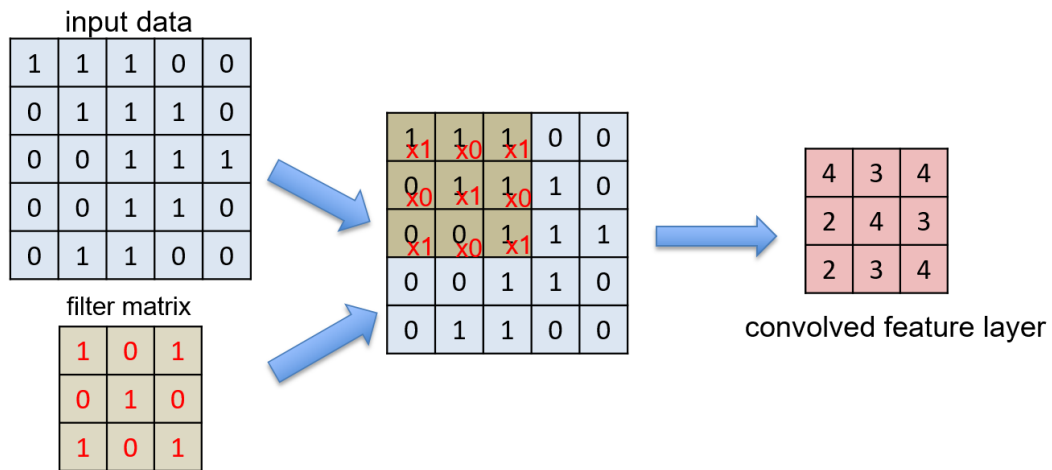


Figure 6.15: The input data is convoluted with the filter matrix and mapped to feature layer. Here, a 3×3 filter matrix is used for edge detection. The 3×3 pink matrix represents the feature layer.

the averaged weight, or the summed weight, then maps selected weights to the subsample layer. In Figure 6.16, the maximum weight is chosen for pooling process. Synapses in the subsample layer finally connects to the fully connected layer. Synapses in the fully connected layer take all inputs from the subsample layer with the summation function. In Figure 6.16, red arrows present negative weights and blue arrows present positive weights. The classification result is based on outputs from the fully connected layer.

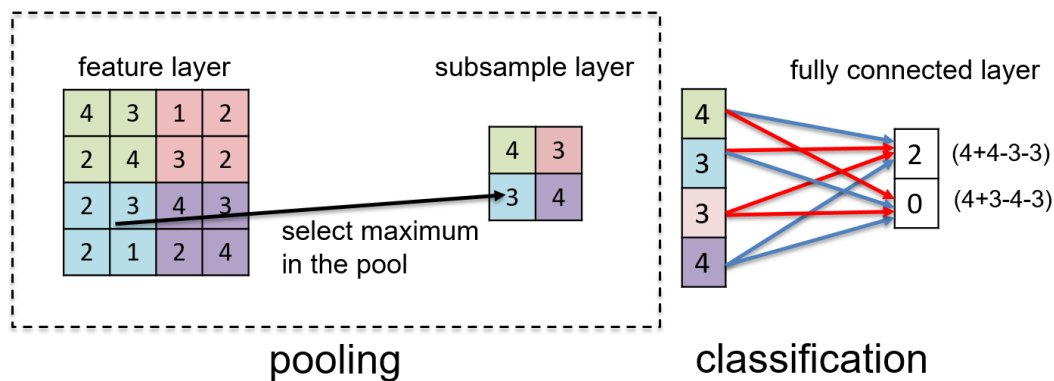


Figure 6.16: The synapses in the feature layer are divided into several pools, and pooling process chooses the maximum weight in each pool then mapping all to a subsample layer. The pool size is  $2 \times 2$  and the size of feature layer is  $4 \times 4$ , therefore, the subsample layer is  $2 \times 2$ . Synapses in the subsample layer connects to a fully connected layer for classification. Synapses in fully connected layer sums up weights with blue arrows minus weights with red arrows. The fully connected layer provides classification result.

The filter matrix and weights in the fully connected layer in this section is shown in Figure 6.17. Generally, these parameters are determined by back-propagation. Backpropagation is a common algorithm to train ANNs together with some optimization methods such as gradient descent[105]. The algorithm

basically consists of two phase cycle, propagation and weight update. While training an ANN, the input data propagated through whole layers and generate prediction output. The output is then compared to the label (answer), then generate an error value. This is called propagation. The second phase is to use the error value to calculate the gradient of the loss function. This gradient can be used to update the parameters such as filter matrix and summation function in the fully connected layer. To simplify the experiment complexity, the filter matrix and weights in the fully connected layer is adapted from TensorFlow simulation with similar CNN architecture and training data set. Figure 6.18 and 6.19 shows the training and testing scheme for CNN image recognition. In this experiment,  $10 \times 10$  binary pixel images of letter "H" and "L" with ambiguous images presented in training data sets. Two  $8 \times 8$  RRAM arrays were used to update weights in the feature layer for "H" and "L", respectively. The pooling layer and the fully connected layer are combined for comparing output current to provide the prediction result. There are two testing data sets used in the experiment: One is the normal test set, which contains only ideal "H", "L" images and flipped pixels images with maximum 3 flipped pixels. The other one is the difficult test set, which contains 50% ambiguous images and 50% ideal and flipped pixels images with maximum 3 flipped pixels. Figure 6.20 shows testing results on both cases between single hidden layer ANN and CNN. It can be seen that in Figure 6.20(a), both ANN and CNN can achieve high accuracy given enough training images. However, CNN outperforms ANN in Figure 6.20(b) for the difficult data set. Figure

6.20(b) demonstrates the power of CNN, even the architecture of the CNN is simple, it still can improve the prediction accuracy significantly. Figure 6.21(a) shows prediction accuracy of CNNs with different filter matrices. The result in Figure 6.21(a) shows that the prediction accuracy can change significantly by simply changing filter matrix. This implies that the extraction of important features is important for achieving high accuracy when using CNN algorithms. Identity filter matrix in the CNN removes features at the edge of the image, and thus improve the prediction accuracy. Figure 6.21(b) illustrates reason of the lower prediction accuracy of using sharpen filter matrix in Figure 6.21(a), that is, the failure of predicting "H" results the low prediction accuracy.

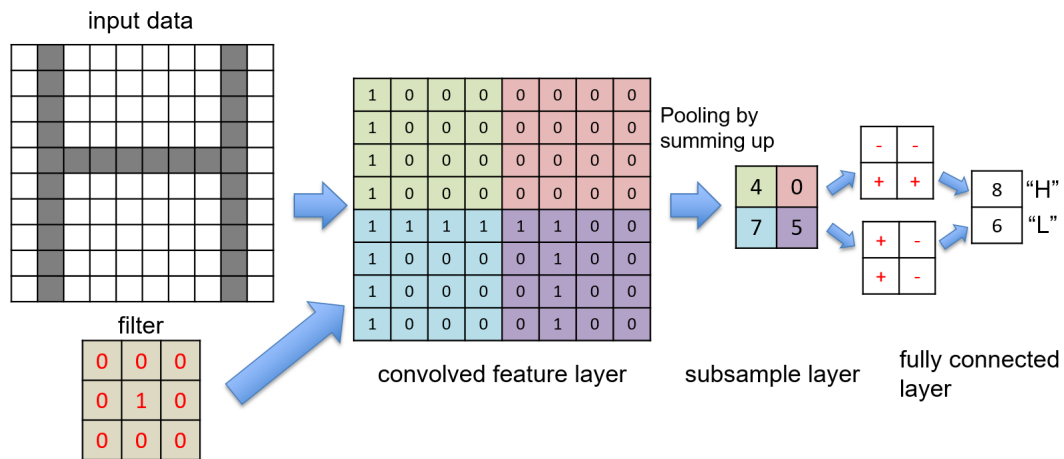


Figure 6.17: The filter matrix and connection between the subsample layer and the fully connected layer is adapted from the result by TensorFlow simulation.

In the demonstration, RRAM arrays are used for storing and updating weights in the feature layer. Using RRAM arrays can be beneficial for implementing hardware neuromorphic systems because RRAM arrays can be



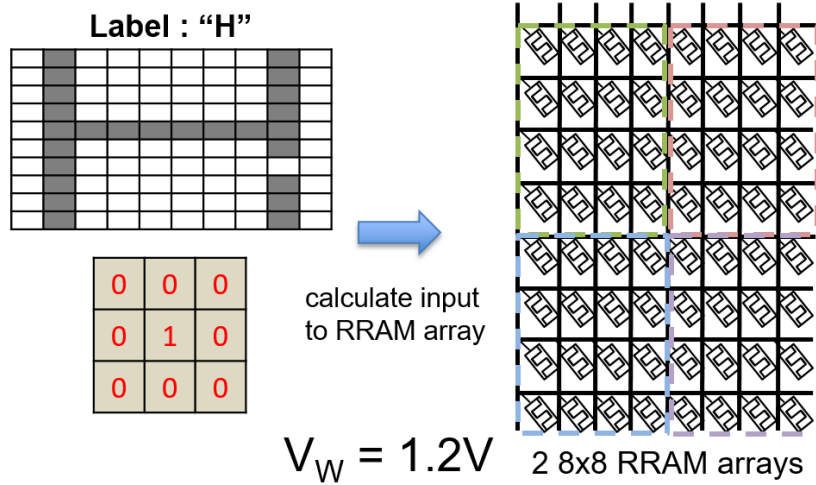


Figure 6.18: Training scheme of the CNN. Note that during the training, the write voltage is always 1.2V and no connection to the subsample layer.

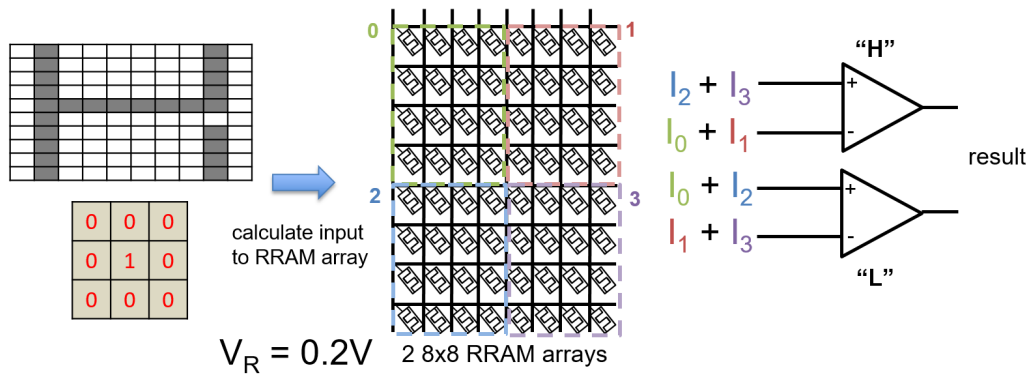


Figure 6.19: Testing scheme of the CNN. The read voltage is 0.2V and the output is based on the comparison of current from each subsample layer.

integrated into BEOL CMOS processes, which means tremendous reduction in delay between logic and memory. Comparing to conventional NVM technologies, RRAM consumes much lower energy and has higher writing speed. This is desired for low-power System on a Chip (SoC) with pattern recognition

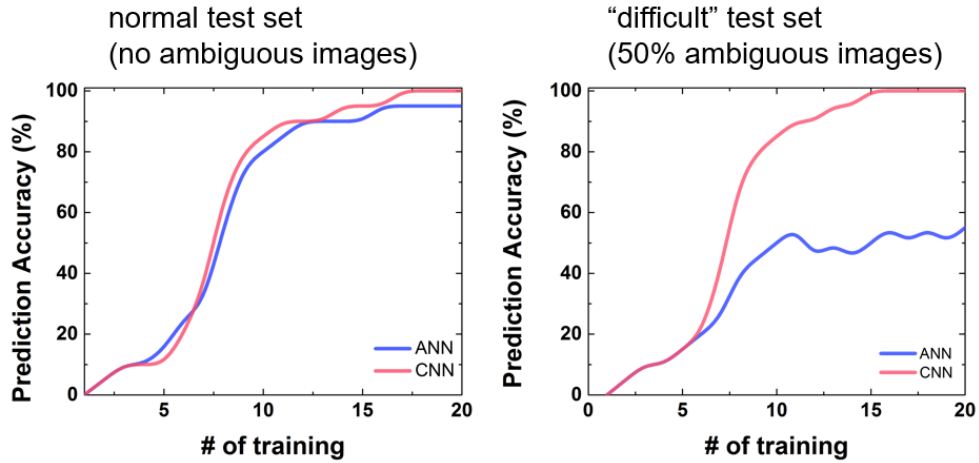


Figure 6.20: (a) Prediction accuracy of regular testing data set between single hidden layer ANN and CNN. (b) Prediction accuracy of 50% ambiguous testing data set between single hidden layer ANN and CNN.

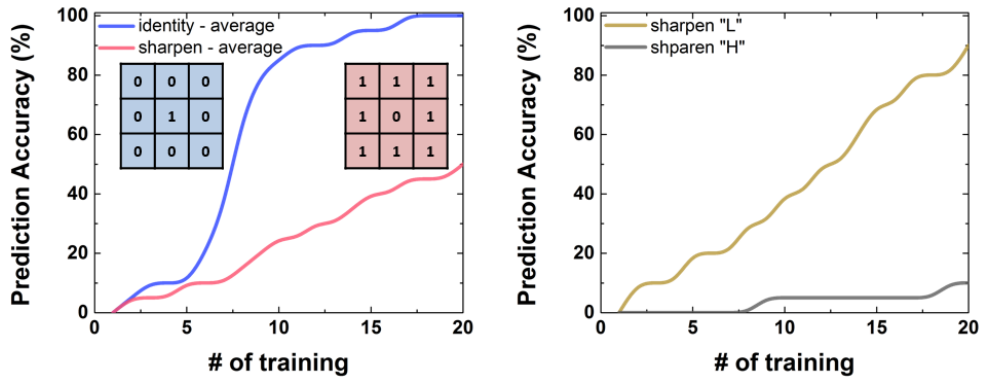


Figure 6.21: (a) Prediction accuracy of CNN with different filter matrix for ambiguous testing data set. The filter matrix is also shown in (a), the one at left is identity filter and the one at right is sharpen filter. (b) Prediction accuracy of each class for CNN with sharpen filter.

tasks.

## **6.8 Conclusion**

This chapter demonstrates RRAM arrays for image recognition by realizing ANN algorithms. Both binary and multiclass classification with single hidden layer ANN are demonstrated. Furthermore, a simple CNN algorithm has been demonstrated with RRAM arrays and it shows improved accuracy for recognizing ambiguous images compared to single hidden layer ANN. This also indicates that using RRAM arrays to store data in feature layers would be energy-saving and faster than pure CMOS approach.

# Chapter 7

## Conclusions

### 7.1 Summary

The performance requirements for NVM in Big Data era are increasing, thus a new NVM technology is needed for developing future data center and mobile devices. RRAM is one of potential candidates and it has several advantages over CMOS based NAND flash, such as switching speed, switching energy, area density and etc.. This thesis explores the feasibility of  $\text{CeO}_x$  for RRAM device and the potential of using  $\text{CeO}_x$  RRAM as synaptic device for neuromorphic computing. The contributions of this thesis include:

1. The fabrication process for  $\text{CeO}_x$  RRAM device. MBE and PVD thin film growth method has been studied. wet etching method  $\text{CeO}_x$  has been demonstrated. The basic performance parameters of  $\text{CeO}_x$  from different growth method has been studied, such as operating voltage, on/off window, data retention and endurance.

2. Study of switching mechanism in  $\text{CeO}_x$  RRAM. This thesis addresses the switching mechanism of  $\text{CeO}_x$  by analyzing vertical and horizontal scaling of  $\text{CeO}_x$  RRAM. The results indicate that set and reset voltage of  $\text{CeO}_x$  RRAM are independent of vertical and horizontal scaling, and forming voltage is de-

pendent of vertical scaling. It can be concluded that one-dimensional filament type is very likely for resistive switching in  $\text{CeO}_x$  RRAM. TEM cross-sectional images of MBE grown  $\text{CeO}_x$  RRAM reveal  $\text{Ce}_2\text{O}_3$  nanophase in switching layer, and it plays an important role for oxygen vacancies in  $\text{CeO}_x$ .

3. Performance improvement by introducing the bilayer structure. The  $\text{CeO}_x$  RRAM doesn't show analog switching in reset operation. Adding a non-stoichiometric  $\text{HfO}_x$  layer can not only reduce operating voltage by Hf metallic bonds around  $\text{HfO}_x$  and  $\text{CeO}_x$  interface and enhanced joule heating, but also provide analog switching in reset operation. With capability of analog switching in both set and reset,  $\text{HfO}_x/\text{CeO}_x$  bilayer RRAM can emulate synaptic behavior and demonstrate Hebbian learning. This leads to neuromorphic applications in Chapter 6.

4. Chapter 4 demonstrates highly nonlinear MSM diode as selector device and propose I-V characteristics model of it. The Ti-a-Si-Ti MSM diode with Schottky emission provides very high nonlinear ratio and low turn-on voltage, which can help eliminate sneak path leakage in RRAM array and MSM diode with over  $10^5$  nonlinear ratio is capable for building gigabyte (GB) crossbar array. The asymmetric I-V characteristics can be modeled by different Schottky barriers, which indicates that top electrode interface and bottom electrode interface can cause asymmetry I-V curves.

5. Study of short-term relaxation of 1R and 1S1R structure with adaptive programming algorithms. Chapter 5 observes the different magnitude of short-term relaxation between 1R and 1S1R, and proposes the root cause the

worse relaxation effect in 1S1R. A potential solution is provided with long forming pulses, which reduces the short-term relaxation with sacrificing long term reliability.

6. Demonstration of pattern recognition with RRAM array as artificial neural network. Training and testing scheme for binary/multiclass classification are proposed and demonstrated. For large scale pattern recognition, peripheral RRAM array circuits are proposed and circuit level simulation is demonstrated for single layer neural network for pattern recognition.

## 7.2 Recommendations for Future Work

It is a great time to be researching emerging memory technology like RRAM. The potential replacement by novel NVM technologies is exciting and it is happening. These novel technologies could be revolutionary in memory system design. The next big thing comes after Big Data will be artificial intelligence and neuromorphic applications. RRAM is a great tool to implement brain-liked computing hardware.

The above works outlined a comprehensive study of  $\text{CeO}_x$  which provides a solid foundation for more exciting future work in several directions. The bilayer study of  $\text{HfO}_x/\text{CeO}_x$  shed light on materials engineering for RRAM devices. There are several combination of materials worth to try out. Not only switching layers, even different electrode materials can change switching behavior. There are a lot of room for research.

For selector device, novel low-dimensional materials such as  $\text{MoSe}_2$ , could be potential material for MSM diode selector because of its native thin film properties and lower bandgap. The ultra low turn-on voltage would be desirable when scaling RRAM arrays.

With respect to neuromorphic applications, the learning algorithms applied on crossbar RRAM arrays have been tested and verified. The next step can be tape out circuit level design and validate functionality of the chip. Multilayer ANN with System on a Chip(SoC) applications is very attractive for Internet of Things(IoT) devices.

# Index

*Selector Device for RRAM*, 45

*Short Term Relaxation of RRAM*, 65

Abstract, ix

*Acknowledgments*, v

*Bibliography*, 120

*Bilayer Cerium Oxide RRAM*, 27

*Characterization of Cerium Oxide RRAM*,

13

*Conclusions*, 97

*Dedication*, iv

*Electrical and Physical Characteriza-*

*tion*, 15

*Fabrication*, 14

*Hafnium Oxide and Cerium Oxide Stacked*

*Resistive Random Access Mem-*

*ory*, 30

*Introduction*, 1, 13

*Introduction*, 27

101

*Neuromorphic Applications*, 75



## Bibliography

- [1] H.-S.P. Wong, S. Raoux, SangBum Kim, Jiale Liang, John P. Reifenberg, B. Rajendran, Mehdi Asheghi, and Kenneth E. Goodson. Phase Change Memory. *Proc. IEEE*, 98(12):2201–2227, 2010.
- [2] Sung Hyun Jo, Ting Chang, Idongesit Ebong, Bhavitavya B. Bhadviya, Pinaki Mazumder, and Wei Lu. Nanoscale Memristor Device as Synapse in Neuromorphic Systems. *Nano Lett.*, 10(4):1297–1301, 2010.
- [3] Leon Chua. If its pinched its a memristor. *Semiconductor Science and Technology*, 29(10):104001, 2014.
- [4] J.N. Greeley and J.A. Smythe. Resistive ram devices and methods, May 27 2014. US Patent 8,735,211.
- [5] E.V. Karpov, B.S. Doyle, C.C. Kuo, R.S. Chau, E.R. Dickey, M.S. Bowen, and S.S. Sun. Low voltage embedded memory having conductive oxide and electrode stacks, January 5 2016. US Patent 9,231,204.
- [6] R. Pillarisetty, P. Majhi, U. Shah, N. Mukherjee, E.V. Karpov, B.S. Doyle, and R.S. Chau. Techniques for filament localization, edge effect reduction, and forming/switching voltage reduction in rram devices, December 29 2016. US Patent App. 14/752,934.

- [7] A. Foong and F. Hady. Storage as fast as rest of the system. In *2016 IEEE 8th International Memory Workshop (IMW)*, pages 1–4, May 2016.
- [8] In-Sung Park, Kyong-Rae Kim, Sangsul Lee, and Jinho Ahn. Resistance Switching Characteristics for Nonvolatile Memory Operation of Binary Metal Oxides. *Jpn. J. Appl. Phys.*, 46(4B):2172–2174, apr 2007.
- [9] Hiroyuki Akinaga and Hisashi Shima. Resistive Random Access Memory (ReRAM) Based on Metal Oxides. *Proc. IEEE*, 98(12):2237–2251, dec 2010.
- [10] Eike Linn, Roland Rosezin, Carsten Kügeler, and Rainer Waser. Complementary resistive switches for passive nanocrossbar memories. *Nat. Mater.*, 9(5):403–6, may 2010.
- [11] T. W. Hickmott. Lowfrequency negative resistance in thin anodic oxide films. *Journal of Applied Physics*, 33(9):2669–2682, 1962.
- [12] Yong-Mu Kim and Jang-Sik Lee. Reproducible resistance switching characteristics of hafnium oxide-based nonvolatile memory devices. *Journal of Applied Physics*, 104(11), 2008.
- [13] Wen-Yuan Chang, Yen-Chao Lai, Tai-Bor Wu, Sea-Fue Wang, Frederick Chen, and Ming-Jinn Tsai. Unipolar resistive switching characteristics of zno thin films for nonvolatile memory applications. *Applied Physics Letters*, 92(2), 2008.

- [14] Chih-Yang Lin, Sheng-Yi Wang, Dai-Ying Lee, and Tseung-Yuen Tseng. Electrical properties and fatigue behaviors of zro2 resistive switching thin films. *Journal of The Electrochemical Society*, 155(8):H615–H619, 2008.
- [15] Z. Wei, Y. Kanzawa, K. Arita, Y. Katoh, K. Kawai, S. Muraoka, S. Mitani, S. Fujii, K. Katayama, M. Iijima, T. Mikawa, T. Ninomiya, R. Miyanaga, Y. Kawashima, K. Tsuji, A. Himeno, T. Okada, R. Azuma, K. Shimakawa, H. Sugaya, T. Takagi, R. Yasuhara, K. Horiba, H. Kumigashira, and M. Oshima. Highly reliable TaO<sub>x</sub>/ReRAM and direct evidence of redox reaction mechanism. *2008 IEEE Int. Electron Devices Meet.*, pages 1–4, 2008.
- [16] Qi Liu, Weihua Guan, Shibing Long, Rui Jia, Ming Liu, and Junning Chen. Resistive switching memory effect of zro2 films with zr+ implanted. *Applied Physics Letters*, 92(1), 2008.
- [17] H. Y. Lee, P. S. Chen, T. Y. Wu, Y. S. Chen, F. Chen, C. C. Wang, P. J. Tzeng, C. H. Lin, M. J. Tsai, and C. Lien. Hfo<sub>x</sub> bipolar resistive memory with robust endurance using alcu as buffer electrode. *IEEE Electron Device Letters*, 30(7):703–705, July 2009.
- [18] Yao-feng Chang, Pai-yu Chen, Burt Fowler, Yen-ting Chen, Fei Xue, Yanzhen Wang, Fei Zhou, and C Jack. Understanding the resistive switching characteristics and mechanism in active SiO<sub>x</sub>-based resistive switching memory. *J. Appl. Phys.*, 112(May 2014):123702, 2013.

- [19] Yao-Feng Chang, Burt Fowler, Ying-Chen Chen, Yen-Ting Chen, Yanzhen Wang, Fei Xue, Fei Zhou, and Jack C. Lee. Intrinsic SiO<sub>x</sub>-based unipolar resistive switching memory. II. Thermal effects on charge transport and characterization of multilevel programming. *J. Appl. Phys.*, 116(4):043709, 2014.
- [20] Adnan Younis, Dewei Chu, and Sean Li. Oxygen level: the dominant of resistive switching characteristics in cerium oxide thin films. *J. Phys. D. Appl. Phys.*, 45(35):355101, sep 2012.
- [21] I. G. Baek, D. C. Kim, M. J. Lee, H. J. Kim, E. K. Yim, M. S. Lee, J. E. Lee, S. E. Ahn, S. Seo, J. H. Lee, J. C. Park, Y. K. Cha, S. O. Park, H. S. Kim, I. K. Yoo, U. Chung, J. T. Moon, and B. I. Ryu. Multi-layer cross-point binary oxide resistive memory (oxrram) for post-nand storage application. In *IEEE International Electron Devices Meeting, 2005. IEDM Technical Digest.*, pages 750–753, Dec 2005.
- [22] Rohit S Shenoy, Geoffrey W Burr, Kumar Virwani, Bryan Jackson, Alvaro Padilla, Pritish Narayanan, Charles T Rettner, Robert M Shelby, Donald S Bethune, Karthik V Raman, Matthew BrightSky, Eric Joseph, Philip M Rice, Teya Topuria, Andrew J Kellock, Blent Kurdi, and Kailash Gopalakrishnan. Miec (mixed-ionic-electronic-conduction)-based access devices for non-volatile crossbar memory arrays. *Semiconductor Science and Technology*, 29(10):104005, 2014.

- [23] M.-J. Lee, Y. Park, D.-S. Suh, E.-H. Lee, S. Seo, D.-C. Kim, R. Jung, B.-S. Kang, S.-E. Ahn, C.B. Lee, D.H. Seo, Y.-K. Cha, I.-K. Yoo, J.-S. Kim, and B.H. Park. Two series oxide resistors applicable to high speed and high density nonvolatile memory. *Advanced Materials*, 19(22):3919–3923, 2007.
- [24] DerChang Kau, S. Tang, I. V. Karpov, R. Dodge, B. Klehn, J. A. Kalb, J. Strand, A. Diaz, N. Leung, J. Wu, Sean Lee, T. Langtry, Kuo wei Chang, C. Papagianni, Jinwook Lee, J. Hirst, S. Erra, E. Flores, N. Righos, H. Castro, and G. Spadini. A stackable cross point phase change memory. In *2009 IEEE International Electron Devices Meeting (IEDM)*, pages 1–4, Dec 2009.
- [25] L. Zhang, A. Redolfi, C. Adelman, S. Clima, I. P. Radu, Y. Y. Chen, D. J. Wouters, G. Groeseneken, M. Jurczak, and B. Govoreanu. Ultrathin metal/amorphous-silicon/metal diode for bipolar rram selector applications. *IEEE Electron Device Letters*, 35(2):199–201, Feb 2014.
- [26] L. Chua. Memristor-the missing circuit element. *IEEE Transactions on Circuit Theory*, 18(5):507–519, Sep 1971.
- [27] Muhammad Ismail, Ijaz Talib, Chun-Yang Huang, Chung-Jung Hung, Tsung-Ling Tsai, Jheng-Hong Jieng, Umesh Chand, Chun-An Lin, Ejaz Ahmed, Anwar Manzoor Rana, Muhammad Younus Nadeem, and Tseung-Yuen Tseng. Resistive switching characteristics of pt/ceo x /tin

- memory device. *Japanese Journal of Applied Physics*, 53(6):060303, 2014.
- [28] Sambhaji S. Warule, Nilima S. Chaudhari, Bharat B. Kale, Kashinath R. Patil, Pankaj M. Koinkar, Mahendra A. More, and Ri-ichi Murakami. Organization of cubic ceo2 nanoparticles on the edges of self assembled tapered zno nanorods via a template free one-pot synthesis: significant cathodoluminescence and field emission properties. *J. Mater. Chem.*, 22:8887–8895, 2012.
- [29] Michiko Yoshitake, Michal Vaclavu, Mykhailo Chundak, Vladimir Matolin, and Toyohiro Chikyow. Epitaxial CeO<sub>2</sub> thin films for a mechanism study of resistive random access memory (ReRAM). *J. Solid State Electrochem.*, 17(12):3137–3144, aug 2013.
- [30] a. Kossoy, M. Greenberg, K. Gartsman, and I. Lubomirsky. Chemical Reduction and Wet Etching of CeO<sub>2</sub> Thin Films. *J. Electrochem. Soc.*, 152(2):C65, 2005.
- [31] Rickard Fors, Sergey Khartsev, and Alexander Grishin. Giant resistance switching in metal-insulator-manganite junctions: Evidence for Mott transition. *Phys. Rev. B*, 71(4):045305, jan 2005.
- [32] L. Goux, Y.-Y Chen, L. Pantisano, X.-P. Wang, G. Groeseneken, M. Jurczak, and D. J. Wouters. On the Gradual Unipolar and Bipolar Resistive Switching of TiN/HfO<sub>2</sub>/Pt Memory Systems. *Electrochem. Solid-State Lett.*, 13(6):G54, 2010.

- [33] K. Kinoshita, T. Tamura, M. Aoki, Y. Sugiyama, and H. Tanaka. Bias polarity dependent data retention of resistive random access memory consisting of binary transition metal oxide. *Applied Physics Letters*, 89(10), 2006.
- [34] Anupam Roy, Samaresh Guchhait, Sushant Sonde, Rik Dey, Tanmoy Pramanik, Amritesh Rai, Hema C. P. Movva, Luigi Colombo, and Sanjay K. Banerjee. Two-dimensional weak anti-localization in  $\text{Bi}_2\text{Te}_3$  thin film grown on  $\text{Si}(111)-(7\times 7)$  surface by molecular beam epitaxy. *Applied Physics Letters*, 102(16), 2013.
- [35] Tomo Hasegawa, Syed Mohammad Fakruddin Shahed, Yasuyuki Sainoo, Atsushi Beniya, Noritake Isomura, Yoshihide Watanabe, and Tadahiro Komeda. Epitaxial growth of  $\text{CeO}_2(111)$  film on  $\text{Ru}(0001)$ : Scanning tunneling microscopy (STM) and x-ray photoemission spectroscopy (XPS) study. *J. Chem. Phys.*, 140(4):044711, 2014.
- [36] Faal Larachi, Jrme Pierre, Alain Adnot, and Alain Bernis. Ce 3d {XPS} study of composite  $\text{Ce}_x\text{Mn}_y\text{O}_2$  wet oxidation catalysts. *Applied Surface Science*, 195(14):236 – 250, 2002.
- [37] O. Gunnarsson and K. Schönhammer. Electron spectroscopies for ce compounds in the impurity model. *Phys. Rev. B*, 28:4315–4341, Oct 1983.
- [38] Peter Burroughs, Andrew Hamnett, Anthony F. Orchard, and Geoffrey Thornton. Satellite structure in the x-ray photoelectron spectra of some

- binary and mixed oxides of lanthanum and cerium. *J. Chem. Soc., Dalton Trans.*, pages 1686–1698, 1976.
- [39] Yang Yin Chen, Masanori Komura, Robin Degraeve, Bogdan Govoreanu, Ludovic Goux, Andrea Fantini, Naga Raghavan, Sergiu Clima, Leqi Zhang, Attilio Belmonte, Augusto Redolfi, Gouri Sankar Kar, Guido Groeseneken, Dirk J. Wouters, and Malgorzata Jurczak. Improvement of data retention in HfO<sub>2</sub>/Hf 1T1R RRAM cell under low operating current. *Tech. Dig. - Int. Electron Devices Meet. IEDM*, pages 252–255, 2013.
- [40] Dmitri B. Strukov, Gregory S. Snider, Duncan R. Stewart, and R. Stanley Williams. The missing memristor found. *Nature*, 453(7191):80–83, 2008.
- [41] Rainer Waser and Masakazu Aono. Nanoionics-based resistive switching memories. *Nat. Mater.*, 6(11):833–40, nov 2007.
- [42] Duygu Kuzum, Shimeng Yu, and H.-S. Philip Wong. Synaptic electronics: materials, devices and applications. *Nanotechnology*, 24(38):382001, 2013.
- [43] H-S Philip Wong and Sayeef Salahuddin. Memory leads the way to better computing. *Nature nanotechnology*, 10(3):191–194, 2015.
- [44] Timothée Masquelier, Rudy Guyonneau, and Simon J Thorpe. Competi-



- tive STDP-based spike pattern learning. *Neural Comput.*, 21:1259–1276, 2009.
- [45] Shimeng Yu and H. S Philip Wong. A phenomenological model for the reset mechanism of metal oxide RRAM. *IEEE Electron Device Lett.*, 31(12):1455–1457, 2010.
- [46] Mirko Prezioso, Farnood Merrikh-Bayat, BD Hoskins, GC Adam, Konstantin K Likharev, and Dmitri B Strukov. Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature*, 521(7550):61–64, 2015.
- [47] J. Joshua Yang, Dmitri B. Strukov, and Duncan R. Stewart. Memristive devices for computing. *Nat. Nanotechnol.*, 8(1):13–24, 2012.
- [48] Giacomo Indiveri, Bernabé Linares-Barranco, Robert Legenstein, George Deligeorgis, and Themistoklis Prodromakis. Integration of nanoscale memristor synapses in neuromorphic computing architectures. *Nanotechnology*, 24(38):384010, 2013.
- [49] Shimeng Yu, Duygu Kuzum, and H.-S. Philip Wong. Design considerations of synaptic device for neuromorphic computing. *2014 IEEE Int. Symp. Circuits Syst.*, pages 1062–1065, 2014.
- [50] Katsumasa Kamiya, Moon Young Yang, Seong-Geon Park, Blanka Magyari-Kope, Yoshio Nishi, Masaaki Niwa, and Kenji Shiraishi. ON-OFF switching mechanism of resistive random access memories based on

the formation and disruption of oxygen vacancy conducting channels. *Appl. Phys. Lett.*, 100(7):073502, 2012.

- [51] J. Joshua Yang, Matthew D. Pickett, Xuema Li, Douglas A. A. Ohlberg, Duncan R. Stewart, and R. Stanley Williams. Memristive switching mechanism for metal/oxide/metal nanodevices. *Nat. Nanotechnol.*, 3(7):429–433, 2008.
- [52] C. Schindler, G. Staikov, and R. Waser. Electrode kinetics of  $\text{CuSiO}_2$ -based resistive switching cells: Overcoming the voltage-time dilemma of electrochemical metallization memories. *Applied Physics Letters*, 94(7), 2009.
- [53] Shimeng Yu, Hong-Yu Chen, Bin Gao, Jinfeng Kang, and H-S Philip Wong. Hfox-based vertical resistive switching random access memory suitable for bit-cost-effective three-dimensional cross-point architecture. *ACS nano*, 7(3):2320–2325, 2013.
- [54] L. Goux, a. Fantini, G. Kar, Y. Y. Chen, N. Jossart, R. Degraeve, S. Clima, B. Govoreanu, G. Lorenzo, G. Pourtois, D. J. Wouters, J. a. Kittl, L. Altimime, and M. Jurczak. Ultralow sub-500nA operating current high-performance  $\text{TiN}/\text{Al}_2\text{O}_3/\text{HfO}_2/\text{Hf}/\text{TiN}$  bipolar RRAM achieved through understanding-based stack-engineering. *Dig. Tech. Pap. - Symp. VLSI Technol.*, pages 159–160, 2012.
- [55] L. Goux, A. Fantini, A. Redolfi, C. Y. Chen, F. F. Shi, R. Degraeve, Y. Y. Chen, T. Witters, G. Groeseneken, and M. Jurczak. Role of the ta

- scavenger electrode in the excellent switching control and reliability of a scalable low-current operated 1T1R rram device. In *VLSI Technology (VLSI-Technology): Digest of Technical Papers, 2014 Symposium on*, pages 1–2, June 2014.
- [56] Yang Dan and Mu ming Poo. Spike timing-dependent plasticity of neural circuits. *Neuron*, 44(1):23 – 30, 2004.
- [57] Sanghyun Ban and Ohyun Kim. Improvement of switching uniformity in HfO<sub>x</sub>-based resistive random access memory with a titanium film and effects of titanium on resistive switching behaviors. *Jpn. J. Appl. Phys.*, 53, 2014.
- [58] W. L. Scopel, Antônio J R Da Silva, W. Orellana, and a. Fazzio. Comparative study of defect energetics in HfO<sub>2</sub> and SiO<sub>2</sub>. *Appl. Phys. Lett.*, 84(9):1492–1494, 2004.
- [59] Christopher Hardacre, Gerard M. Roe, and Richard M. Lambert. Structure, composition and thermal properties of cerium oxide films on platinum {111}. *Surf. Sci.*, 326(1-2):1–10, 1995.
- [60] J Guy, G Molas, P Blaise, C Carabasse, M Bernard, A Roule, G Le Carval, V Sousa, H Grampeix, V Delaye, A Toffoli, J Cluzel, P Briancaeu, O Pollet, S Barraud, O Cueto, G Ghibaud, F Clermidy, B De Salvo, and L Perniola. Experimental and theoretical understanding of Forming, SET and RESET operations in Conductive Bridge RAM (CBRAM)

for memory stack optimization. *2014 IEEE Int. Electron Devices Meet.*, pages 6.5.1–6.5.4, 2014.

- [61] Marat Khafizov, In Wook Park, Aleksandr Chernatynskiy, Lingfeng He, Jianliang Lin, John J. Moore, David Swank, Thomas Lillo, Simon R. Phillpot, Anter El-Azab, and David H. Hurley. Thermal conductivity in nanocrystalline ceria thin films. *J. Am. Ceram. Soc.*, 97(2):562–569, 2014.
- [62] Stephan Menzel. *Modeling and simulation of resistive switching devices*. Lehrstuhl für Werkstoffe der Elektrotechnik II und Institut für Werkstoffe der Elektrotechnik, 2013.
- [63] M.a. Panzer, M. Shandalov, J.a. Rowlette, Y. Oshima, Yi Wei Chen Yi Wei Chen, P.C. McIntyre, and K.E. Goodson. Thermal Properties of Ultrathin Hafnium Oxide Gate Dielectric Films. *IEEE Electron Device Lett.*, 30(12):2009–2011, 2009.
- [64] Fabien Alibart, Elham Zamanidoost, and Dmitri B Strukov. Pattern classification by memristive crossbar circuits using ex situ and in situ training. *Nature communications*, 4, 2013.
- [65] Sukru Burc Eryilmaz, Duygu Kuzum, Rakesh Jeyasingh, SangBum Kim, Matthew BrightSky, Chung Lam, and H-S Philip Wong. Brain-like associative learning using a nanoscale non-volatile phase change synaptic device array. *arXiv preprint arXiv:1406.4951*, 2014.

- [66] Bipin Rajendran, Yong Liu, Jae-sun Seo, Kailash Gopalakrishnan, Leland Chang, Daniel J Friedman, and Mark B Ritter. Specifications of nanoscale devices and circuits for neuromorphic computational systems. *IEEE Transactions on Electron Devices*, 60(1):246–253, 2013.
- [67] Paolo Cappelletti. Non Volatile Memory Evolution and Revolution. *2015 IEEE Int. Electron Devices Meet.*, pages 241–244, 2015.
- [68] Adenilson J Chiquito, Cleber a Amorim, Olivia M Berengue, Luana S Araujo, Eric P Bernardo, and Edson R Leite. Back-to-back Schottky diodes: the generalization of the diode theory in analysis and extraction of electrical parameters of nanodevices. *J. Phys. Condens. Matter*, 24:225303, 2012.
- [69] Yexin Deng, Peng Huang, Bing Chen, Xiaolin Yang, Bin Gao, Juncheng Wang, Lang Zeng, Gang Du, Jinfeng Kang, and Xiaoyan Liu. RRAM crossbar array with cell selection device: A device and circuit interaction study. *IEEE Trans. Electron Devices*, 60(2):719–726, 2013.
- [70] Jiantao Zhou, Kuk-Hwan Kim, and Wei Lu. Crossbar RRAM Arrays: Selector Device Requirements During Read Operation. *Electron Devices, IEEE Trans.*, 61(5):1369–1376, 2014.
- [71] Jiun Jia Huang, Yi Ming Tseng, Wun Cheng Luo, Chung Wei Hsu, and Tuo Hung Hou. One selector-one resistor (1S1R) crossbar array for high-density flexible memory applications. In *Tech. Dig. - Int. Electron Devices Meet. IEDM*, pages 733–736, 2011.

- [72] Chun Li Lo, Mei Chin Chen, Jiun Jia Huang, and Tuo Hung Hou. On the potential of CRS, 1D1R, and 1S1R crossbar RRAM for storage-class memory. In *2013 Int. Symp. VLSI Technol. Syst. Appl. VLSI-TSA 2013*, pages 0–1, 2013.
- [73] V. S S Srinivasan, S. Chopra, P. Karkare, P. Bafna, S. Lashkare, P. Kumbhare, Y. Kim, S. Srinivasan, S. Kuppurao, S. Lodha, and Udayan Ganguly. Punchthrough-diode-based bipolar RRAM selector by Si epitaxy. *IEEE Electron Device Lett.*, 33(10):1396–1398, 2012.
- [74] Ting Zhang, Xin Ou, Weifeng Zhang, Jiang Yin, Yidong Xia, and Zhiguo Liu. High- $k$ -rare-earth-oxide  $\text{Eu}_2\text{O}_3$  films for transparent resistive random access memory (RRAM) devices. *J. Phys. D: Appl. Phys.*, 47(6):065302, feb 2014.
- [75] Leo Esaki. New phenomenon in narrow germanium  $p - n$  junctions. *Phys. Rev.*, 109:603–604, Jan 1958.
- [76] E. F. Schubert, J. E. Cunningham, and W. T. Tsang. Perpendicular electronic transport in doping superlattices. *Applied Physics Letters*, 51(11):817–819, 1987.
- [77] X. Zhu, X. Zheng, M. Pak, M. O. Tanner, and K. L. Wang. A si bistable diode utilizing interband tunneling junctions. *Applied Physics Letters*, 71(15):2190–2192, 1997.

- [78] Stanford R. Ovshinsky. Reversible electrical switching phenomena in disordered structures. *Phys. Rev. Lett.*, 21:1450–1453, Nov 1968.
- [79] Daniele Ielmini. Threshold switching mechanism by high-field energy gain in the hopping transport of chalcogenide glasses. *Phys. Rev. B*, 78:035308, Jul 2008.
- [80] J. L. Hartke. The threedimensional poolefrenkel effect. *Journal of Applied Physics*, 39(10):4871–4873, 1968.
- [81] Peter L. Young. dc electrical conduction in thin ta2o5 films. i. bulklimited conduction. *Journal of Applied Physics*, 47(1):235–241, 1976.
- [82] John G. Simmons. Poole-frenkel effect and schottky effect in metal-insulator-metal systems. *Phys. Rev.*, 155:657–660, Mar 1967.
- [83] Peter Mark and Thomas E. Hartman. On distinguishing between the schottky and poolefrenkel effects in insulators. *Journal of Applied Physics*, 39(4):2163–2164, 1968.
- [84] M Moravej and Et al. Plasma enhanced chemical vapour deposition of hydrogenated amorphous silicon at atmospheric pressure. *Plasma Sources Sci. Technol.*, 13(1):8, 2004.
- [85] Ryo Nouchi. Extraction of the Schottky parameters in metal-semiconductor-metal diodes from a single current-voltage measurement. *J. Appl. Phys.*, 116(18), 2014.

- [86] Deniz Bozyigit, Weyde M M Lin, Nuri Yazdani, Olesya Yarema, and Vanessa Wood. A quantitative model for charge carrier transport, trapping and recombination in nanocrystal-based solar cells. *Nat. Commun.*, 6:6180, 2015.
- [87] Paul R. Berger. MSM photodiodes, 1996.
- [88] A. Fantini, G. Gorine, R. Degraeve, L. Goux, C. Y. Chen, A. Redolfi, S. Clima, A. Cabrini, G. Torelli, and M. Jurczak. Intrinsic program instability in hfo2 rram and consequences on program algorithms. In *2015 IEEE International Electron Devices Meeting (IEDM)*, pages 7.5.1–7.5.4, Dec 2015.
- [89] Cheng-Chih Hsieh, Anupam Roy, Yao-Feng Chang, Davood Shahrjerdi, and Sanjay K. Banerjee. A sub-1-volt analog metal oxide memristive-based synaptic device with large conductance change for energy-efficient spike-based computing systems. *Applied Physics Letters*, 109(22):223501, 2016.
- [90] C. I. Bliss. The method of probits. *Science*, 79(2037):38–39, 1934.
- [91] L. K. J. Vandamme. Noise as a diagnostic tool for quality and reliability of electronic devices. *IEEE Transactions on Electron Devices*, 41(11):2176–2187, Nov 1994.
- [92] S. Clima, Y. Y. Chen, A. Fantini, L. Goux, R. Degraeve, B. Govoreanu, G. Pourtois, and M. Jurczak. Intrinsic tailing of resistive states



- distributions in amorphous hfox and taox based resistive random access memories. *IEEE Electron Device Letters*, 36(8):769–771, 2015.
- [93] Zhongrui Wang, Saumil Joshi, Sergey E Savelev, Hao Jiang, Rivu Midya, Peng Lin, Miao Hu, Ning Ge, John Paul Strachan, Zhiyong Li, et al. Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nature Materials*, 2016.
- [94] N. Raghavan, R. Degraeve, L. Goux, A. Fantini, D. J. Wouters, G. Groeseneken, and M. Jurczak. Rtn insight to filamentary instability and disturb immunity in ultra-low power switching hfox and alox rram. In *2013 Symposium on VLSI Technology*, pages T164–T165, 2013.
- [95] A. Chen and M. R. Lin. Reset switching probability of resistive switching devices. *IEEE Electron Device Letters*, 32(5):590–592, 2011.
- [96] Aqyan A. Bhatti, Cheng-Chih Hsieh, Anupam Roy, Leonard F. Register, and Sanjay K. Banerjee. First-principles simulation of oxygen vacancy migration in hfox/ceox, and at their interfaces for applications in resistive random-access memories. *Journal of Computational Electronics*, 15(3):741–748, 2016.
- [97] Shimeng Yu, Yi Wu, Rakesh Jeyasingh, Duygu Kuzum, and H. S Philip Wong. An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation. *IEEE Trans. Electron Devices*, 58(8):2729–2737, 2011.

- [98] Zhongqiang Wang, Stefano Ambrogio, Simone Balatti, and Daniele Ielmini. A 2-transistor/1-resistor artificial synapse capable of communication and stochastic learning in neuromorphic systems. *Front. Neurosci.*, 8(January):438, jan 2014.
- [99] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin, 2001.
- [100] Bernhard Schölkopf and Alexander J Smola. Learning with kernels: support vector machines, regularization, optimization, and beyond, 2002.
- [101] Maureen Caudill. Neural networks primer, part i. *AI Expert*, 2(12):46–52, December 1987.
- [102] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [103] S. Park, A. Sheri, J. Kim, J. Noh, J. Jang, M. Jeon, B. Lee, B. R. Lee, B. H. Lee, and H. Hwang. Neuromorphic speech systems using advanced rram-based synapse. In *2013 IEEE International Electron Devices Meeting*, pages 25.6.1–25.6.4, Dec 2013.

- [104] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.
- [105] Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Back-propagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.

## Vita

Cheng-Chih Hsieh was born in Taipei City, Taiwan on 24 March 1988, the son of Shih-Lung Hsieh and Hsiang-Ju Lan. He received the Bachelor of Science degree in Materials Science and Engineering from the National Taiwan University in June, 2010. He started his mandatory military service as platoon leader in Taiwan Army in July, 2010. After retiring from his military service, he moved to United States and received the Master of Engineering degree in Electrical Engineering and Computer Science from University of California, Berkeley. He joined the University of Texas at Austin for the Doctorate program in August, 2012. He completed two summer internships at Hewlett Packard Labs in Palo Alto, CA during his Doctorate program. He received the Doctorate degree in Electrical and Computer Engineering from University of Texas at Austin, May 2017.

Permanent address: andreashsieh@utexas.edu

This dissertation was typeset with L<sup>A</sup>T<sub>E</sub>X<sup>†</sup> by the author.

---

<sup>†</sup>L<sup>A</sup>T<sub>E</sub>X is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's T<sub>E</sub>X Program.