

# **Process Development and Electrical Characterization of CMOS-Integrated Memristive Devices for Emerging Non-Volatile Memory Applications**

Von der Fakultät 1 - MINT - Mathematik, Informatik, Physik,  
Elektro- und Informationstechnik  
der Brandenburgischen Technischen Universität Cottbus-Senftenberg  
genehmigte Dissertation zur Erlangung des akademischen Grades eines

Dr.-Ing.

vorgelegt von

Mamathamba Kalishettyhalli Mahadevaiah  
geboren am 16.09.1990 in Madenahalli (India)

Vorsitzende des Prüfungsausschusses: Prof. Dr. Inga Fischer

Gutachter: Prof. Dr. Christian Wenger

Gutachter: Prof. Dr. Andreas Mai

Gutachter: Prof. Dr. Martin Ziegler

Tag der mündlichen Prüfung: 20.11.2023

<https://doi.org/10.26127/BTUOpen-6557>

*Dedicated to "that which is not"*

---

## Abstract

---

Energy efficiency is vital for future low-power electronic applications. This ultra-low power consumption requirement enables the research beyond the conventional charge-based memories. Further, reliability, high scalability, fast switching, CMOS compatibility, high endurance, etc., are some of the characteristics envisaged by the new generation of emerging non-volatile memories (NVMs). A memristive device or OxRAM is one among the many emerging NVMs, which can exhibit the aforementioned characteristics, and it has the potential to replace the power-hungry conventional NVMs.

The memristive devices have the advantage of monolithic integration with the CMOS logic, which enables the widening of their application areas. Despite their various advantages, the reliability, forming voltages, and variability of the devices pose a hurdle to their wide commercial usage. Hence, it is crucial to identify these factors and mitigate them. This thesis addresses these issues through fabrication process improvements, electrical characterization techniques, and device-engineering methods.

The improvements in the fabrication processes reduced the pristine state currents of the memristive devices. It impacted the reliability and resistive switching performance of the memristive devices directly. To further improve the performance, the memristive devices are integrated into the 130 nm BiCMOS baseline technology of

IHP. Additionally, dedicated test structures are developed to monitor and control the fabrication process steps through in-line electrical characterization.

Further, the forming current and voltage values, along with their dispersions in the 4 kbit memristive arrays, were reduced by utilizing the electrical characterization techniques. Accordingly, the forming operations were performed at high operating temperatures using incremental step pulse and verify algorithm (ISPVA). In contrast to the well-known method of increasing the current compliance (1R) or the gate voltage of the transistor (1T-1R) to increase the conduction filament size, a thin layer of Al<sub>2</sub>O<sub>3</sub> is added. This device engineering technique reduced the variability in both LRS and HRS currents of the memristive devices. Additionally, the conduction filament properties in both states are modeled by using the quantum point contact (QPC) model. Finally, harnessing the intrinsic variability of the memristive devices for neuromorphic computing applications is demonstrated. The reliability of the devices is assessed through endurance and retention characteristics.

**Keywords:** memristive devices, forming voltage, variability, BiCMOS, fabrication process, dispersion, memristive arrays, incremental step pulse and verify algorithm (ISPVA), conduction filament (CF), quantum point contact (QPC), neuromorphic applications and reliability.

---

## Zusammenfassung

---

Energieeffizienz ist für künftige elektronische Anwendungen mit geringem Stromverbrauch von entscheidender Bedeutung. Dieser extrem niedrige Stromverbrauch ermöglicht die Forschung über die herkömmlichen ladungsbasierten Speicher hinaus. Darüber hinaus sind Zuverlässigkeit, hohe Skalierbarkeit, schnelles Schalten, CMOS-Kompatibilität, hohe Lebensdauer usw. einige der Eigenschaften, die von der neuen Generation der aufkommenden nichtflüchtigen Speicher (NVMs) angestrebt werden. Ein Memristive Bauelement oder OxRAM ist einer der vielen aufkommenden NVMs, die die oben genannten Eigenschaften aufweisen können, und er hat das Potenzial, die stromfressenden konventionellen NVMs zu ersetzen.

Die memristiven Bauelemente haben den Vorteil, dass sie monolithisch in die CMOS-Logik integriert werden können, was eine Ausweitung ihrer Anwendungsbereiche ermöglicht. Trotz ihrer zahlreichen Vorteile stellen die Zuverlässigkeit, die Formationsspannungen und die Variabilität der Bauelemente eine Hürde für ihre breite kommerzielle Nutzung dar. Daher ist es von entscheidender Bedeutung, diese Faktoren zu identifizieren und sie zu entschärfen. Die vorliegende Arbeit befasst sich mit diesen Problemen durch Verbesserungen des Herstellungsprozesses, elektrische Charakterisierungstechniken und Methoden der Bauelementekonstruktion.

Die Verbesserungen in den Herstellungsprozessen haben die Ströme im Urzustand der memristiven Bauelemente reduziert. Dies wirkte sich direkt auf die Zuverlässigkeit und die Widerstandsschaltleistung der memristiven Bauelemente aus. Um die Leistung weiter zu verbessern, werden die memristiven Bauelemente in die 130 nm BiCMOS-Basistechnologie des IHP integriert. Darüber hinaus werden spezielle Teststrukturen entwickelt, um die Herstellungsprozesse durch elektrische Inline-Charakterisierung zu überwachen und zu steuern.

Darüber hinaus wurden die Werte für den Formierungsstrom und die Formierungsspannung sowie deren Streuungen in den memristiven Arrays mit 4 kbit durch den Einsatz der elektrischen Charakterisierungstechniken reduziert. Dementsprechend wurden die Umformvorgänge bei hohen Betriebstemperaturen unter Verwendung des ISPVA-Algorithmus (Incremental Step Pulse and Verify) durchgeführt. Im Gegensatz zu der bekannten Methode, die Stromnachgiebigkeit (1R) oder die Gatespannung des Transistors (1T-1R) zu erhöhen, um die Größe des Leitungsfilaments zu vergrößern, wird eine dünne Schicht Al<sub>2</sub>O<sub>3</sub> hinzugefügt. Durch diese Technik wurde die Variabilität der LRS- und HRS-Ströme der memristiven Bauelemente verringert. Darüber hinaus werden die Eigenschaften des Leitungsfilaments in beiden Zuständen mit Hilfe des Quantenpunktmodell (QPC) modelliert. Schließlich wird demonstriert, wie die intrinsische Variabilität der memristiven Bauelemente für neuromorphe Computeranwendungen genutzt werden kann. Die Zuverlässigkeit der Bauelemente wird anhand der Ausdauer und der Retentionseigenschaften bewertet.

Keywörter: Memristiven Bauelemente, Formierspannung, Variabilität, BiCMOS, Herstellungsprozess, Dispersion, memristive Arrays, inkrementeller Schritimpuls- und Verifizierungsalgorithmus (ISPVA), Leitungsdraht (CF), Quantenpunktmodell (QPC), neuromorphe Anwendungen und Zuverlässigkeit.

---

## Acknowledgements

---

I would like to take this opportunity to thank a great lot of people who supported me in completing my doctoral thesis.

First of all, I would like to express my sincerest gratitude to my supervisor Prof. Dr. Christian Wenger, for his constant support and watchful supervision in every step of this thesis. He always saw the ray of hope under the pile of my results. This work would not have been possible without your motivation, guidance, freedom, and insightful comments, from the very beginning of this work. I would not have wished for any better supervision than what I received from you.

Secondly, I would like to thank Prof. Dr. Andreas Mai for his supervision, insightful discussions, comments and motivation throughout the thesis. His constant support towards implementing the MEMRES module into the technology platform and improving it was immense. I truly admire and respect him for the opportunities he provided me to grow professionally within and outside the IHP.

I would like to thank the chairperson of the PhD committee Prof. Dr. Inga Fischer for her support and interest in my work.



I would like to thank my direct supervisor for RRAM-related work, Dr. Eduardo Perez for his constant supervision since the beginning of this work. His guidance regarding the theoretical and practical aspects of this work, insightful discussions, and comments helped me to improve the work. I would like to express my gratitude from the bottom of my heart to Eduardo for reviewing all my work thoroughly.

I would like to express my gratitude to my direct supervisor for technology-related work, Dr. Marco Lisker for his watchful supervision since day one. His guidance, support, motivation, and insightful discussions in the areas of physics and CMOS technologies helped me gain a deeper understanding of the subject areas. My sincerest thanks to him for being my supervisor, well-wisher, good friend, and buddy.

I would like to thank Dr. Martin Drost for supporting me throughout the writing phase of my thesis. He stood strong for me with his smiling face, personally and professionally and constantly motivated me that I can do better. He made sure that I am not stressed and created a conducive atmosphere for me to write down my thesis. I bow down to you for going through my thesis and providing your insightful comments.

I would like to thank all the members of the technology department and process and device research group for being very supportive. I would like to extend my sincere gratitude to Dr. Bernd Heinemann, Dr. Holger Rücker, Dr. Marco Lisker, Dr. Steffen Marschmeyer, Dr. Dirk Wolansky, Dr. Dirk Hagen, Andreas Krüger, Philip Kulse, Alexander Göritz, and Dr. Oksana Fursenko for supporting me with the RRAM module integration process. My heartfelt thanks to Mirko Fraschke, a Marvel etcher and kind person, for the RIE etching of the RRAM stack and supporting the experiments related to RRAM throughout this work. I am also deeply thankful to Dr. Frank Vater, Christian Mai, and Rahul Kumar Yadav for helping me with the PDK and cadence-related work. My sincere thanks to Elise Funke for her support in various administrative works. I am grateful to Burkhard Neumann and Mirko Berfelde for their wonderful IT-services. I would also like to thank Katja Albani, Kathleen Schulte, Christine Richter, and Nicole Götze for their support.

I would like to express my gratitude to “the competence team” members Thomas Bartz, Jörg Franz, Ulrich Ewald and the former team member Felix Ness for their continuous support in the clean room. My heartfelt thanks to all the operators in the clean room for their support. Especially Anita, Vera, Trixi, Martin. S, Jens, Julian, Mendy, Tobias, and Martin. B for their support with the ALD tool and understanding my bad German and carrying out the work perfectly. I would like to thank Dr. Ioan Costina, Florian Bärwolf, Dr. Marvin Zöllner, Jens Katzer, Dr. Andreas Schubert,

David Stolarek, and Beate Kuck for offline characterizations. I am very grateful to Florian Teply for the automatic electrical measurements inside and outside the clean room. I am thankful to Heike Silz and Nicole Ewald for their help regarding dicing the wafers.

Further, I would like to thank my friends and colleagues who were and are at IHP, Pouya, Max, Emilio, Tommy, Andrea, Keerthi, Navaneetha, Rizwan, Rasuole, Martina, Christin, Christian, Martin, Mirko, Anne, Sebastian, Peter and many others who supported me in any regard during the completion of this work.

Further, I would like to thank all the colleagues and management of IHP for being supportive during this entire period. The opportunities IHP provided me to grow with it were immense. I consider myself lucky to be a part of IHP with its outstanding personnel, excellent research facilities and down to the earth scientific community. Knowledge grows by sharing it is what one can experience and learn at IHP.

I would like to thank the DfG project partners Prof. Dr. Hermann Kohlstedt and Finn Zahari from CAU Kiel, and Prof. Dr. Martin Ziegler from TU Ilmenau for the valuable discussions and meetings.

I would also like to express my gratitude to Prof. Dr. Enrique Miranda from Universitat Autònoma de Barcelona (UAB), Spain for his support regarding the assessment of the QPC model in Bi-layer memristive devices.

This work was initially funded by Deutsche Forschungsgemeinschaft (DFG) in the frame of research group FOR2093 and later by NeuroFOR. Further, I would also like to thank the Leibniz Association for open access funding of one of my journals.

My heartfelt thanks to Jürgen Berthold and Andrea Berthold for their support during the writing of this thesis. Thanks to Guru for helping me with the basics of cadence, programming and electronics. Finally, I would like to thank all my friends and family for their support.

---

## List of Publications

---

This thesis has been written by Mamathamba Kalishettyhalli Mahadevaiah under the supervision of Prof. Dr. Christian Wenger and Prof. Dr. Andreas Mai. The publications related to this thesis have been published in international conferences and peer-reviewed journals. The data presented in this thesis was partially published in the below publications.

The data presented in Chapter 3 have been published as:

- M. K. Mahadevaiah *et al.*, “(Invited) Optimized HfO<sub>2</sub>-Based MIM Module Fabrication for Emerging Memory Applications,” *ECS Trans*, vol. 92, no. 4, pp. 211–221, Jul. 2019, doi: 10.1149/09204.0211ecst.

The data presented in Chapter 4 have been published as:

- M. Kalishettyhalli Mahadevaiah *et al.*, “Integration of memristive devices into a 130 nm CMOS baseline technology” in *Bio-inspired Information Pathways - From Neuroscience to Neurotronics*, Springer Series on Bio- and Neurosystems, 2023.

The data presented in Chapter 5 have been published as:

- M. K. Mahadevaiah, E. Perez, and C. Wenger, "Influence of specific forming algorithms on the device-to-device variability of memristive Al-doped HfO<sub>2</sub> arrays," *Journal of Vacuum Science & Technology B*, vol. 38, no. 1, p. 013201, Jan. 2020, doi: 10.1116/1.5126936.
- M. Kalishettyhalli Mahadevaiah *et al.*, "Modulating the Filamentary-Based Resistive Switching Properties of HfO<sub>2</sub> Memristive Devices by Adding Al<sub>2</sub>O<sub>3</sub> Layers," *Electronics (Basel)*, vol. 11, no. 10, p. 1540, 2022.

The data presented in Chapter 6 have been published as:

- M. K. Mahadevaiah *et al.*, "Reliability of CMOS Integrated Memristive HfO<sub>2</sub> Arrays with Respect to Neuromorphic Computing," in *2019 IEEE International Reliability Physics Symposium (IRPS)*, 2019, pp. 1–4. doi: 10.1109/IRPS.2019.8720552.

---

# Table of Contents

---

Abstract.....	i
Zusammenfassung.....	iii
Acknowledgements .....	v
List of Publications.....	viii
Table of Contents.....	x
List of Figures .....	xiv
List of Tables .....	xxii
List of Acronyms and Variables.....	xxiv
1 Introduction .....	1
1.1 Emerging Non-Volatile Memories.....	3
1.2 Classification of Resistive Switching Memories .....	6
1.3 Memristive device applications .....	13

1.4	Thesis overview.....	14
1.4.1	Thesis outline.....	15
2	Experimental.....	17
2.1	Atomic Layer Deposition (ALD).....	18
2.1.1	Overview of ASM's Polygon® 8200 ALD cluster tool.....	19
2.1.2	Growth of Hafniumoxide (HfO <sub>2</sub> ) and its film properties.....	20
2.1.3	Growth of Aluminiumoxide (Al <sub>2</sub> O <sub>3</sub> ) and its film properties.....	24
2.1.4	Growth of HfAlO and its film properties.....	27
2.2	Electrical Characterization.....	28
2.2.1	Electrical Measurement Setup of Keithley 4200A-SCS.....	29
2.2.2	Electrical Measurement Setup of RIFLE SE.....	31
2.3	Summary.....	33
3	Process Development of Memristive Devices.....	35
3.1	Optimization of memristor module fabrication process.....	36
3.2	Dedicated test structures of memristor module.....	39
3.3	Anisotropic dry etching of MIM stack.....	40
3.4	Electrical characterization of the PrCM test structures.....	41
3.5	Electrical characterization of the memristive devices.....	46
3.6	Conclusion.....	50
4	Process Integration of Memristive Devices.....	51
4.1	IHP's BiCMOS Baseline Technologies.....	52
4.2	Integrated MEMRES module in 250 nm BiCMOS technology.....	55
4.3	Integration of the MEMRES module into a 130 nm BiCMOS technology	57

4.4	Layout design of 1T-1R test structures in 130 nm technology.....	60
4.5	Layout design of PrCM test structures in 130 nm technology.....	62
4.6	Impact of MEMRES module integration on a 130 nm baseline technology 65	
4.7	Conclusion .....	67
5	Variability in Memristive Devices .....	68
5.1	Variability in integrated 1T-1R memristive arrays.....	69
5.1.1	Programming algorithms and the memristive array operation .....	70
5.1.2	Variability in forming currents .....	72
5.1.3	Weibull distribution.....	76
5.2	Variability in individual 1T-1R memristive devices .....	81
5.2.1	Fabrication of memristive devices .....	82
5.2.2	Variability in memristive device operations .....	84
5.2.3	Modeling of conduction filament properties .....	89
5.2.4	Conclusion.....	93
6	Towards Neuromorphic Computing by Exploiting Switching Variabilities in HfO <sub>2</sub> -based Memristive Devices.....	95
6.1	Memristive devices as synapses.....	97
6.2	Fabrication and electrical characterization.....	98
6.3	Evaluation of variability in memristive devices.....	101
6.4	Evaluation of reliability in memristive devices .....	104
6.5	Conclusion .....	107
7	Conclusion and Future Outlook .....	109
8	Bibliography .....	112

Appendix A.....	131
Appendix B .....	134



---

## List of Figures

---

Figure 1.1: The memory hierarchy pyramid structure of typical computer systems. Adapted from [6]. The values on the left of the pyramid indicates the access time and on the right of the pyramid indicates the capacity of the memory.....	2
Figure 1.2: Partial classification of eNVM technologies. ....	4
Figure 1.3: The prominent eNVM key players including semiconductor foundries and IDMs [26]. ....	6
Figure 1.4: Schematic representation of (a) a MIM stack of a RRAM device, and its I-V characteristics in (a) unipolar and (b) bi-polar operation modes [14] [38]. ....	8
Figure 1.5: Classification of the redox-based resistive switching memories based on the voltage polarity of operation and resistive switching phenomenon. ....	8
Figure 1.6: Schematic illustration of the redox-based resistive switching mechanisms in (a) TCM [33] and (b) ECM [41] devices, along with their I-V characteristics.....	9
Figure 1.7: Schematic representation of forming in strong filament-based VCM devices along with its I-V characteristic. Adapted from [33]. ....	11

Figure 1.8: Schematic representation of reset and set operations in strong filament-based VCM devices along with its I-V characteristics. The conduction filament depicted in the schematic is not according to any physical model but instead a generic representation. Adapted from [33].	12
Figure 2.1: Schematic illustration of the industrial applications of ALD. Adapted from [84].	18
Figure 2.2: Schematic overview of ASM's Polygon® 8200 ALD cluster tool consisting of two process modules.	20
Figure 2.3: Schematic illustration of an ALD cycle used for the growth of HfO <sub>2</sub> films by using the TALD process. The surface reactions of the precursor (HfCl <sub>4</sub> ) and the co-reactant (H <sub>2</sub> O) molecules are represented in two half-cycles. The bond lengths and the bond angles are not drawn to the scale.	21
Figure 2.4: Wafermap of the (a) thickness and (b) the refractive index of TALD grown 8 nm HfO <sub>2</sub> film on a 200 mm Si-substrate.	22
Figure 2.5: TEM image of HfO <sub>2</sub> film grown using the TALD process on a 3D structure fabricated in a silicon substrate.	22
Figure 2.6: XPS depth profile analysis of an 8 nm HfO <sub>2</sub> film grown using the TALD process.	23
Figure 2.7: AFM image of an as-deposited HfO <sub>2</sub> film grown using the TALD process.	23
Figure 2.8: Schematic illustration of an ALD cycle used for the growth of Al <sub>2</sub> O <sub>3</sub> films by using the TALD process. The surface reactions of the precursor (Al(CH <sub>3</sub> ) <sub>3</sub> ) and the co-reactant (H <sub>2</sub> O) are represented in two half-cycles. The bond lengths and the bond angles are not drawn to the scale.	24
Figure 2.9: Wafermap of the (a) thickness and (b) the refractive index of TALD grown 8 nm Al <sub>2</sub> O <sub>3</sub> layer on a 200 mm Si-substrate.	25
Figure 2.10: TEM image of an Al <sub>2</sub> O <sub>3</sub> film grown using the TALD process on a 3D structure fabricated in a silicon substrate.	26
Figure 2.11: XPS depth profile of an 8 nm Al <sub>2</sub> O <sub>3</sub> film grown using the TALD process.	26

Figure 2.12: AFM image of an as-deposited Al <sub>2</sub> O <sub>3</sub> film grown using the TALD process. .....	26
Figure 2.13: Schematic illustration of a typical ALD supercycle represented using a pressure versus time pulse-scheme diagram, for the deposition of Al doped HfO <sub>2</sub> layers. Adapted from [71]. .....	27
Figure 2.14: Wafermap of (a) the thickness and (b) the refractive index of TALD grown 9 nm HfAlO layers on a 200 mm Si-substrate.....	28
Figure 2.15: Plot of refractive index versus the Al doping percentage, with a linear fit. The investigation was performed on HfAlO layers of similar thickness but with different Al doping. ....	28
Figure 2.16: DC signal waveforms generated by the SMU which are utilized for the electrical characterization of 1T-1R and 1R devices: (a) Linear staircase sweep; (b) dual voltage sweep; (c) DC bias waveform. Adapted from [94]. .....	29
Figure 2.17: Quasi-static measurement setup: (a) SUSS MicroTec PMV200 manual probe station along with its vacuum pump and chuck cooling unit; (b) chuck stage unit with its manipulators in X, Y and Z directions; (c) Keithley 4200A SCS. ....	30
Figure 2.18: The pulsing schemes programmed using RIFLE-SE system (a) single pulse (b) incremental step pulse (c) incremental step pulse with verify. Adapted from [95]. .....	31
Figure 2.19: Pulsed measurement setup for the electrical characterization of 4 kbit memristive arrays consisting of (a) the Cascade Microtech PA200 semi-automatic probe station along with its chuck cooling unit, and (b) the measurement setup for packaged devices: the RIFLE SE. (c) Shows the 4 kbit array on a wafer placed on the setup for measuring the devices. ....	32
Figure 3.1: Schematic illustration of the three short loop process flows of the memristor module fabrication [98]. .....	37
Figure 3.2: The flow chart of the process steps involved in the three different fabrication approaches of the memristor module [98]. .....	38
Figure 3.3: Schematic illustration of (a) class A and (b) class B test structures [98]. ...	39
Figure 3.4: Etching profiles of the MIM stack: (a) tapered etch profile; (b) notching at the metal-insulator interface. The etch profiles are highlighted in red. ....	40

Figure 3.5: Wafer-scale leakage current measurement of class A devices which are fabricated by using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of the leakage currents of type 1, 2 and 3 devices measured at 0.1 V [98]. ..... 42

Figure 3.6: Wafer-scale leakage current measurement of class B devices which are fabricated by using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of the leakage currents of type 1, 2 and 3 devices measured at 0.1 V [98]. ..... 42

Figure 3.7: The schematic illustration of an equivalent circuit of a memristive device with its parasitic components. Adapted from [108]...... 43

Figure 3.8: EDX-based chemical compositional maps of the side wall of the class A test structures fabricated using type (a) 1, (b) 2, and (c) 3 approaches [98]. ..... 44

Figure 3.9: Wafer-scale breakdown voltage measurements of class A devices at (a) 1 pA/ $\mu\text{m}^2$ , (b) 10 pA/ $\mu\text{m}^2$  and class B devices at (c) 1 pA/ $\mu\text{m}^2$ , (d) 10 pA/ $\mu\text{m}^2$  fabricated using three different approaches [98]...... 45

Figure 3.10: Capacitance measurements of (a) class A, (b) class B devices fabricated using three different approaches [98]...... 45

Figure 3.11: Schematic illustration of an electrical measurement of a memristive device. .... 46

Figure 3.12: Wafer-scale forming I-V curves of class A devices fabricated using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of forming voltages of class A devices [98]...... 47

Figure 3.13: Wafer-scale first-reset I-V curves of class A devices fabricated using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of first-reset voltages of class A devices [98]...... 48

Figure 3.14: The distribution of maximum currents in the first reset operation of the memristive devices fabricated using the three different approaches [98]. ..... 48

Figure 3.15: Wafer-scale set I-V curves of class A devices fabricated using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of set voltages of class A devices [98]. .... 49

Figure 3.16: The memory window after 1<sup>st</sup> and 50<sup>th</sup> cycle, of the class A devices fabricated using three different approaches [98]...... 50

Figure 4.1: Schematic cross-section of the components in the SGB25V BiCMOS process technology of IHP. .... 53

Figure 4.2: Schematic cross-section of the components in the SG13S BiCMOS process technology of IHP. ....	53
Figure 4.3: Process flow chart of the fabrication of BEOL interconnects in the SGB25V and SG13S technologies. ....	54
Figure 4.4: Layout design rules for the integration of the MEMRES module into the 250 nm BiCMOS baseline technology. ....	56
Figure 4.5: TEM cross section of a MEMRES module integrated between M2 and M3 in the BEOL interconnects of the 250 nm BiCMOS technology.....	56
Figure 4.6: Layout design rules for the integration of the MEMRES module into the 130 nm BiCMOS baseline technology. ....	57
Figure 4.7: TEM cross section of a MEMRES module integrated between M2 and M3 in the BEOL interconnects of the 130 nm BiCMOS technology.....	58
Figure 4.8: XRD patterns of as-deposited and sinter annealed 8 nm HfO <sub>2</sub> layers deposited using TALD. Adapted from [122]. ....	59
Figure 4.9: XRD patterns of as-deposited and sinter annealed 6 nm Hf <sub>1-x</sub> Al <sub>x</sub> O <sub>y</sub> layers deposited using TALD. Adapted from [122]. ....	59
Figure 4.10: Schematic layout of CMOS integrated 1T-1R memristive test structures in the 130 nm BiCMOS technology. ....	62
Figure 4.11: Schematic layout of a 1T-1R test structure with its contact terminals. ....	62
Figure 4.12: Schematic layout of PrCM test structures dedicated for the MEMRES module in the 130 nm BiCMOS technology. ....	63
Figure 4.13: PrCM test structures (a) a, (b) b, (c) c, and (d) d of the MEMRES module. ....	64
Figure 4.14: Matrix of MIM structures represented as structure e in the PrCM block dedicated for the MEMRES module.....	65
Figure 4.15: Structure f in the PrCM block dedicated for the MEMRES module (a) comprising of comb structures and (b) the schematic illustration of its electrical characterization. ....	65

Figure 5.1: Schematic illustration of (a) ISPVA and (b) CAPVA, pulsed based algorithms [130].	71
Figure 5.2: Block diagram of integrated 4kbit memristive arrays in 250 nm BiCMOS technology.	72
Figure 5.3: (a) Cross-sectional TEM image and (b) the schematic of the integrated 1T-1R memristive device [130].	73
Figure 5.4: D2D CDFs of forming currents of the memristive devices (a) at different temperatures programmed using ISPVA and (b) at different voltage amplitudes programmed using CAPVA [130].	74
Figure 5.5: Yield of formed memristive devices after the forming operation as function (a) of temperature by using ISPVA and (b) of voltage amplitudes by using CAPVA [130].	74
Figure 5.6: Mean values of the read-out currents with error bars measured after the forming procedure of memristive devices (a) at different temperatures by using ISPVA and, (b) at different voltage amplitudes by using CAPVA [130].	75
Figure 5.7: General illustration of the drain current ( $I_d$ ) of the transistor at $V_G = 1.5$ V and $V_{DS} = 1$ V, without the integrated memristive device and the on current ( $I_{on}$ ) of the memristive device without the transistor contribution at various operating temperatures.	76
Figure 5.8: D2D (a) CDFs and (b) the mean values, of forming voltages with error bars of the memristive devices at four different temperatures programmed using ISPVA [130].	78
Figure 5.9: Weibull fitted CDFs of forming voltages of the memristive devices at four different operating temperatures programmed using ISPVA [130].	79
Figure 5.10: (a) The D2D CDFs of the forming times and (b) time to switch 63% of the memristive devices at different pulse amplitudes programmed using CAPVA [130].	80
Figure 5.11: Weibull fitted CDFs of the forming times of the memristive devices at three different voltage amplitudes programmed by using CAPVA [130].	81
Figure 5.12: Cross-sectional TEM images with EDX elemental mapping of memristor modules in (a) V1 (b) V2 and (c) V3 variants [131].	83

Figure 5.13: XRD patterns of rapid thermal annealed 6 nm Al <sub>2</sub> O <sub>3</sub> layer deposited using ALD.....	83
Figure 5.14: Wafermaps of TALD grown Al <sub>2</sub> O <sub>3</sub> layers of thickness (a) 1 nm and (b) 2 nm on a 200 mm Si-substrate. ....	84
Figure 5.15: Forming I-V characteristics of (a) V1, (b) V2 and (c) V3 device variants. The characteristics of individual devices from each variant are represented in grey and the computed median curves are represented in blue [131]......	85
Figure 5.16: Mean values of forming voltages with error bars versus the total dielectric thickness of memristive device variants fabricated in 130 nm CMOS technology of IHP [131]......	86
Figure 5.17: The single layer Al <sub>2</sub> O <sub>3</sub> based MIM devices with and without a Ti layer are (a) electrically characterized for their breakdown voltages (the inset images illustrate the schematic of the layer stack) and (b) material characterized by using cross-sectional TEM images with EDX elemental mapping [131]. ....	86
Figure 5.18: I-V characteristic of a set operation carried out under DC conditions. ...	87
Figure 5.19: LRS currents extracted from DC set operations of V1, V2 and V3 memristive devices [131]......	88
Figure 5.20: HRS currents extracted from DC set operations of V1, V2 and V3 memristive devices [131]......	88
Figure 5.21: Experimental I–V characteristics (symbols) associated with the mean LRS currents extracted from DC set operations of V1, V2 and V3 memristive devices. The solid lines are simulated characteristics using the QPC model for LRS [131]......	92
Figure 5.22: Experimental I–V characteristics (circles) associated with the mean HRS currents extracted from DC reset operations of V1, V2 and V3 memristive devices. The solid lines are simulated characteristics using the QPC model for HRS [131]......	92
Figure 5.23: Schematic representation of the energy band diagram of the conductive filament potential barrier in (a) V1, (b) V2 and (c) V3 memristive devices. E is the energy of electrons, x is the direction of current flow in the filament, E <sub>F</sub> is the Fermi level and φ is the potential barrier height with respect to the Fermi level [131]. ....	93
Figure 6.1: (a) Schematic representation of a biological synapse. (b) Schematic representation of a memristive device which can be used as an electronic synapse. .	97

Figure 6.2: Photograph of the packaged 4 kbit memristive array placed in the socket of the experimental set-up [181].	99
Figure 6.3: Current map of the 4,096 memristive devices after forming, in a packed 64 × 64 array [181].	99
Figure 6.4: CDF of forming currents of the devices programmed by using ISPVA [181].	100
Figure 6.5: Schematic illustration of one-single-pulse mode operation. Adapted from [181].	100
Figure 6.6: Evolution of the D2D distributions of the 1T-1R cells after applying a series of 100 pulses with amplitudes of – 0.6 V (a), -0.9 V (b) and – 1.2 V (c). The dotted lines represent the HRS threshold current value of 10 μA [181].	102
Figure 6.7: Evolution of the D2D distributions of the 1T-1R cells after applying a series of 100 pulses with amplitudes of 0.6 V (a), 0.9 V (b) and 1.2 V (c). The dotted lines represent the LRS threshold current value of 20 μA [181].	102
Figure 6.8: Switching probability plots for the reset (a) and set (b) operations as function of voltage amplitudes and number of pulses [181].	104
Figure 6.9: HRS cumulative distributions after 1 set/reset cycle at different retention times [181].	105
Figure 6.10: LRS cumulative distributions after 1 set/reset cycle at different retention times [181].	105
Figure 6.11: HRS mean (a) and dispersion (b) read currents as a function of retention time and number of endurance cycles [181].	106
Figure 6.12: LRS mean (a) and dispersion (b) read currents as a function of retention time and number of endurance cycles [181].	107



---

## List of Tables

---

Table 1: First generation of commercially available eNVMS. ....	5
Table 2: Comparison of the electrical characteristics of the silicon-based conventional memories and eNVMS. Adapted from [26], [27], [31], [32]. ....	6
Table 3: DC sweep parameters for resistive switching operation of the memristive devices .....	46
Table 4: Size of vias in the BEOL of the 130 nm BiCMOS technology.....	58
Table 5: Dimensions of the six different low voltage transistors of layout section DIGI01.....	61
Table 6: Dimensions of the six different high voltage transistors of layout section HIVT01. ....	61
Table 7: Summary of Weibull fit parameters of the memristive devices programmed using ISPVA [130]. ....	79
Table 8: Summary of Weibull fit parameters and forming time of the memristive devices programmed using CAPVA [130]. ....	81
Table 9: Variants of memristive devices with respective layer thicknesses [131]. ....	82

Table 10: Electrical parameters for operation of V1, V2, and V3 memristive device variants .....	84
Table 11: The mean values of memory window (MW) of V1, V2 and V3 devices determined from the respective mean values of LRS and HRS currents [131]. .....	88
Table 12: The fitting parameter N extracted from the QPC model fit for LRS curves [131].....	91
Table 13: The fitting parameters $\alpha$ and $\phi$ extracted from the QPC model fit for HRS curves and the determined ratios of $T_B/R_B$ [131].....	93

---

## List of Acronyms and Variables

---

### Acronyms

1R	One-Resistor
1T-1R	One-Transistor-One-Resistor
ADC	Analog to Digital Converters
AFM	Atomic Force Microscopy
AI	Artificial Intelligence
ALD	Atomic Layer Deposition
ANN	Artificial Neural Network
ARC	Anti-Reflection Coating
AVD	Atomic Vapor Deposition
BE	Bottom Electrode
BEOL	Back-End-of-Line
BL	Bit Line
C2C	Cycle-to-Cycle
CAPVA	Constant Amplitude Pulse and Verify Algorithm
CCP	Capacitively Coupled Plasma
CDD	Critical Defect Density
CC	Current Compliance
CF	Conduction Filament
CM	Carrier Module

CMOS	Complementary Metal-Oxide-Semiconductor
CMP	chemical mechanical planarization
CNT	carbon nanotube
CPU	Central Processing Unit
CVD	Chemical Vapor Deposition
D2D	Device-to-Device
DC	Direct Current
DLL	Diffusion Limiting Layer
DNN	Deep Neural Network
DRAM	Dynamic Random-Access Memory
DUT	Device Under Test
ECM	Electrochemical Metallization
EDX	Energy Dispersive X-ray
EOL	End-Of-Line
FEOL	Front-End-Of-Line
FeRAM	Ferroelectric Random-Access Memory
FPGA	Field-Programmable Gate Arrays
GB	Grain Boundary
GPU	Graphics Processing Unit
GUI	Graphical User Interface
ICP	Inductively Coupled Plasma
IDM	Integrated Device Manufacturer
ILD	Interlayer Dielectric
IoT	Internet of Things
ISP	Incremental Step Pulse
ISPV	Incremental Step Pulse with Verify
ISPVA	Incremental Step Pulse and Verify Algorithm
LRS	Low Resistance State
M1	Metal 1
M2	Metal 2
M3	Metal 3
MEMS	Microelectromechanical Systems
MIM	Metal-Insulator-Metal
MLP	Multi-Level Programming
MOS	Metal-Oxide-Semiconductor
MW	Memory Window
NOE	Neutral Oxide Etch
NVM	Non-Volatile Memories
PCM	Phase Change Memory

PEALD	Plasma Enhanced Atomic Layer Deposition
PID	Plasma-Induced Damage
PLD	Pulse Laser Deposition
PM	Process Module
PMU	Pulse Measuring Unit
PrCM	Process Control Monitor
PVD	Physical Vapor Deposition
QPC	Quantum Point Contact
RF	Radio Frequency
RIE	Reactive-Ion-Etching
RMS	Root Mean Square
RRAM	Resistive Random-Access Memory
RTA	Rapid Thermal Annealed
SCM	Storage Class Memory
SDM	Source-Delay-Measure
SL	Source Line
SMU	Semiconductor Measurement Units
SNN	Spiking Neural Network
SP	Single Pulse
SPC	Statistical Process Control
SPA	Semiconductor Parameter Analyzer
SRAM	Static Random-Access Memory
SSD	Solid-State Drive
StochANN	Stochastic Artificial Neural Network
STT-MRAM	Spin Transfer Torque Magneto-Resistance Random Access Memory
TALD	Thermal Atomic Layer Deposition
TCM	Thermochemical Memory
TE	Top Electrode
TEL	Thermal-Enhanced Layer
TEM	Transmission Electron Microscopy
TM	Transfer Module
TM1	Top Metal 1
TM2	Top Metal 2
ULP	Ultra-Low Power
VCM	Valence Change Memory
V1	Via 1
V2	Via 2
V3	Via 3
WF	Work Function

WL	Word Line
XPS	X-ray Photoelectron Spectroscopy

## Variables

$W$	Transistor Width
$L$	Transistor Length
$\Sigma$	Standard Deviation
$\mu$	Mean
$I_d$	Drain Current
$V_{DS}$	Drain to Source Voltage
$V_G$	Gate Voltage
$B$	Shape parameter of Weibull Distribution
$H$	Scale parameter of Weibull Distribution
$I$	Current
$V$	Voltage
$E$	Electron Charge
$H$	Planck's Constant
$G/G_0$	Conductance Parameter
$N$	Number of Conducting Channels
$\Phi$	Potential Barrier Height
$T_B$	Potential Barrier Thickness
$A$	Parameter related to $T_B$
$R$	Resistance
$m^*$	Effective Electron Mass
$z_0$	First Order Bessel Function
$R_B$	Radius of the CF Constriction
$E$	Energy of Electrons
$E_F$	Fermi Energy

# Chapter I

---

## 1 Introduction

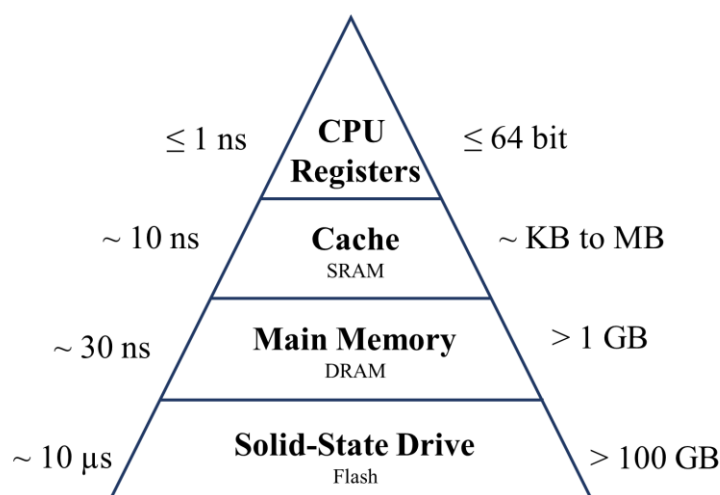
---

Data storage is considered to be of great relevance since ancient times. Stone inscriptions, cave paintings etc., were used to store the information in prehistoric times, whereas parchments were used in the middle ages. The widespread use of printing in the early modern times was an important enabler of the scientific revolution. In the 1900s, punch cards, vacuum tubes, and magnetic core memories became popular. The first semiconductor memory chip was introduced by the Intel Corporation in 1969 [1]. Since then, the world has seen continuous improvements regarding data storage techniques starting with magnetic disks, optical discs, solid-state drives, etc.

The semiconductor memories based on silicon are classified mainly into two types, namely volatile and non-volatile memories (NVM) [2]. The volatile memory holds the data as long as the power supply is on, while the non-volatile holds the data even without the power supply. Both memory types form an integral part of the electronic gadgets used in our day-to-day lives. Inside any electronic computing system, different types of memories are used at various levels for specific applications [3]. The hierarchy of conventional memories with their programming speeds and data storage capacity is illustrated in the form of a pyramid shown in Figure 1.1.

Flash memory, dynamic random-access memory (DRAM), and static random-access memory (SRAM) are the three main conventional silicon-based solid-state

memories used in a computing system [2]. Flash is a NVM, whereas SRAM and DRAM are volatile memories. However, all three are charge-based memories where the information stored is in an electronic charge format. In the case of Flash memory, the information is stored in a floating gate. Flash can store large data, but it is used as a secondary memory in a conventional computer architecture due to its long programming times [4]. DRAM uses one transistor and one capacitor structure to store the data. It requires periodic refreshing of the memory to compensate for the capacitor charge leakage. DRAM is faster than Flash, hence it is used as the main memory. It stores the data and the applications which are currently being used by a processor [5]. SRAM uses six transistors to store the data and hence consumes higher power. It is faster than DRAM and used in the cache memory of the computer for fast access to data and frequently used programs [3]. The registers are memories with a small capacity built into the central processing unit (CPU). The register size determines the speed of the CPU.



**Figure 1.1: The memory hierarchy pyramid structure of typical computer systems. Adapted from [6]. The values on the left of the pyramid indicates the access time and on the right of the pyramid indicates the capacity of the memory.**

The fast data generation rate from electronic gadgets such as smartphones and wearable electronics, internet of things (IoT) and edge artificial intelligence (AI) applications demands fast computation, expansion of data centers, and low power consumption memory technologies [7]. The complementary metal-oxide-semiconductor (CMOS) technology is scaling down continuously to cater to the needs, obeying Moore's law. Accordingly, improvements are being made in memory storage technologies as well. However, the scaling of CMOS and Flash technologies beyond 22 nm becomes challenging mainly due to the high electric field effects, short channel effects, and charge leakage through thin gate oxides [8]. Further, the other challenges such as the von Neumann bottleneck, memory wall, speed gap between the main

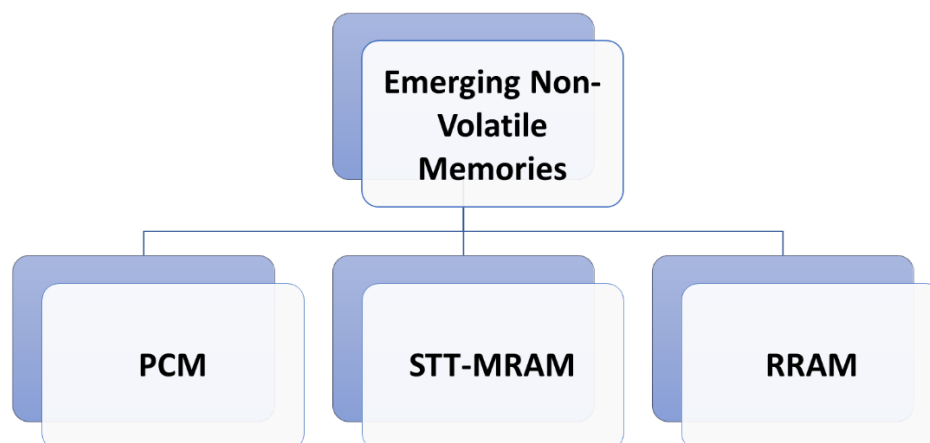


memory (DRAM) and the secondary memory (Flash), ultra-low power (ULP) consumption in standby operation mode, etc., promote the research beyond the charge-based semiconductor memories. This next generation of memories aims for smaller, cheaper, faster, and reliable device performance, and is termed emerging NVM (eNVMs) [3]. They are expected to fill the gaps created by conventional charge-based memories [9].

## 1.1 Emerging Non-Volatile Memories

The characteristic features such as, low power consumption, high scalability, high performance, non-volatility, CMOS compatibility, fast switching operations, etc., are desired from the eNVMs [10]. Various memories based on different physical switching mechanisms are currently under research for different eNVM applications. Memories such as ferroelectric random-access memory (FeRAM) [11], phase change memory (PCM) [12], spin transfer torque magneto-resistance random access memory (STT-MRAM) [13], resistive random-access memory (RRAM) [14], carbon nanotube (CNT) based memories [15], molecular memories [16], silicon nanowire memories [17], and many others have been considered as eNVMs. Out of these, the three memory types illustrated in Figure 1.2 got strong emphasis due to their potential to overcome the limitations of conventional NVMs. These three eNVMs are commonly called memristive devices.

A memristive device is a two-terminal non-linear device that stores the memory in the form of different resistance states, namely low resistance state (LRS) and high resistance state (HRS) [18]. The applied electrical stimuli change the resistance states [19]. In the case of the PCM, the chalcogenide material phase determines the resistance change. The phase change from amorphous to polycrystalline sends the PCM cell from HRS to LRS and vice versa [20]. In the case of the STT-MRAM, the orientation of the ferromagnetic electrodes separated by a tunnel barrier determines the resistance change. The parallel orientation of the electrodes results in LRS, whereas anti-parallel orientation results in the HRS of the device [21]. The orientation can be changed using a spin-polarized current [13]. The resistance change in RRAM devices is due to the distribution of oxygen ions at the interfaces or the movement of oxygen ions or metal cations at the bulk of the switching material [10]. The classification of RRAM devices is diverse, which is discussed in detail in the next sections.



**Figure 1.2: Partial classification of eNVM technologies.**

A summary of the first generation of eNVM commercial products is listed in Table 1. They were introduced into the commercial market at the beginning of 2012. PCM became commercially available in 2017, released by Intel and Micron Technology [22]. In 2016, Everspin released STT-MRAM-based commercial chips [23]. Adesto and Panasonic released RRAM-based commercial chips in 2012 and 2013, respectively [24], [25]. From the first generation of emerging NVMs until now, the capacity and performance of emerging NVM devices were continuously improved.

Further, the number of key players who want to pursue eNVMs, semiconductor foundries and integrated device manufacturers (IDMs) who adopted the fabrication processes has been continuously increasing. This adoption of the eNVM is essential for its mass production [26]. The three prominent eNVM technologies in the commercial market and their corresponding key players are shown in Figure 1.3. The replacement of Flash technology remains the target of emerging NVMs. However, due to the previously mentioned challenges associated with computing systems, the emerging NVMs are actively sought in two main fields of memory applications: (a) Stand-alone memories and (b) Embedded memories.

The stand-alone memory is a separate memory module with its primary function being data storage, and the storage capacity of it is in Gigabytes. Flash memory and DRAM are typical examples of conventional stand-alone memory types [6]. Embedded memory is a small on-chip memory in Megabytes capacity integrated into the core of a processor, controller, or any other logic. The main goal of the embedded memory is to assist the intended functionality of the chip. Flash (Eflash NOR) memory and SRAM are typical examples of conventional embedded memory types [6].

**Table 1: First generation of commercially available eNVMs.**

Memory Type	Industry	Commercial Name	Capacity	Application	Year
PCM	Intel and Micron Technology	3D XPoint™ (Optane)	32 GB	SSD	2017
STT-MRAM	Everspin	EMD3D064M	64 MB	SSD	2016
RRAM	Adesto	RM24EP	32 kB – 1 MB	IoT	2012
	Panasonic	MN101L	64 kB	Embedded memory in MCU	2013

The three prominent eNVMs are particularly attractive due to their non-volatility, CMOS compatibility, high scalability, zero power consumption in standby mode, high endurance, and low write voltages. Further, the eNVMs can be integrated into the back-end-of-line (BEOL) with negligible impact on the front-end-of-line (FEOL) devices [27]. The electrical characteristics of the conventional memories and the eNVMs are summarized in Table 2. PCM is mainly attractive for stand-alone applications due to its high integration density, fast switching, low power consumption, and high endurance features. Accordingly, the capacity of the PCM-based 3D Xpoint solid-state drive (SSD) grew from 32 GB to 1.5 TB within 5 years dominating the stand-alone eNVM market [26]. Due to the growing ULP consumption requirement of embedded systems in standby mode, STT-MRAM is seen as a potential replacement for SRAM in cache memories [28]. Scalability, low programming voltages, and high endurance are some of the features responsible for its adoption into embedded applications. The STT-MRAM dominates the embedded eNVM market with its presence in low-power electronic wearables, microcontrollers (MCUs), AI processors and accelerators, and many others [26]. Along with non-volatility, multi-bit storage, and high endurance, RRAM has the advantage of low power consumption and high scalability in comparison to PCM and STT-MRAM [29]. Hence, RRAM is actively sought for both stand-alone and embedded memory applications as a replacement for Flash and on-chip memory for various embedded systems, respectively. Apart from the initial commercial products of RRAM in 2012 and 2013, Nuvoton-Panasonic released a product consisting of embedded RRAM to the commercial market in 2020 for security applications. Further, TSMC has adopted the fabrication of RRAM in its 22 nm CMOS technology and is currently offering it to customers [30].

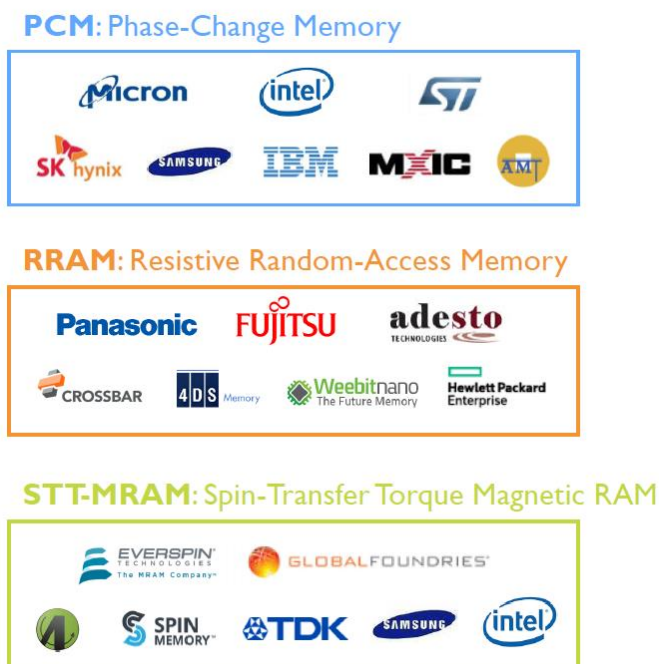


Figure 1.3: The prominent eNVM key players including semiconductor foundries and IDMs [26].

Table 2: Comparison of the electrical characteristics of the silicon-based conventional memories and eNVMs. Adapted from [26], [27], [31], [32].

Features	SRAM	DRAM	Flash	PCM	MRAM	RRAM
			NAND		STT-MRAM	
Integration	FEOL	FEOL	FEOL	BEOL	BEOL	BEOL
Multi-bit	1	1	3	>2	1	2-8
Device Density	120-150 F <sup>2</sup>	4-6 F <sup>2</sup>	< 4 F <sup>2</sup>	4-30 F <sup>2</sup>	6-50 F <sup>2</sup>	<4 F <sup>2</sup>
Write Voltage	< 1 V	< 1 V	> 10 V	< 3 V	< 1.5 V	< 3 V
Write Time	< 1 ns	~ 10 ns	~ 0.1 ms	~ 50 ns	< 10 ns	< 10 ns
Write Energy	~ 1 fJ	~ 10 fJ	< 0.1 fJ	~ 10 pJ	~ 1 pJ	~ 1 pJ
Endurance	~ 10 <sup>16</sup>	~ 10 <sup>16</sup>	> 10 <sup>4</sup>	~ 10 <sup>12</sup>	~ 10 <sup>15</sup>	~ 10 <sup>12</sup>

## 1.2 Classification of Resistive Switching Memories

The RRAM devices are mainly comprised of a Metal-Insulator-Metal (MIM) stack as shown in Figure 1.4 (a). The metal electrode is a good electron conductor and the insulator is usually an oxide, higher chalcogenide, or ionic solid [33]. The selection of the metal electrodes and the insulator layer determines the type of resistive switching.

RRAM devices store the data in the form of resistance change which takes place due to the redox reactions. The redox process is a combination of reduction and oxidation reactions taking place in the MIM stack of the RRAM device. These reactions can be triggered by electrical stimuli, temperature, or both [34].

As shown in Figure 1.5, the RRAM devices can be broadly classified based on the voltage polarity required for their programming and the type of resistive switching mechanisms involved in their operation. Depending on the voltage polarity requirement, the RRAM devices are classified into two types: (a) unipolar switching and (b) bipolar switching. The ideal I-V characteristics of both switching types are shown in Figure 1.4. The unipolar devices can be switched on and off using the same voltage polarity, whereas the bipolar devices require two different voltage polarities [14]. Further, the switching mode is the intrinsic property of the oxide and the electrode/oxide or oxide/oxide interfaces [35].

The RRAM devices are classified further into three main types based on their resistive switching mechanisms. The classification is shown in Figure 1.5. Accordingly, thermochemical memory (TCM), electrochemical metallization (ECM) memory, and valence change memory (VCM) resistive switching mechanisms are discussed in the following sections.

A TCM cell consists of electrochemically inert top and bottom metal electrodes of the same material (e.g. Pt) and an oxide layer (e.g. NiO, TiO<sub>2</sub>) in between them. The resistance change in a TCM cell takes place due to the variation in the stoichiometry of the oxide caused by the dominating thermochemical redox processes over the electrochemical redox processes [36]. An initial electroforming step creates a conduction filament (CF) in the oxide layer of the MIM stack through thermoelectric breakdown. The current through the filament in the on-state is controlled by current compliance (CC). The switching off of the device results in a partial rupture of the CF [37]. The symmetry of the top and bottom electrodes combined with the domination of the thermochemical redox process results in the unipolar resistive switching behavior of the TCM devices [33]. The schematic of the TCM process along with its current-voltage characteristics is shown in Figure 1.6 (a). The usage of single voltage polarity to program the TCM devices is an advantage. However, the high current and large switching time requirement to switch off the device inhibits its usage in low-power and high-speed applications.

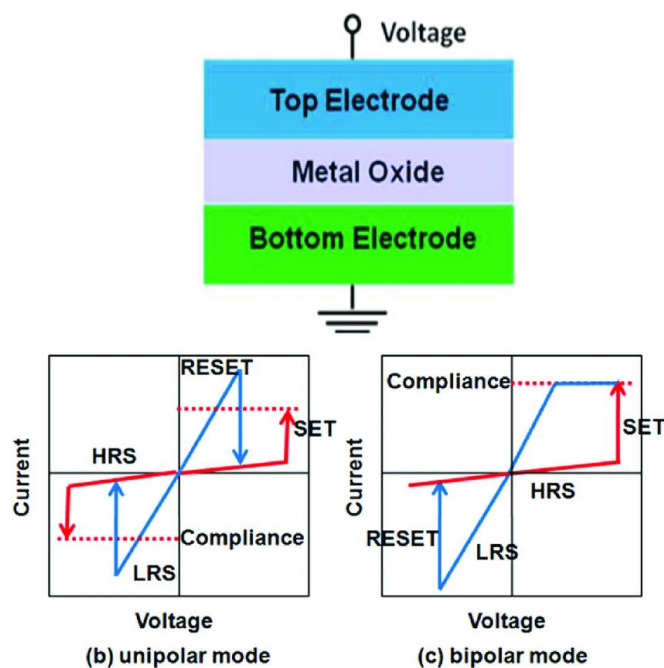


Figure 1.4: Schematic representation of (a) a MIM stack of a RRAM device, and its I-V characteristics in (a) unipolar and (b) bi-polar operation modes [14] [38].

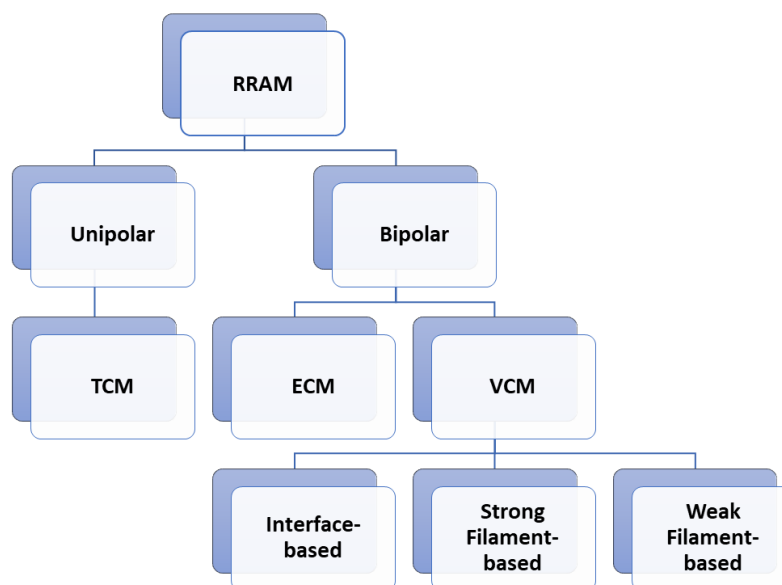
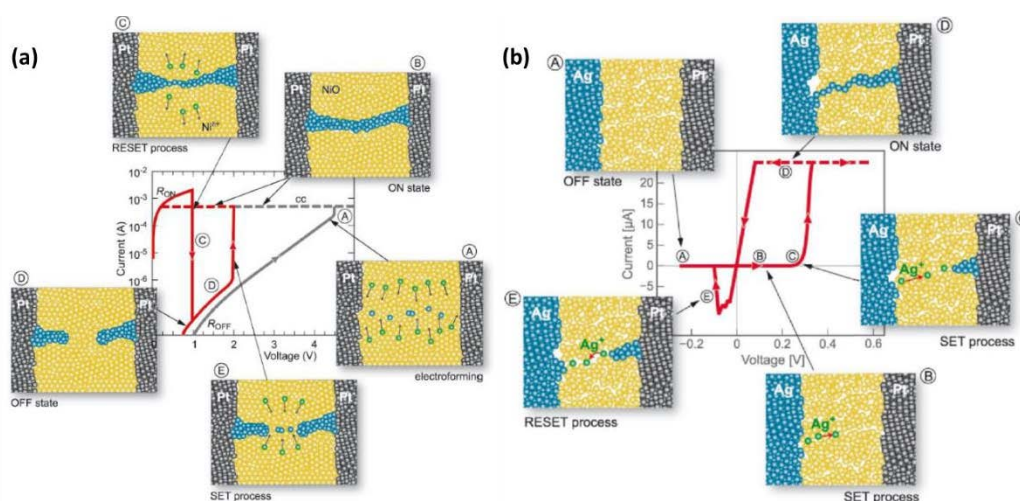


Figure 1.5: Classification of the redox-based resistive switching memories based on the voltage polarity of operation and resistive switching phenomenon.

An ECM cell consists of an electrochemically active (e.g. Ag, Cu) and inactive (e.g., Pt, Au, W, Ir) metal as its top and bottom metal electrodes, respectively, and an ion conducting layer (e.g., GeSe, SiO<sub>2</sub>, Cu<sub>2</sub>S, WO<sub>3</sub>) in between them. Together they form a MIM structure. The resistance change observed in the ECM cell is due to the migration of the metal cations from the active metal electrode to the inert metal electrode through the ion-conducting solid electrolyte layer [39]. Similar to TCM, an electroforming step

is necessary to create an initial CF. Under the influence of a high electric field, the electrochemically active metal electrode undergoes an oxidation reaction and loses electrons. The metal cations generated from the oxidation travel through the solid electrolyte and get reduced on the inert electrode [34]. The process is called electro-crystallization, where a CF is formed in the form of metal dendrites [33]. It represents the on-state of the ECM cell and the current through the metal filament is limited by the CC. Further, the ECM cell can be switched off by reversing the applied voltage polarity in comparison with the electroforming process. During the switch-off process, the metal CF is electrochemically dissolved and redeposited on the active metal electrode. Due to the asymmetry of the metal electrodes and the domination of the electrochemical redox process, bipolar resistive switching is observed in ECM cells [37]. The schematic of the ECM process along with its current-voltage characteristics are shown in Figure 1.6 (b). Only two materials (e.g., Ag and Cu) are dominantly used as the active electrodes due to their lower electrode potentials for electrochemical dissolution. However, Ag is not CMOS compatible. The chalcogenide materials used for the ion-conducting layer have glass transition temperatures between 200 °C and 300 °C which makes them incompatible with BEOL integration in a CMOS process. Further, ECM devices suffer from poor retention, high variability switching and low endurance characteristics [40].



**Figure 1.6: Schematic illustration of the redox-based resistive switching mechanisms in (a) TCM [33] and (b) ECM [41] devices, along with their I-V characteristics.**

The VCM-based RRAM devices consist of a metal-oxide dielectric layer sandwiched between two asymmetric metal electrodes in a MIM stack. A metal-oxide layer consists of a single, or multiple layers of the dielectric stack, deposited using various techniques such as atomic layer deposition (ALD) [42], physical vapor deposition (PVD) [43], pulse laser deposition (PLD) [14], etc. Unlike other RRAM switching mechanisms, the resistance change in a VCM cell takes place due to the

---

movement of oxygen vacancies which is a consequence of the redox reactions [33]. Based on the oxygen vacancy movement, the VCM devices can be further classified into three subcategories as shown in Figure 1.5.

The resistive switching in interface-based RRAM devices takes place due to the distribution of oxygen vacancies at the electrode/oxide or oxide/oxide interfaces [44]. The oxygen vacancies migrate through the entire area of the device, resulting in the modulation of Schottky or tunneling barriers [10]. Although interface-based RRAM devices do not require forming, driving the oxygen vacancies through the entire device area demands larger pulse widths [45]. Further, due to their low current level operation, interface-based devices face retention challenges. Hence, the devices are not suitable for applications where high operation speed and working at high temperatures are of primary concern.

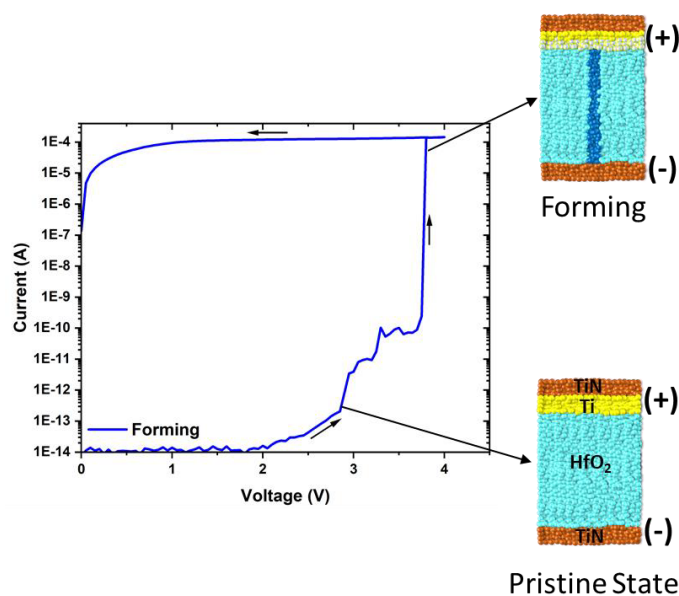
The resistive switching in a strong-filament-based RRAM device takes place due to the movement of oxygen vacancies in a localized region inside a metal-oxide dielectric [46]. It is placed between a chemically active metal electrode with high oxygen affinity and low work function (WF) and a chemically inert metal electrode with comparatively higher WF [33]. Various kinds of binary metal-oxides such as Ta<sub>2</sub>O<sub>5</sub>, NiO, TiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, HfO<sub>2</sub>, etc., are reported in the literature [14]. Among them, HfO<sub>2</sub> layers are extensively studied due to their CMOS compatibility and well-developed deposition processes with high uniformity. Further, the low oxygen affinity metals such as Pt, Au, Ir, TiN, TaN, and W are used as inert electrodes. The high oxygen affinity metals such as Ti, Ta, Hf, and W are used as active metal electrodes/ohmic electrodes [33]. The strong filament-based RRAM is also widely known as OxRAM.

The HfO<sub>2</sub>-based memristive devices integrated into the CMOS baseline technology results in polycrystalline HfO<sub>2</sub> layers. It is caused due to the BEOL thermal budget. The grain boundary conduction in these devices leads to large device-to-device variabilities [47]. Doping HfO<sub>2</sub> layers with aluminum reduces the formation energy of oxygen vacancies and result in reduced variabilities of the switching parameters. However, the retention behavior of memristive devices is affected [48]. To mitigate this issue, instead of doping, a separate layer of thin Al<sub>2</sub>O<sub>3</sub> layers is used, along with HfO<sub>2</sub> layers resulting in a bi-layer memristive device stack. The thin Al<sub>2</sub>O<sub>3</sub> layers act as a tunneling barrier, resulting in improved thermal stability and endurance characteristics [49].

Besides the Ta<sub>2</sub>O<sub>5</sub>-based memristive stack, the HfO<sub>2</sub>-based memristive stack (TiN/HfO<sub>2</sub>/Ti/TiN) is widely investigated by research institutes, foundries, and IDMs



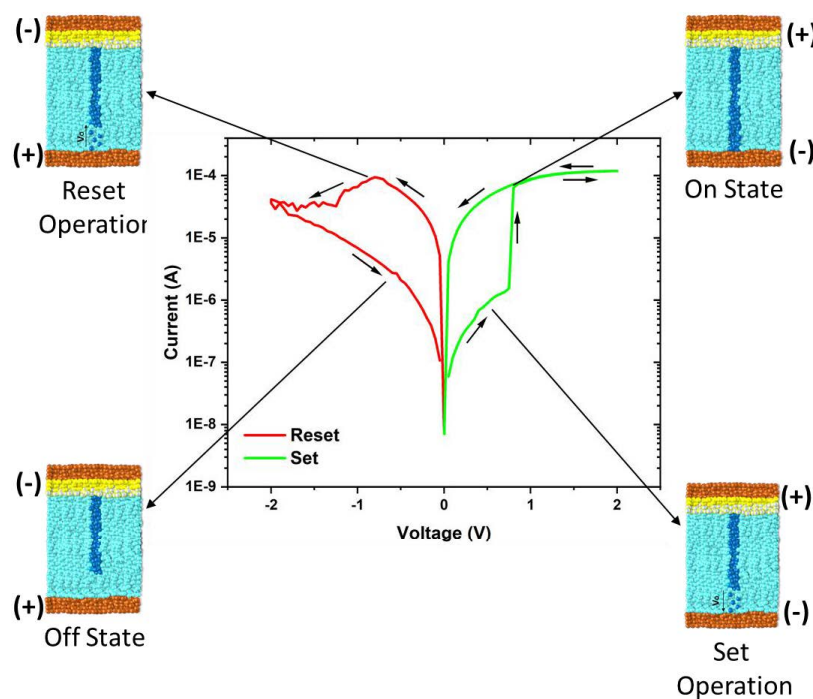
due to their CMOS compatibility in the BEOL fabrication processes. This particular stack is considered for the following discussion. Initially, an electroforming step is performed on a pristine state of a device as a part of its preconditioning. Forming is a one-time operation where a high enough electric field is applied, which kicks the oxygen atoms out of the lattice [50]. The metal-oxide bonds are broken, resulting in the reduction of an oxide ( $\text{HfO}_2$ ) switching layer. This bond breakage results in the formation of an oxygen vacancy and oxygen ion Frenkel pair [51]. Under the influence of the high electric field, the oxygen ions drift towards the anode (Ti), resulting in its oxidation [52]. The ohmic electrode acts as an oxygen reservoir. The deficiency of oxygen in the metal oxide ( $\text{HfO}_2$ ) layer leads to the formation of metal precipitates or oxygen vacancies in a localized region whose conduction is ohmic in nature [53]. This localized region is termed CF. Due to the large concentration of oxygen vacancies in this region, the device type is categorized as strong-filament-based RRAM. The forming operation is schematically illustrated along with its I-V characteristic in Figure 1.7.



**Figure 1.7:** Schematic representation of forming in strong filament-based VCM devices along with its I-V characteristic. Adapted from [33].

The CF created in forming step can be destroyed and recreated multiple times through reset and set operations, respectively. The reset operation sends the device to the off-state or HRS. The set operation sends the device to the on-state or LRS. The ideal I-V characteristic of a bi-polar resistive switching is shown in Figure 1.4 (c). The forming operation is succeeded by a reset operation during which a reverse voltage polarity is applied on the ohmic electrode. The oxygen ions return to the dielectric layer, and the CF in the localized region gets partially re-oxidized by reducing the

conduction level of a device [52]. During the set operation, the oxygen ions are re-captured by the ohmic electrode by leaving the partially oxidized part of the CF during the reset operation to become oxygen deficient again [50]. Similar to forming operation, ohmic conduction mechanism is observed in the resulting conductive level of the set operation as well. The schematic representation of the reset and set operations in a MIM stack along with their experimental I-V characteristics and the corresponding voltage polarities are illustrated in Figure 1.8. The conduction mechanism in HRS has been explained using models such as trap-assisted tunneling [54], Poole-Frenkel emission [52], Schottky emission [55], Fowler-Nordheim tunneling [56], quantum point contact (QPC) [57], and many others. However, for the LRS of the device, mainly the ohmic conduction type is reported [14]. Compared to interface-based devices, strong filament-based RRAM devices have the advantage of good retention and fast switching speed.



**Figure 1.8:** Schematic representation of reset and set operations in strong filament-based VCM devices along with its I-V characteristics. The conduction filament depicted in the schematic is not according to any physical model but instead a generic representation. Adapted from [33].

The MIM stack of weak-filament-based RRAM devices is similar to strong-filament-based RRAM devices. However, they consist of additional layers such as diffusion limiting layer (DLL) [58] or thermal-enhanced layer (TEL) [59]. Further, doping elements such as aluminum (Al) [46] and manganese (Mn) [60] can be incorporated into the binary metal-oxide. The underlying physical mechanisms

behind each of them are different. However, they aim to achieve multiple weak conduction filaments through high oxygen vacancy density in a narrow region compared to interface-based switching, and low oxygen vacancy density in a wider region compared to strong-filament-based switching [45]. In addition to good retention, high operation speed, and CMOS compatibility, the devices exhibit multi-level conduction through analog switching. The analog switching behavior in RRAM devices is one of the main features desired for neuromorphic computing applications. The weak filament-based RRAM devices are an ongoing area of research and require deep insights into the device physics before their wide acceptance.

Among the various kinds of eNVMs discussed so far, the OxRAM devices have the unique advantage of CMOS compatibility, ease of integration in BEOL processes, scalability, and many other previously mentioned features. Further, OxRAM-based device modeling has been an active area of research. A wide range of atomistic, physical and compact models exist. Such models can be utilized to understand the device physics as well as simulate their behavior in macros. This thesis is focused mainly on OxRAM devices, and henceforth they are referred to as memristive devices.

### **1.3 Memristive device applications**

The characteristic features of memristive devices discussed previously make them potential candidates for various memory-related applications. Similar to other eNVMs, memristive devices are also actively pursued in stand-alone and embedded memory application areas. Additionally, they are being widely investigated for neuromorphic computing applications. As mentioned before in section 1.1, memristive devices are used for data storage and as on-chip memory in the case of stand-alone and embedded applications, respectively. Whereas, they are used for computation in neuromorphic applications.

As shown in Figure 1.1, a large speed-capacity gap exists between the main memory (DRAM) and the secondary memory (Flash). To mitigate this latency gap, a new category of memory storage called storage class memory (SCM) is introduced between working memory and storage memory. The main function of the SCM is to improve the throughput of the computing systems by enabling faster data access time [6]. The memristive devices-based SCM is expected to have the scalability and operation speed of DRAM, non-volatility of Flash, and low-power operation. Further, the memristive device integration steps are lower compared to standard Flash [29]. Utilizing the memristive devices for SCM is an active area of research.

The ULP consumption requirements for applications such as IoT, smart wearables, mobile phones, etc., demand highly scalable, enhanced performance, and low power-consuming memristive devices [27]. Extensive research has been going on to embed memristive devices into integrated circuits such as field-programmable gate arrays (FPGA) [61], logic circuits [62], microprocessors [63], and cache memories [64], as on-chip memories. Further, the memristive devices are used together with SRAM and Flip-Flops, which makes them non-volatile and suppresses the standby power consumption [62]. The memristive devices have been used as embedded memory in commercial microcontrollers from Panasonic [25].

In conventional computing systems, the logic and memory units are separated. The data is always fetched from memory, processed in the logic, and stored back into the memory. The data transfer between logic and memory units for every operation result in a latency delay. It can be mitigated with new computing architectures. In-memory computing or neuromorphic computing is one such architecture. It is a highly parallel and efficient computation technique inspired by the human brain [65]. Artificial synapses and neurons are the two main components of neuromorphic computing, whose behavior is emulated by using BEOL memristive devices and FEOL CMOS logic, respectively. Further, they can be monolithically integrated on the same CMOS baseline technology [66]. The memristive devices allow multi-level programming (MLP), which is one of the most desired features for neuromorphic computing [67].

## 1.4 Thesis overview

The memristive devices have demonstrated excellent scalability, fast switching, good endurance, and retention characteristics. However, there are open challenges associated with memristive devices that must be addressed for their wide acceptance in commercial applications. The device characteristics, such as forming voltages, stochastic variabilities, and reliability, pose a hurdle to commercial device usage. This thesis focuses on the aforementioned issues of the memristive devices and addresses them through process development and electrical characterization techniques.

Firstly, under process development, the memristive devices are fabricated using improved process techniques. The technique is adapted during the memristor module integration into the 130 nm CMOS baseline technology of IHP. Secondly, as a part of the characterization technique, integrated 4kbit memristive arrays are electrically characterized using two different pulse-programming schemes, and their impact on the forming operation is assessed. Further, a device-engineering technique is

employed on the memristive devices to mitigate variability. Additionally, the conduction filament properties of these devices are modeled using a physics-based analytical model. Finally, the stochastic variability in the memristive devices is harnessed for neuromorphic computing applications. In particular, the analog synaptic functionality is realized using such stochastic variability of the memristive devices.

### 1.4.1 Thesis outline

This thesis consists of seven chapters and it is structured as below:

Chapter 1 gives a general overview of conventional semiconductor memories and the necessity to pursue emerging non-volatile memories (NVMs). The switching mechanisms behind PCM, STT-MRAM, and RRAM emerging NVMs, along with their application areas are briefly discussed. The chapter reviews the classification of RRAM devices based on the voltage polarity requirement and resistive switching mechanisms. Since the thesis is focused mainly on VCM-type RRAM devices (memristive devices), its resistive switching mechanisms behind the forming, reset, and set operations are discussed. Finally, the specific application areas where the memristive devices have the potential to replace conventional memories are briefly discussed.

Chapter 2 describes the experimental work carried out in this thesis. A general description of the Polygon® 8200 tool from ASM is given. It is used to deposit the memristive switching layers utilized in this work. The physical mechanism behind the growth process of the layers is explained concerning the specific precursors used for their deposition. The properties of the films are investigated using various material characterization techniques. Further, the characterization tools and the operation of the probe stations used for the electrical measurement of different types of memristive devices are briefly described.

The fabrication process variations are one of the reasons behind the device variabilities observed in memristive devices. It can be mitigated through optimized fabrication techniques and narrow process window variations. This is realized in Chapter 3 by fabricating the memristive devices using three different approaches in the MEMRES module. The impact of the fabrication approaches on the memristive devices is tested on the process control monitor structures. Further, the memristive devices are electrically characterized to correlate the resistive switching properties of the devices with the fabrication approaches.

The memristive devices fabricated using the improved fabrication technique from Chapter 3 have to be integrated into the CMOS baseline technology to enhance the performance of the devices. Chapter 4 describes the integration process of the MEMRES module into a 130 nm BiCMOS baseline technology of IHP and its benefits. Further, the layout design rules concerning the integration process, selection of BEOL metal levels for the module integration, and influence of the BEOL thermal budget are comprehensively reviewed. Two independent layouts are explicitly designed for individual 1T-1R memristive devices and the process control monitor test structures. Finally, the impact of MEMRES module integration on the FEOL devices and BEOL processes is investigated through standard electrical measurements.

Chapter 5 reports the variability issue in memristive devices, one of the main challenges to be mitigated for its mass production and commercial usage. The first section of the chapter explains the variability in the forming operation of the integrated memristive arrays. An electrical characterization technique with two different programming algorithms consisting of verify schemes is used to evaluate the forming operation. The parameters of the forming operation from both programming algorithms are fitted using the Weibull distribution. A device engineering technique is employed in the second section of the chapter to mitigate the intrinsic variability issue in memristive devices. The conduction properties and the resistive switching variabilities in the single and bi-layer memristive devices are compared. The conduction filament properties in the LRS and HRS of the memristive device are explained using the QPC model.

Utilizing the stochastic switching variability in the memristive devices to emulate the behavior of artificial synapses is explored in Chapter 6. The integrated memristive devices are electrically characterized by using two different programming schemes. The forming operation is performed through incremental step pulse and verification programming scheme. The switching variability of the memristive devices are evaluated using a one-single-pulse programming scheme. The switching probability values of the memristive devices are extracted and evaluated at different voltage pulse amplitudes for synaptic information processing. Finally, the reliability of the memristive devices is evaluated through endurance and retention measurements.

## Chapter II

---

## 2 Experimental

---

This chapter focuses mainly on the experimental part of the work. Firstly, the fundamentals of atomic layer deposition (ALD) technique are discussed. The ALD cluster tool utilized for the deposition of HfO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub> and Al doped HfO<sub>2</sub> dielectric layers is briefly presented. The aforementioned layers are mainly used as the memristive switching layers in this work. The self-terminating surface reactions of ALD HfO<sub>2</sub> and Al<sub>2</sub>O<sub>3</sub> layers is explained with respect to their specific precursors and the co-reactants utilized for the process. Further, the doping of HfO<sub>2</sub> layers with elemental Al carried out by using the individual HfO<sub>2</sub> and Al<sub>2</sub>O<sub>3</sub> cycles in an ALD super cycle is briefly discussed. In addition, the ALD layers are material characterized for their uniformity, stoichiometry, conformality and surface topography by using various metrology techniques.

Secondly, the chapter discusses the experimental setups used for the electrical characterization. Further, two different measurement techniques used for the electrical characterization of memristive devices under this framework are presented. At first, the quasi-static direct current (DC) measurement technique used for the characterization of one-resistor (1R) and one-transistor-one resistor (1T-1R) devices is explained. The characterization of 4 kbit memristive arrays is discussed with respect to the pulse measurement technique. Finally, the chapter presents briefly the signal waveforms used for both measurement techniques.

## 2.1 Atomic Layer Deposition (ALD)

Atomic layer deposition (ALD) is one of the promising technologies used for the growth of thin films [42]. It is used as a key process for various industrial applications due to its excellent process controllability and well-established deposition techniques [68]–[70]. Other factors such as compatibility with CMOS technologies, availability of the precursor chemistries and production grade tools with various configurations also contribute to its successful usage in diverse application areas [71]. Figure 2.1 shows the wide areas of industrial applications of ALD, from large scale manufacturing to niche markets.

Firstly, ALD finds its application in integrated electronic circuits as high-k material for gate dielectrics [72], memristive switching layers [73], interconnect liners [74], waveguide material [75] etc. among many others. Further, ALD layers are used in energy technologies for electrode materials [76], protective coatings [77], passivation layers [78] and as an encapsulation material [79]. Finally, in the field of medicine, ALD films are used as anti-corrosion layers [80], protective layers [81] and diffusion barrier layers [82]. ALD technology has been considered as the enabler for scaling down of the More-than-Moore devices [83]. Despite its lower growth rate and throughput, it is widely used as a key process in the microelectronics industry due to the deposition of high-quality thin films.

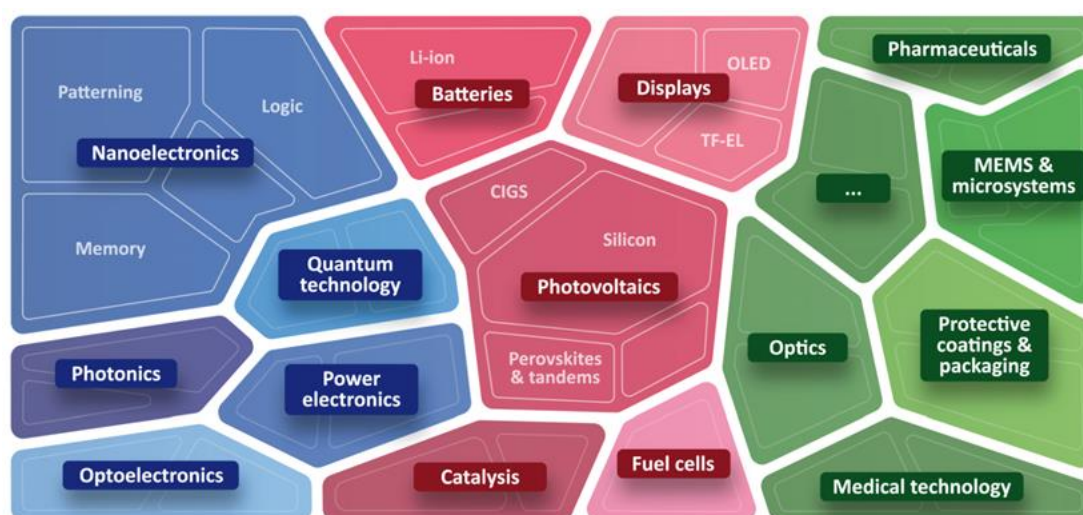


Figure 2.1: Schematic illustration of the industrial applications of ALD. Adapted from [84].

ALD is one of the most popular techniques used in the fabrication of emerging non-volatile memories such as oxide-based resistive random-access memory (RRAM) or otherwise called as memristive devices [85]. Along with the uniformity, stoichiometry and conformality of the deposited films, the deposition temperatures of



ALD being compatible with the BEOL processes are the main reasons behind their widespread usage [86]. Further, ALD technology enables to deposit various oxide layers in the form of nano-laminate structures and to introduce dopants with variable concentrations into the oxide layers through ALD super cycles [71]. Due to its aforementioned key features, ALD is prominently attractive for depositing memristive switching layers in case of 1T-1R arrays [87], cross-bar arrays [88], 3D integrated memristive arrays [89] and many others.

### 2.1.1 Overview of ASM's Polygon® 8200 ALD cluster tool

Figure 2.2 shows the layout overview of ASM's Polygon® 8200 ALD cluster tool. The memristive switching layers used in this work were mainly deposited by using this tool. It consists of two process modules (PM) or chambers: PM2/EmerALD® and PM4/Pulsar®. EmerALD® is a front-end-of-line (FEOL) chamber and Pulsar® is a BEOL chamber. In order to prevent any possible cross contaminations, the two chambers have their own load locks/carrier modules (CMs). CM1 is dedicated for EmerALD® processes and CM2 is dedicated for Pulsar® processes. The PM5 is the cooling station which is used for cooling the wafers coming out of the PM2 or PM4. The TM1 is the transfer module (TM) which is used for transferring the wafers between the load locks, the cooling station and the two process modules.

EmerALD® is a plasma enhanced atomic layer deposition (PEALD) chamber. The deposition process is assisted by thermal and chemical energies [70]. The chamber is used for the deposition of SiO<sub>2</sub> and SiN layers. As the layers deposited in this chamber were not used for this thesis, further discussion on this chamber and the processes related to it is out of scope of this work.

Pulsar® is a Thermal Atomic Layer Deposition (TALD) chamber. The energy required for the surface reactions is provided by thermal energy alone. The chamber is suitable for the deposition of HfO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub> and a mixture of both the films. The TALD deposition temperature window for HfO<sub>2</sub> and Al<sub>2</sub>O<sub>3</sub> films ranges between 200 °C and 350 °C. The layers from the Pulsar® chamber find their application mainly as switching layers in memristive devices which are mostly used for neuromorphic and embedded applications.

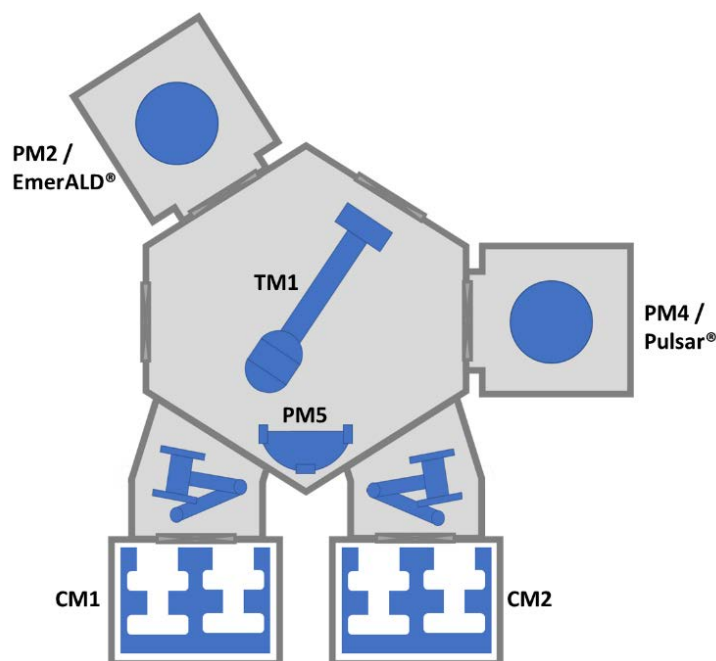


Figure 2.2: Schematic overview of ASM's Polygon® 8200 ALD cluster tool consisting of two process modules.

### 2.1.2 Growth of Hafniumoxide ( $\text{HfO}_2$ ) and its film properties

Hafnium oxide films are grown in the Pulsar® chamber using the TALD process at 300 °C. During the deposition process, the precursor and the co-reactant are fed into the process chamber sequentially [71]. Hafnium tetrachloride ( $\text{HfCl}_4$ ) and water ( $\text{H}_2\text{O}$ ) are used as the precursor and the co-reactant, respectively. The two feeds are separated by an inert gas purge to avert possible gas phase reactions. The deposition takes place through a self-terminating surface reaction process where the films are grown layer by layer [69]. The  $\text{HfCl}_4$  solid precursor is placed in a high-temperature inert gas (HIG) source cabinet which is heated to a certain temperature in order to vaporize it. Nitrogen ( $\text{N}_2$ ) is used as a carrier gas to deliver the vapors of  $\text{HfCl}_4$  and  $\text{H}_2\text{O}$  to the reaction chamber sequentially.

Figure 2.3 demonstrates the growth of a  $\text{HfO}_2$  monolayer through a typical ALD cycle on Si substrate. An ALD reaction cycle consists of two self-limiting half-cycles. In the first half-cycle, a pulse of  $\text{HfCl}_4$  is fed into the reaction chamber in a saturation dose. The precursor molecule chemisorbs on the silicon wafer surface with OH terminations which are formed due to the oxygen and moisture present in the ambient air [90]. Hydrogen chloride ( $\text{HCl}$ ) gas is generated as a volatile by-product of the surface reaction process. The precursor pulse step is subsequently followed by a nitrogen purge step. The purge step removes the unreacted  $\text{HfCl}_4$  and the  $\text{HCl}$  by-product from the process chamber. In the second half-cycle, a co-reactant pulse of  $\text{H}_2\text{O}$

is introduced. The  $\text{H}_2\text{O}$  readily reacts with the adsorbed precursor by substituting the remaining Cl ligands. Thereby, a monolayer of hafnium oxide with OH surface terminations is formed which assists the next surface reaction step. A subsequent nitrogen gas purge will remove the unreacted  $\text{H}_2\text{O}$  and the HCl by-product from the process chamber. The desired thickness of  $\text{HfO}_2$  is grown layer-by-layer by repeating the half-cycles sequentially. The net reaction between  $\text{HfCl}_4$  and  $\text{H}_2\text{O}$  is expressed as:

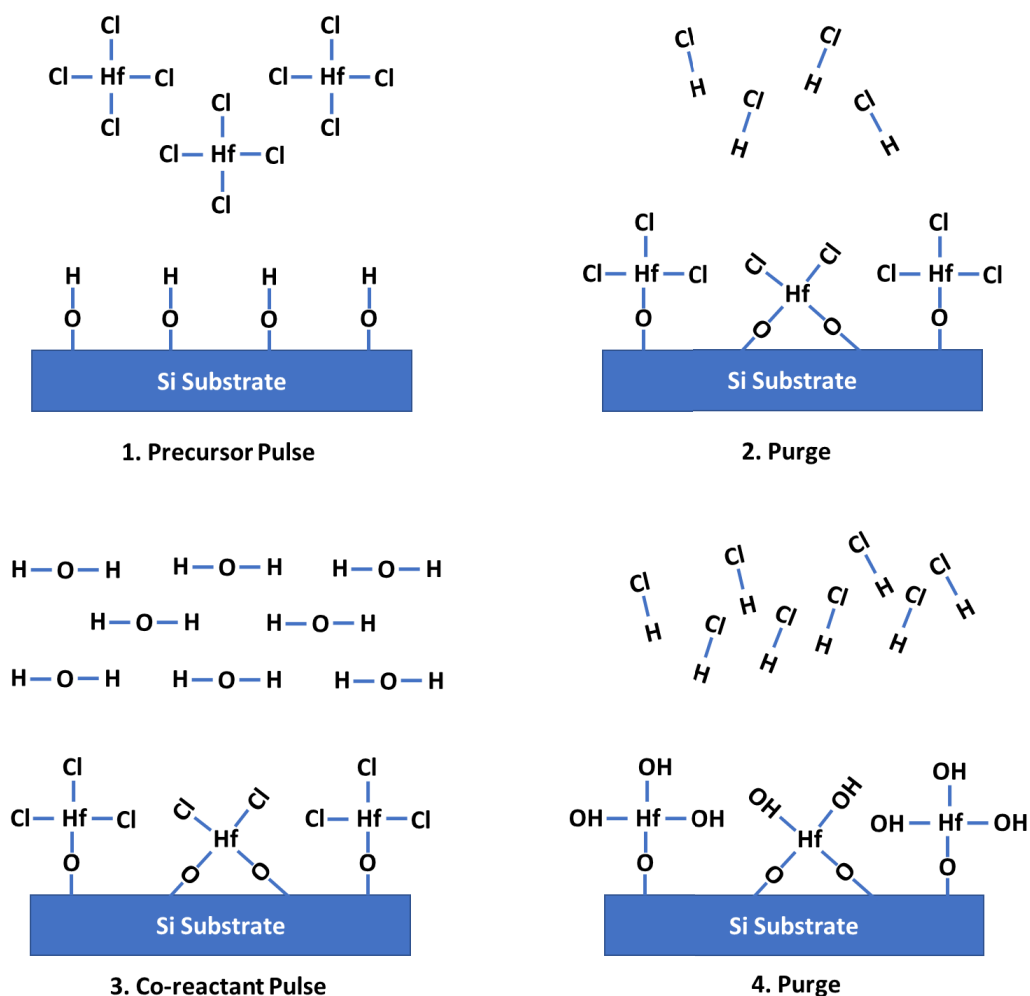


Figure 2.3: Schematic illustration of an ALD cycle used for the growth of  $\text{HfO}_2$  films by using the TALD process. The surface reactions of the precursor ( $\text{HfCl}_4$ ) and the co-reactant ( $\text{H}_2\text{O}$ ) molecules are represented in two half-cycles. The bond lengths and the bond angles are not drawn to the scale.

The  $\text{HfO}_2$  layers deposited using the TALD process have a growth rate of  $\sim 0.6 \text{ \AA}/\text{cycle}$ . For a given temperature and precursor chemistries with sufficient pulse and purge times, the growth rate depends mainly on the number of ALD cycles. Hence, the thickness of the deposited layers even in high aspect ratio structures can be

accurately controlled [83]. Figure 2.4 illustrates the wafermap of 8 nm  $\text{HfO}_2$  films grown using the TALD process. The thickness of all the dielectric layers deposited using the TALD process in this work, is measured by using spectroscopic ellipsometry through X-ray reflectivity. Further, their uniformity is evaluated by measuring the thickness and refractive index at 49 points in a circular contour with 3 mm edge exclusion on a 200 mm wafer. The refractive index of all the dielectric films is measured at a wavelength of 633 nm. The range of  $\text{HfO}_2$  thickness variation across the entire wafer was less than 3 Å and the standard deviation of the deposited layers was below 1% ( $1\sigma$ ) for both thickness and refractive index. The measured refractive index of the  $\text{HfO}_2$  layers was 2.09. The conformality of the deposited  $\text{HfO}_2$  layers is evaluated using a 3D feature fabricated in a silicon substrate as illustrated in the transmission electron microscopy (TEM) image in Figure 2.5. The step, bottom and side wall coverage of the deposited  $\text{HfO}_2$  layers was greater than 95%.

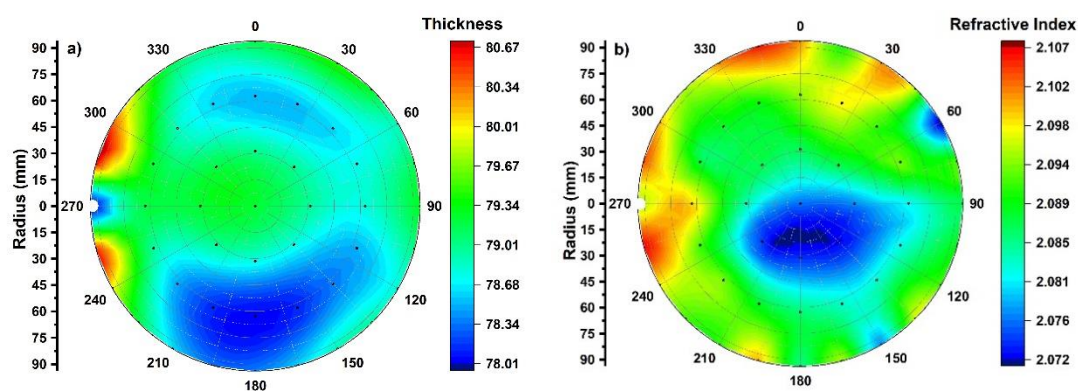


Figure 2.4: Wafermap of the (a) thickness and (b) the refractive index of TALD grown 8 nm  $\text{HfO}_2$  film on a 200 mm Si-substrate.

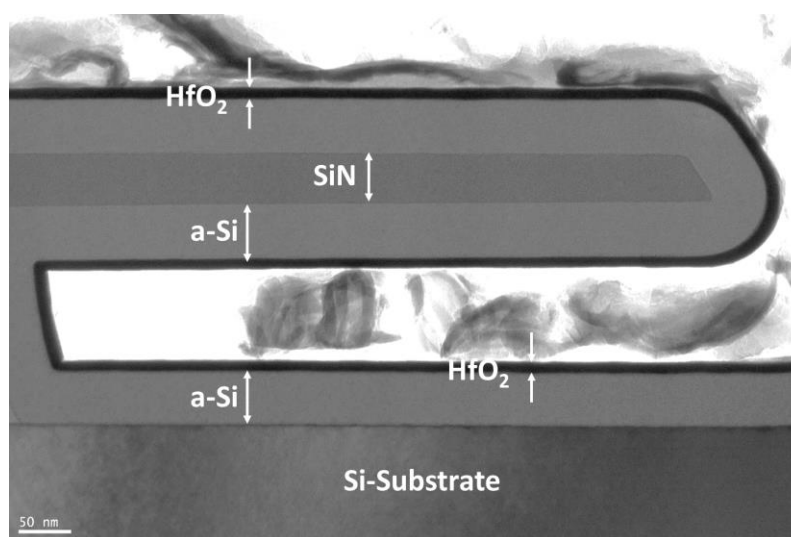


Figure 2.5: TEM image of  $\text{HfO}_2$  film grown using the TALD process on a 3D structure fabricated in a silicon substrate.

The elemental composition of HfO<sub>2</sub> film deposited on Si substrate is investigated by using X-ray photoelectron spectroscopy (XPS) depth profile analysis. Figure 2.6 shows the atomic concentrations of Hf, O, N and Si versus the sputtering time. The determined atomic concentration ratio of O/Hf was ~1.9 indicating that the deposited layers are stoichiometric. The surface morphology of the HfO<sub>2</sub> films was analyzed in 3D by using atomic force microscopy (AFM). The deposited films are found to be very smooth as shown in Figure 2.7 with the measured root mean square (RMS) roughness value of 0.18 nm.

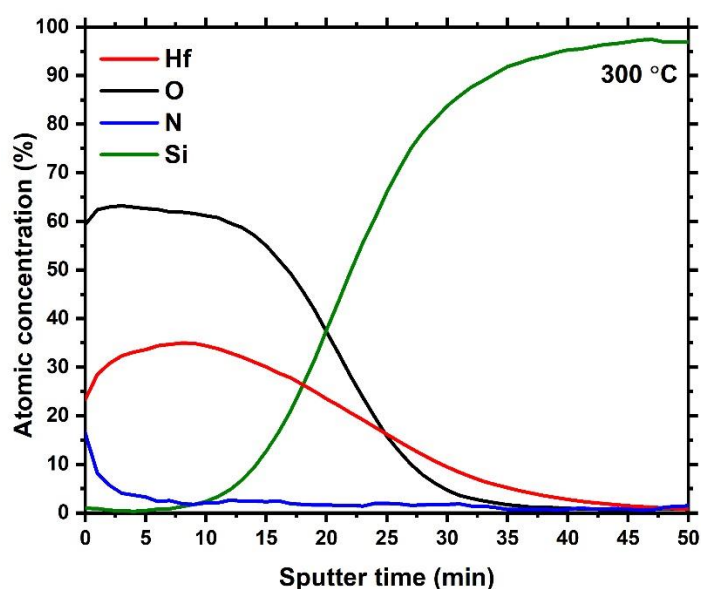


Figure 2.6: XPS depth profile analysis of an 8 nm HfO<sub>2</sub> film grown using the TALD process.

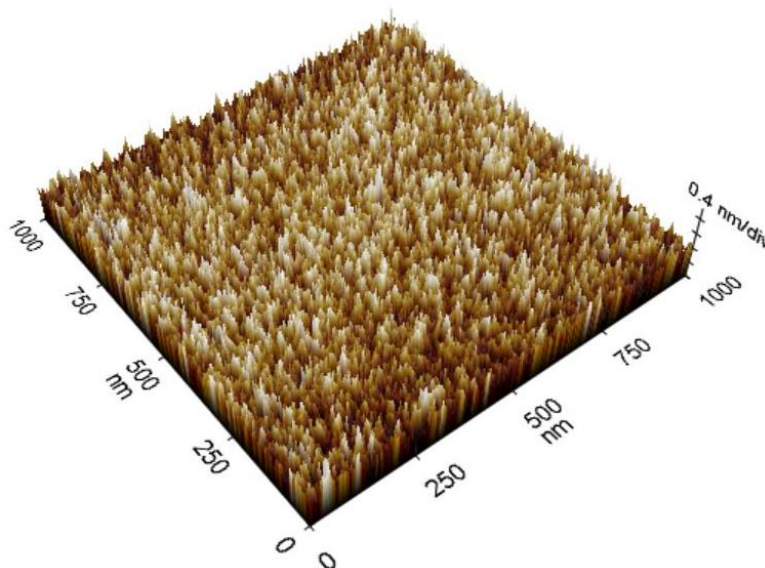


Figure 2.7: AFM image of an as-deposited HfO<sub>2</sub> film grown using the TALD process.

### 2.1.3 Growth of Aluminiumoxide ( $\text{Al}_2\text{O}_3$ ) and its film properties

Similar to  $\text{HfO}_2$ ,  $\text{Al}_2\text{O}_3$  films are also grown in the Pulsar<sup>®</sup> chamber using the TALD process at 300 °C. Trimethylaluminium ( $\text{Al}(\text{CH}_3)_3$ /TMA) is used as the precursor and water ( $\text{H}_2\text{O}$ ) is used as the co-reactant. TMA is a liquid precursor which has high enough vapor pressure at room temperature and does not require additional heating [91]. The vapors of TMA and  $\text{H}_2\text{O}$  are carried sequentially to the reaction chamber by using  $\text{N}_2$  as a carrier gas. Figure 2.8. shows the growth of an  $\text{Al}_2\text{O}_3$  monolayer on Si substrate by using an ALD cycle.

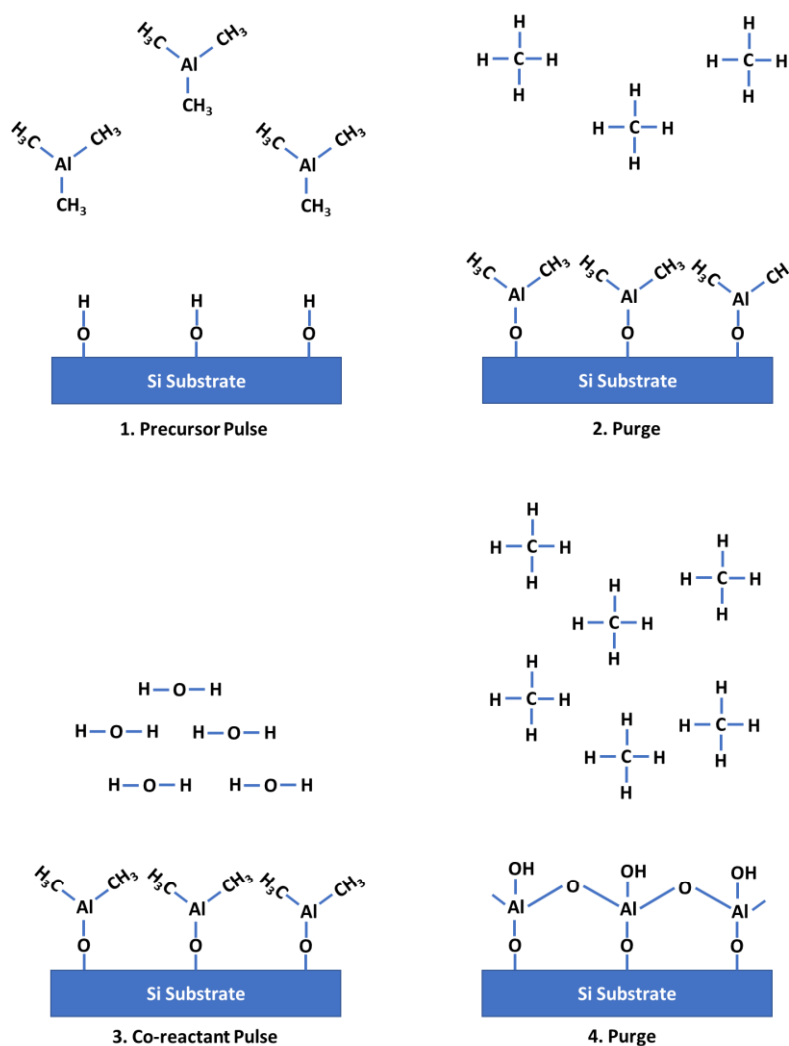
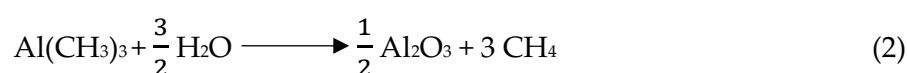


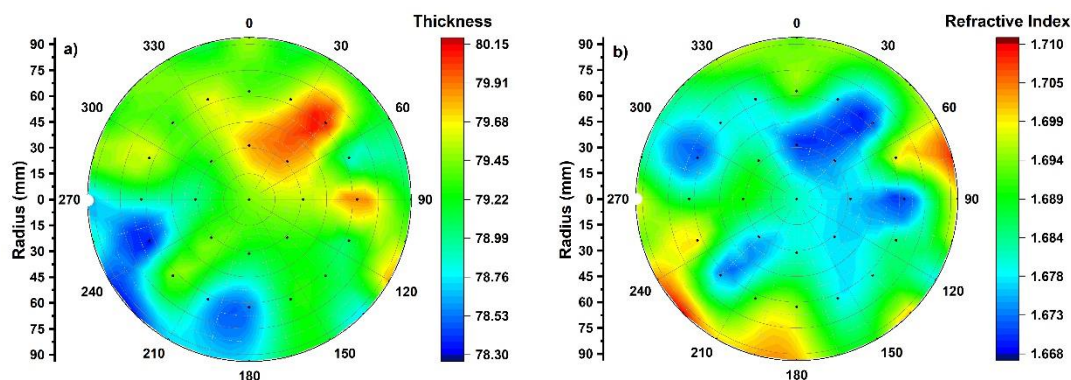
Figure 2.8: Schematic illustration of an ALD cycle used for the growth of  $\text{Al}_2\text{O}_3$  films by using the TALD process. The surface reactions of the precursor ( $\text{Al}(\text{CH}_3)_3$ ) and the co-reactant ( $\text{H}_2\text{O}$ ) are represented in two half-cycles. The bond lengths and the bond angles are not drawn to the scale.

Initially, TMA precursor is dosed into the reaction chamber. It reacts with the surface  $\text{OH}$  groups present on the silicon wafer. In this ligand exchange reaction, one

methyl group of TMA reacts with a OH group by releasing methane ( $\text{CH}_4$ ) as gaseous by-product. A sub-monolayer of  $\text{Al}_2\text{O}_3$  consisting of O-Al bonds is formed at the end of precursor half-cycle [92]. A subsequent  $\text{N}_2$  purge step removes the unreacted TMA and  $\text{CH}_4$  by-product from the reaction chamber, which completes the first half-cycle [92]. Subsequently,  $\text{H}_2\text{O}$  co-reactant is dosed into the reaction chamber which substitutes the remaining methyl ligands. Methane is released as a by-product and the surface is terminated with OH groups again, on which the next chemisorption of TMA can take place. Finally,  $\text{CH}_4$  and the unreacted  $\text{H}_2\text{O}$  are removed from the reaction chamber through a  $\text{N}_2$  purge step. The two half-cycles are repeated until the desired thickness of  $\text{Al}_2\text{O}_3$  layer is obtained. The net reaction between TMA and  $\text{H}_2\text{O}$  is represented as:



The  $\text{Al}_2\text{O}_3$  layers deposited by using the TALD process exhibited a growth rate of  $\sim 0.9 \text{ \AA}/\text{cycle}$ . The thickness and refractive index wafer maps of the  $\text{Al}_2\text{O}_3$  film are as shown in Figure 2.9. The standard deviation of the films was found to be less than 1% ( $1\sigma$ ) for both thickness and refractive index. The measured refractive index of the  $\text{Al}_2\text{O}_3$  films was 1.68. The TEM image shown in Figure 2.10 illustrates the growth of  $\text{Al}_2\text{O}_3$  films on a 3D feature fabricated in a Si-substrate. The conformality of the deposition was evaluated with respect to the step, bottom and side wall coverage of the  $\text{Al}_2\text{O}_3$  layers, which is found to be greater than 97%. Figure 2.11 illustrates the elemental composition of the  $\text{Al}_2\text{O}_3$  films examined by using XPS depth profile technique. The atomic concentrations of Al, O, N, C and Si were measured. The ratio of O/Al atomic concentrations is found to be  $\sim 1.5$  indicating the deposited films are stoichiometric. The roughness of the  $\text{Al}_2\text{O}_3$  films was investigated by using AFM. Figure 2.12 shows the surface topography of the  $\text{Al}_2\text{O}_3$  films in 3D and the RMS value of the measured roughness was 0.19 nm.



**Figure 2.9:** Wafermap of the (a) thickness and (b) the refractive index of TALD grown 8 nm  $\text{Al}_2\text{O}_3$  layer on a 200 mm Si-substrate.

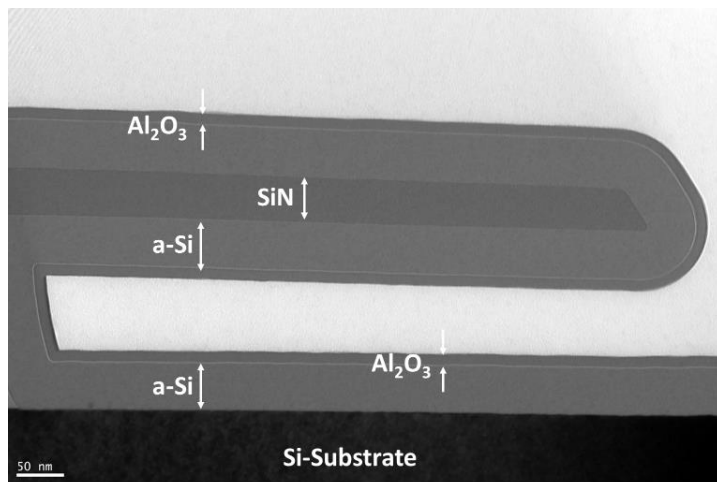


Figure 2.10: TEM image of an Al<sub>2</sub>O<sub>3</sub> film grown using the TALD process on a 3D structure fabricated in a silicon substrate.

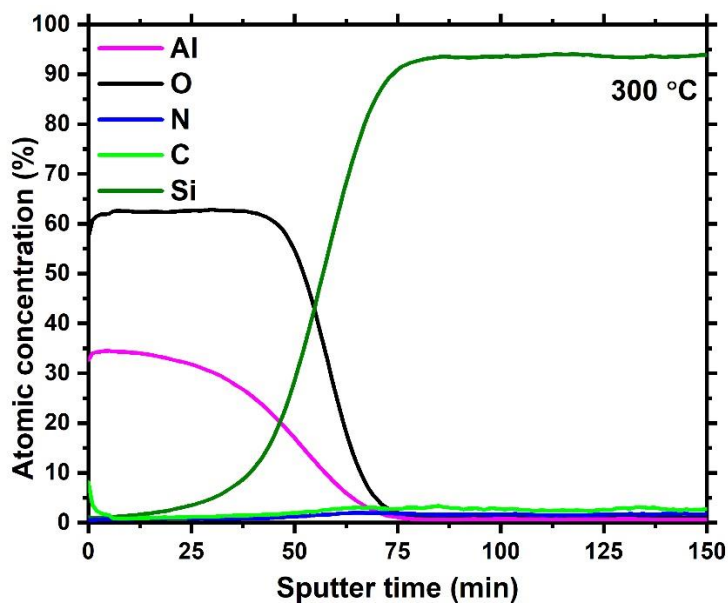


Figure 2.11: XPS depth profile of an 8 nm Al<sub>2</sub>O<sub>3</sub> film grown using the TALD process.

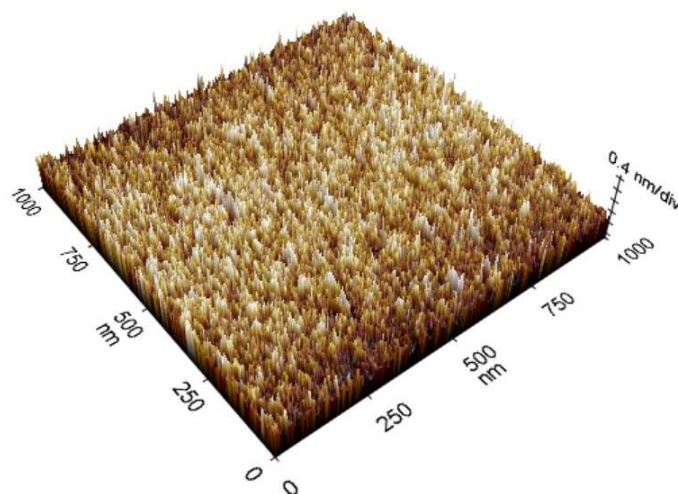
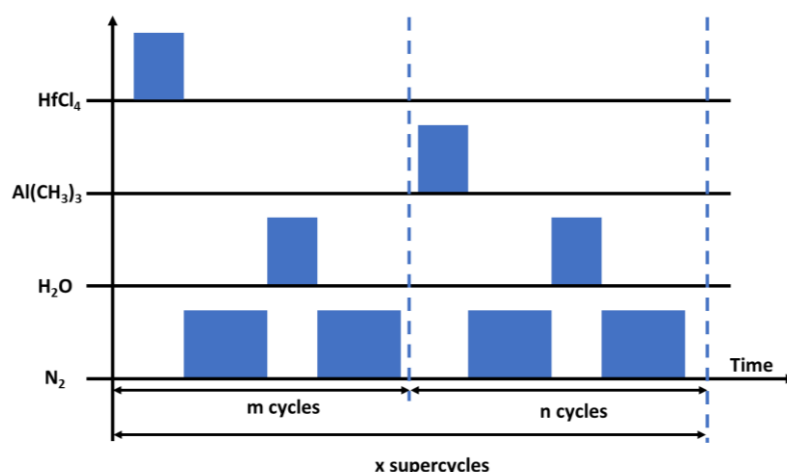


Figure 2.12: AFM image of an as-deposited Al<sub>2</sub>O<sub>3</sub> film grown using the TALD process.



## 2.1.4 Growth of HfAlO and its film properties

Besides the deposition of single layer dielectric films described in section 2.1.2 and 2.1.3, TALD can also be used to deposit their mixtures with specific doping concentrations or in the form of multi-layer dielectric stacks [71]. Aluminum doped HfO<sub>2</sub> layers are deposited in a multistep process by alternating the HfO<sub>2</sub> and Al<sub>2</sub>O<sub>3</sub> cycles sequentially within an ALD super-cycle [93]. The pulse-scheme diagram of a typical ALD super-cycle used for the deposition of Al doped HfO<sub>2</sub> (hereafter referred as HfAlO) layers is illustrated in Figure 2.13. It starts with m number of cycles for the deposition of HfO<sub>2</sub> layers followed by n number of cycles for the deposition of Al<sub>2</sub>O<sub>3</sub> layers. The number of cycles m and n of an ALD supercycle are adjusted depending on the desired concentration of Al within the HfO<sub>2</sub> layers, and are repeated x number of times until the desired thickness of HfAlO layers are obtained [93].



**Figure 2.13:** Schematic illustration of a typical ALD supercycle represented using a pressure versus time pulse-scheme diagram, for the deposition of Al doped HfO<sub>2</sub> layers. Adapted from [71].

Figure 2.14 shows the wafer map of thickness and refractive index of HfAlO layers deposited using 3 cycles of HfO<sub>2</sub> and 1 cycle of Al<sub>2</sub>O<sub>3</sub> in an ALD supercycle. The layers demonstrated a growth rate of ~0.7 Å/cycle. Further, they exhibited less than 1.5 % (1 $\sigma$ ) standard deviation with respect to both the thickness and refractive index. The addition of Al into the HfO<sub>2</sub> layers reduced the effective refractive index of the HfAlO layers. Figure 2.15 shows the linear decrease in the refractive index of the HfAlO layers with respect to increase in the Al doping percentage, while keeping the total number of cycles constant. The percentage of Al doping is determined by equation 2.1:

$$\frac{n \text{ cycles}}{(m + n) \text{ cycles}} \times 100 \quad 2.1$$

This characteristic feature of change in the refractive index of the HfO<sub>2</sub> layers with the incorporation of Al, assists in monitoring and controlling the quality of the deposited layers for a stable technology approach.

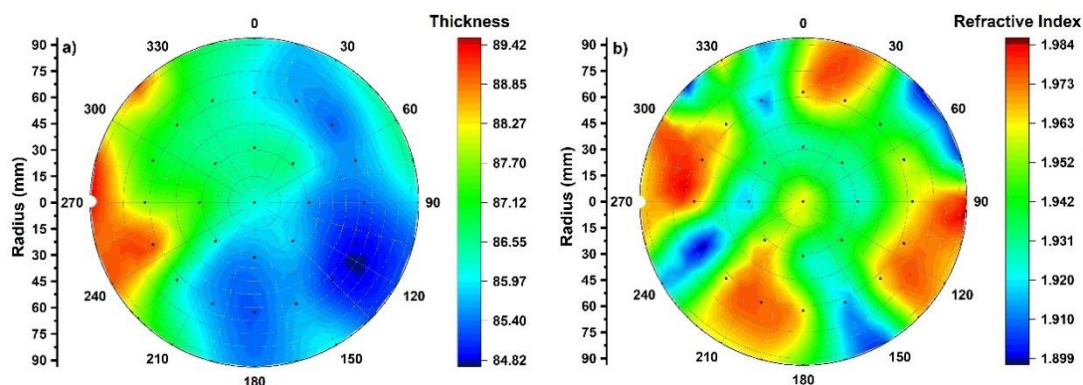


Figure 2.14: Wafermap of (a) the thickness and (b) the refractive index of TALD grown 9 nm HfAlO layers on a 200 mm Si-substrate.

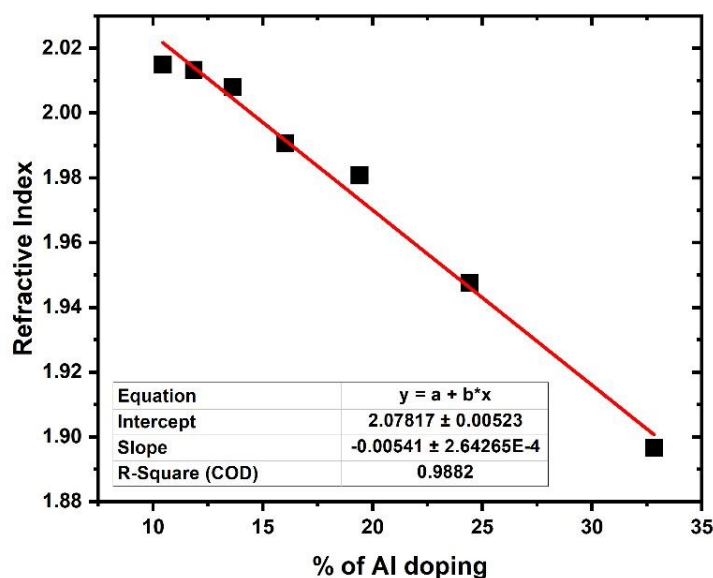


Figure 2.15: Plot of refractive index versus the Al doping percentage, with a linear fit. The investigation was performed on HfAlO layers of similar thickness but with different Al doping.

## 2.2 Electrical Characterization

The memristive devices under the framework of this work are characterized by using two different tools mainly Keithley 4200-SCS and RIFLE SE. The quasi-static measurements on 1T-1R and 1R single devices are performed on a PMV200 probe station using Keithley 4200-SCS characterization tool. The pulsed measurements on

4 kbit memristive arrays are performed on a PA200 probe station using RIFLE SE system.

### 2.2.1 Electrical Measurement Setup of Keithley 4200A-SCS

The measurements are performed by using a Keithley 4200A-SCS semiconductor parameter analyzer (SPA) connected to four vacuum type prober needles of PMV200 through a Keithley 4200A-CVIV multi-switch. The chuck contact of PMV200 probe station is connected directly to the SPA. The Keithley multi-switch module allows switching between I-V and C-V measurements without the necessity to re-cable or lift the prober needles. The Keithley SPA is configured for five semiconductor measurement units (SMUs) in total. The four probers take four SMUs and the fifth SMU is for the chuck contact. The SMU is one of the basic modules of the SPA. It can be operated as a voltage or current source and at the same time it can be used for measuring the current or voltage values. In this work, the SMUs are operated as a voltage source and used for sensing current. The SMU operation takes place in the form of source-delay-measure (SDM) cycles [94]. Once the source output is set, the delay phase will allow the source and the device under test (DUT) to settle before the measurement is performed. The Clarius software serves as a graphical user interface (GUI) for the measurements and the analysis performed using the setup.

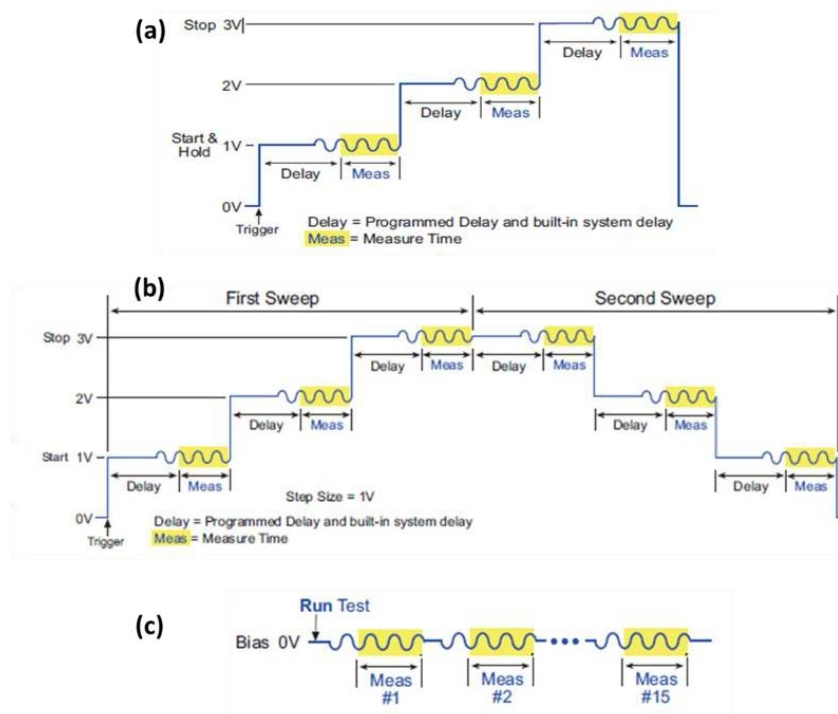
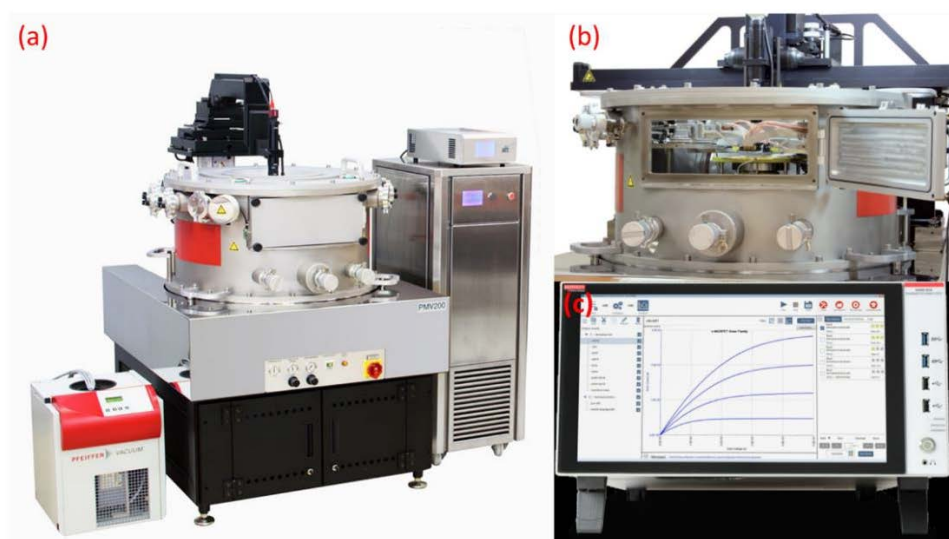


Figure 2.16: DC signal waveforms generated by the SMU which are utilized for the electrical characterization of 1T-1R and 1R devices: (a) Linear staircase sweep; (b) dual voltage sweep; (c) DC bias waveform. Adapted from [94].

The I-V measurements on 1T-1R and 1R devices are mainly performed by using linear staircase DC voltage sweep consisting of SDM cycles which is illustrated in Figure 2.16a. In case of 1R devices only two SMUs are used, one each for top and bottom electrodes. The current compliance for 1R devices is given through Keithley. The I-V measurements on 1R devices are performed using dual voltage sweep signal which is depicted in Figure 2.16b. In case of 1T-1R single devices, four SMUs are used, one each for gate, source, bulk and drain terminals. The transistor limits the current through the memristive device. Hence, providing current compliance through Keithley is not necessary. Similar to 1R devices, the 1T-1R devices require dual voltage sweep but additionally, a DC bias signal as shown in Figure 2.16c is required for the operation of the gate terminal of the transistor. The DC bias signal can be set to any valid bias level. For the DC dual voltage sweep, the start voltage, stop voltage and step voltage values have to be specified.

The PMV200 is a manual probe station which is used for placing the samples under test. It has a wide range of measurement capabilities including direct current (DC), radio frequency (RF), microelectromechanical systems (MEMS) and opto measurements. The measurement setup of PMV200 is shown in the Figure 2.17. The chuck is suitable for testing the devices on small substrates and on wafers up to the size of 200 mm. The temperature of the chamber can be varied from -60 to 300 °C. The temperature of the chuck is regulated by heating the chuck and simultaneously cooling from the chiller. The tool has a vibration isolation frame to protect against the vibrations during measurements. The measurements can be performed under high vacuum down to  $10^{-5}$  mbar.



**Figure 2.17: Quasi-static measurement setup: (a) SUSS MicroTec PMV200 manual probe station along with its vacuum pump and chuck cooling unit; (b) chuck stage unit with its manipulators in X, Y and Z directions; (c) Keithley 4200A SCS.**

## 2.2.2 Electrical Measurement Setup of RIFLE SE

RIFLE-SE test equipment was developed by Active Technologies to test stand alone or embedded non-volatile memories. It allows to investigate the behavior and reliability of the non-volatile memories in the form of single cells and also in the form of arrays. The RIFLE-SE has wide range of applications starting from product development, qualification, product life-time to its failure analysis.

The pulse measurements on the 4 kbit memristive arrays were performed using the pulse measuring unit (PMU) of the RIFLE-SE. The PMU unit consists of eight independent channels with 70 MHz sampling rate which provides high speed current and voltage measurements. The pulse measurements were performed by forcing the voltage as an input in the pulse mode and measuring the current as an output or vice versa. Further, it is possible to perform program and erase operations on the arrays using the test flows on the corresponding LabView software. The test flows can be easily edited and new flows can be configured through LabView block diagrams. Additionally, the LabView software allows to configure and execute the current maps and distributions. The nominal values of the pulse width for the RIFLE-SE system ranges from 10 ns to 10  $\mu$ s.

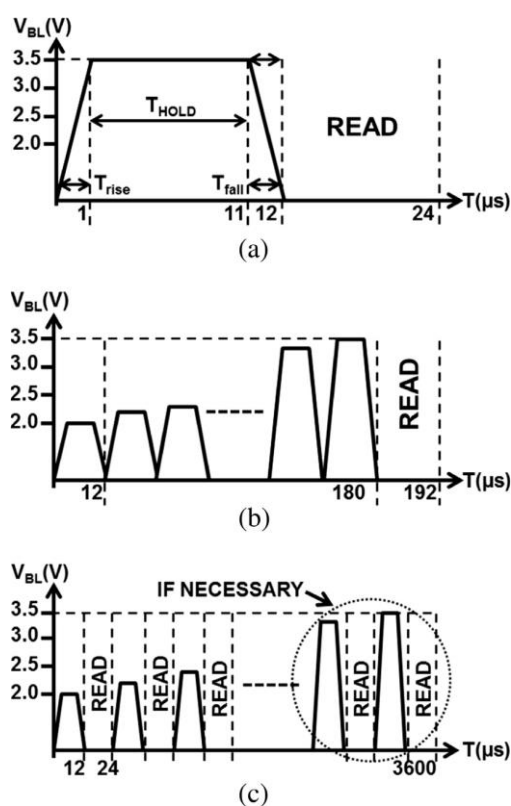
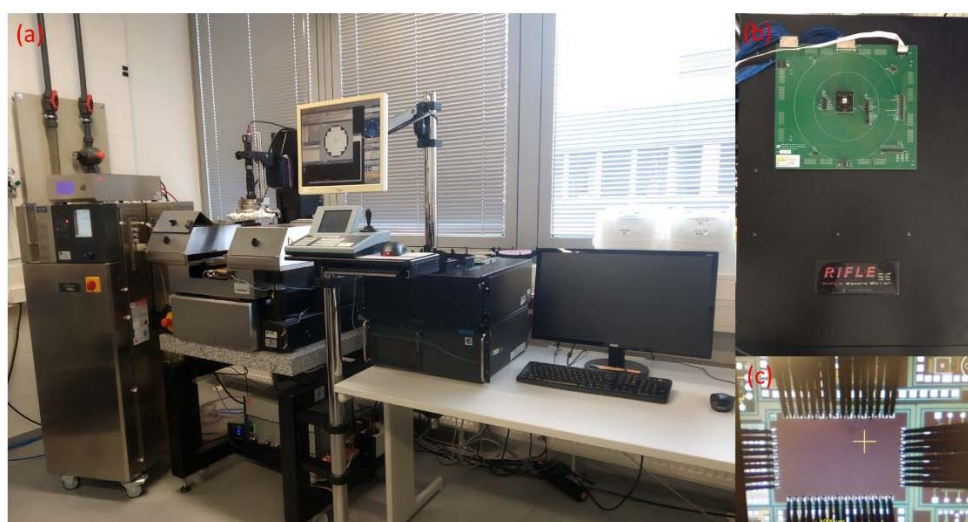


Figure 2.18: The pulsing schemes programmed using RIFLE-SE system (a) single pulse (b) incremental step pulse (c) incremental step pulse with verify. Adapted from [95].

The RIFLE-SE system allows to program different pulsing schemes illustrated in Figure 2.18. A single pulse (SP) scheme consists of a pulse with a specified amplitude and width along with a specific rise and fall time of the pulse in order to avoid the overshoot effects [95]. The incremental step pulse (ISP) scheme consists of a sequence of single pulses with increasing voltage amplitudes and constant pulse width. Further, the incremental step pulse with verify (ISPV) scheme is similar to the ISP scheme but in addition it consists of a current read-out step after every pulse. The current read-out operation generally takes place at lower pulse amplitudes such as 0.2 V in order not to change the state of the device. Finally, the minimum pulse width required for the read-out operation is 1  $\mu$ s.

The 4 kbit memristive array chips on 200 mm wafers can be characterized on PA200 probe station. It is suitable for RF, Failure Analysis (FA), opto-engineering and MEMS measurements. The electrical setup of the probe station is as shown in Figure 2.19. The tool requires a manual loading of the 200 mm wafer on the chuck. The temperature within the probe station can be regulated by using the heating features of the chuck and the cooling of chiller from ATT Systems and can be varied from -60 to 300 °C. The chuck can be easily navigated using a joystick controller. The joystick consists of a color-coded display for the full prober control. The Velox™ probe station control software allows to create a wafer map by utilizing the available information on the die and wafer sizes. Additionally, the Velox™ software provides screen navigation and automation, and it works in synchronization with analyzers and measurement software.



**Figure 2.19:** Pulsed measurement setup for the electrical characterization of 4 kbit memristive arrays consisting of (a) the Cascade Microtech PA200 semi-automatic probe station along with its chuck cooling unit, and (b) the measurement setup for packaged devices: the RIFLE SE. (c) Shows the 4 kbit array on a wafer placed on the setup for measuring the devices.

---

The probe card consists of fifty ad-hoc fingers which are mounted on the probe shield from Cascade Microtech. The measurements are performed by using the RIFLE-SE system from Active Technologies connected to the PA200 probe shield. In order to measure packaged chips, devices are placed in the socket on the device interface board (DIB) from NplusT which is in turn connected to the RIFLE-SE system through cables.

## 2.3 Summary

In this chapter, two main parts of the experimental work carried out under the framework of this thesis were presented in detail. In the first part of the experimental work, the TALD process used for the deposition of dielectric films which are used as memristive switching layers in this work is discussed. In the second part, the two different parameter analyzers and the probe stations used for the electrical characterization of the 1T-1R, 1R single devices, and integrated 4 kbit memristive arrays are presented.

The deposition process of single-layer  $\text{HfO}_2$  and  $\text{Al}_2\text{O}_3$  layers takes place through ALD half-cycles. The chemical precursors and the co-reactant gases are introduced into the reaction chamber sequentially at different times. The dielectric films are grown layer-by-layer through self-terminated surface reactions. Further, the  $\text{HfAlO}$  layers are grown by utilizing both  $\text{HfO}_2$  and  $\text{Al}_2\text{O}_3$  cycles in an ALD supercycle. The uniformity of thickness and refractive index of the deposited dielectric layers are investigated by using a spectroscopic ellipsometer. The conformality of the dielectric layers is investigated by depositing them on a 3D feature fabricated in a silicon substrate. Further, the stoichiometry of the layers is investigated through XPS depth profile analysis and, the surface roughness of the layers is analyzed through the AFM technique. Finally, the dielectric layers deposited through the TALD process are found to be uniform with less than 1.5 % standard deviation, conformal, stoichiometric, and smooth with an RMS value of roughness less than 0.2 nm.

Two different parameter analyzers were used to electrically characterize the memristive devices. The 1T-1R and 1R single devices are characterized using quasi-static DC measurements. A setup consisting of Keithley 4200A-SCS SPA connected to a PMV200 probe station to place the samples is used for the purpose. The SMU module of the SPA is used for the DC signal waveform generation. Mainly the DC bias and the dual voltage sweep signal waveforms are used for the electrical characterization of single devices. The 4 kbit memristive arrays are characterized using the pulsed measurement technique. The measurements were carried out by using the RIFLE-SE

system which is connected to a PA200 probe station to place the samples. Further, the RIFLE-SE system generates the pulses required for the characterization through its PMU module. The incremental step pulse with a verify scheme is used for the characterization of the 4 kbit arrays.



## Chapter III

---

### 3 Process Development of Memristive Devices

---

Memristive devices are regarded as one of the potential candidates for emerging non-volatile memory applications [14]. Despite their characteristic features mentioned in chapter 1, the performance of the devices, the device-to-device (D2D), and the cycle-to-cycle (C2C) variabilities remain a challenge [96]. The factors influencing the variability in the memristive devices are mainly categorized into two groups: extrinsic, and intrinsic. The extrinsic variabilities arise mainly due to the fabrication processes. The discrete nature of the atoms and the randomness associated with the atomistic defects causes intrinsic variabilities [97]. This chapter focuses on the process development of the memristor module with high-quality layers and the narrow window of process variations, thereby targeting the extrinsic factors that can limit the memristive device performance.

The memristive device in this work consists of a metal-insulator-metal (MIM) stack with titanium nitride (TiN) as top and bottom electrodes, titanium (Ti) as an oxygen scavenging layer (OSL), and the HfO<sub>2</sub>-based memristive switching layers. Patterning the MIM stack using a reactive-ion-etching (RIE) process is a crucial step as it could oxidize the Ti layer from the sides, and the impurities from the RIE process, if they reside on the devices, could degrade their performance and reliability. This issue is addressed in this work by patterning the MIM stack using three different fabrication approaches and finally covering the entire stack with Si<sub>x</sub>N<sub>y</sub>O<sub>z</sub> encapsulation layers.

The entire layer stack consisting of MIM and encapsulation layers hereon will be referred to as the MEMRES module.

To convert the fabrication process of the MEMRES module into a stable technology, the quality of the devices has to be monitored and controlled. This is achieved in this work through the electrical characterization of the dedicated process control monitor (PrCM) test structures.

Finally, the HfO<sub>2</sub>-based memristive devices without the CMOS transistors which are integrated into the BEOL of the 250 nm BiCMOS technology of IHP are investigated for their resistive switching performance. Initially, the forming, reset and set operations, followed by 50 subsequent cycles of reset and set operations are carried out on the devices. The electrical measurements are performed on a wafer-scale level and the parameters extracted from the I-V curves are analyzed and interpreted.

*“The investigated samples in this work were fabricated by the cleanroom staff at the IHPs pilot line. The patterning of the memristive devices and their automated electrical measurements after the fabrication was carried out by co-authors Mirko Fraschke and Detlef Schmidt, respectively. I fabricated the memristive switching layers, provided the parameters for automated electrical measurements, and analyzed the data. Parts of this chapter have been published in Ref. [98].”*

### **3.1 Optimization of memristor module fabrication process**

The MEMRES module is integrated between metal 2 (M2) and metal 3 (M3) in the BEOL interconnects of a 250 nm BiCMOS baseline technology at IHP. To integrate any module into a baseline technology, the layers used for its fabrication must be CMOS compatible. Therefore, the fabrication of the MEMRES module and its optimization techniques are discussed in this chapter for the TiN/HfO<sub>2</sub>/Ti/TiN stack consisting of CMOS-compatible layers.

Titanium Nitride (TiN) is widely used in the microelectronics industry as an electrode material for capacitors, memristive devices, ferroelectric devices, etc. The inertness toward chemical reactions, relatively large work function, low resistivity, relatively high thermal stability, and high selectivity during the RIE process are the main reasons behind its wide usage [99]. Titanium (Ti) is mainly used in the microelectronic industry as a liner layer for the via structures and as an OSL in the memristive devices [96], [100]. Due to the fully stoichiometric memristive layers and the inert TiN electrodes, it is not feasible to create a conductive filament (CF) through the soft breakdown process in a memristive device. Hence, an active metal layer such

as Ti is required in the module which acts as an oxygen vacancy reservoir for creating and destroying the CF in the memristive switching layer during the set and reset operations, respectively [52]. A 150 nm thick TiN is used for the top and bottom electrodes. A Ti layer of thickness 7 nm is used for OSL. The Ti and TiN layers are deposited using physical vapor deposition (PVD).

Hafnium oxide ( $\text{HfO}_2$ ) is one of the most commonly used high-k dielectric materials for emerging non-volatile memory devices. As mentioned in Chapter 3, the simplicity of its deposition process, uniformity and conformality of the layers, CMOS compatibility, deposition temperatures being suitable with the BEOL thermal budget etc. are some of the reasons for its widespread usage [71]. A  $\text{HfO}_2$  layer with a thickness of 8.5 nm, deposited using TALD, is used as a switching layer in the MEMRES module.

The fabrication of the MEMRES module begins with the deposition of the aforementioned layers in the form of a MIM stack. The deposited layers are subsequently patterned using lithography and RIE. The patterning of the layers is performed in three different ways, namely type 1, type 2 and type 3. The schematic illustration of their process flows is shown in Figure 3.1 and their corresponding flow chart of the process steps involved is shown in Figure 3.2.

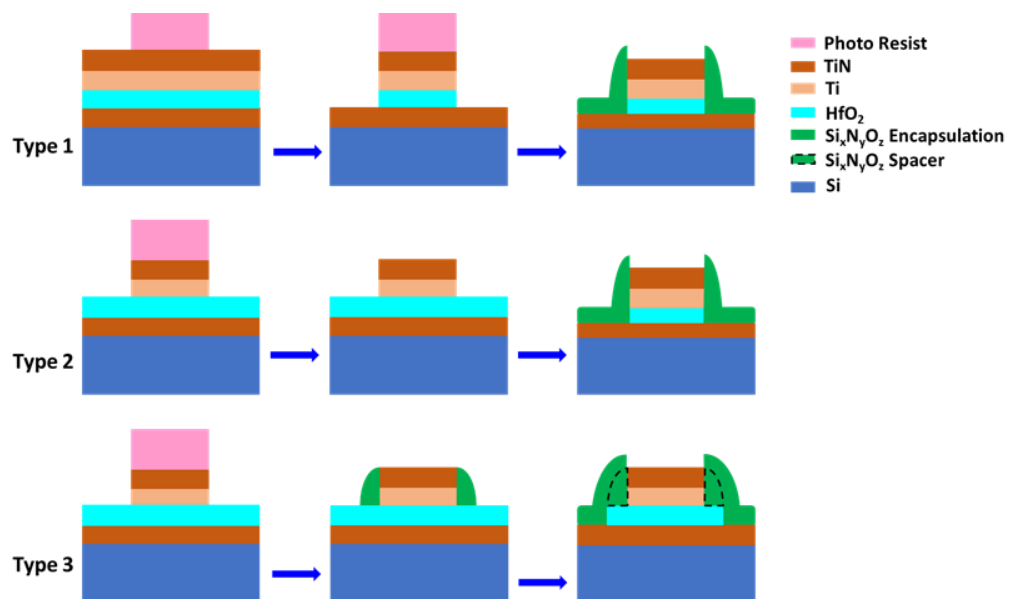
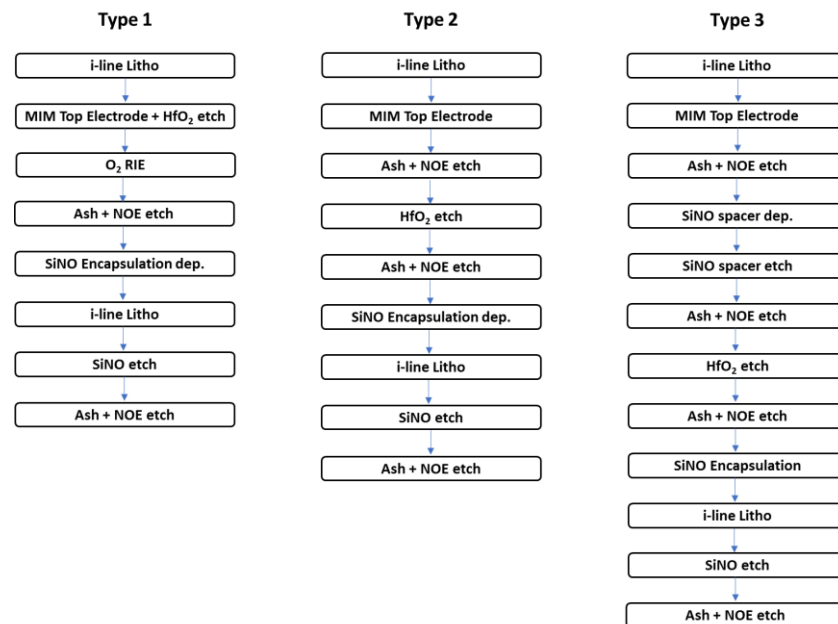


Figure 3.1: Schematic illustration of the three short loop process flows of the memristor module fabrication [98].

The layers in the type 1 approach of the MEMRES module are patterned by using photoresist as the mask. The TE and the  $\text{HfO}_2$  layers of the module are subsequently etched using the RIE process in one step by using  $\text{Cl}_2$  and  $\text{BCl}_3$  chemistry, respectively. Similar to type 1, the TE in the type 2 approach is etched by using photoresist as the

mask. However, the photoresist is removed after the TE etch, and the  $\text{HfO}_2$  layer is etched by using the TiN TE as the hard mask. The memristor module patterning in type 3 is performed similar to type 2 but before etching the memristive switching layer, the side walls of the TE are protected by  $\text{Si}_x\text{N}_y\text{O}_z$  spacers. The spacers are created by conformal deposition of  $\text{Si}_x\text{N}_y\text{O}_z$  layers on the patterned TE and subsequently dry etching them without any mask. Due to the anisotropic etching behavior of the RIE process, the  $\text{Si}_x\text{N}_y\text{O}_z$  layers on the surfaces are removed except for the sidewalls. Further, the  $\text{HfO}_2$  layer is etched by using the TiN hard mask with its side walls protected. Soon after the patterning of the memristive switching layer, the MEMRES modules of all three approaches are protected by a  $\text{Si}_x\text{N}_y\text{O}_z$  encapsulation layer deposited using chemical vapor deposition (CVD). Finally, the  $\text{Si}_x\text{N}_y\text{O}_z$  encapsulation layer and the BE are patterned by using the same photoresist mask but different etch chemistries. The RIE process of the  $\text{Si}_x\text{N}_y\text{O}_z$  layer uses  $\text{CHF}_3/\text{CF}_4$  chemistry while the TiN BE layer uses standard  $\text{Cl}_2$  chemistry. The  $\text{Si}_x\text{N}_y\text{O}_z$  on the TE is etched along with the via hole etching step of the interlayer dielectric two (ILD2). The TE connection of the memristive device is made from the metal 3 (M3) layer through via 2 (V2) contact while the BE connection is made from the metal 2 (M2) layer. As illustrated in the process flow chart of Figure 3.2, the fabrication of the MEMRES module in all three types of approaches until the BE patterning requires two lithography steps [98].



**Figure 3.2:** The flow chart of the process steps involved in the three different fabrication approaches of the memristor module [98].

The RIE process is one of the key steps in MEMRES module fabrication. The oxidizing Ti layer and the thin  $\text{HfO}_2$  layers make the etch process challenging. Additionally, it creates polymer and ionic residuals which could reside on the device

surfaces and degrade their normal functionality. The RIE process steps in all three approaches are immediately followed by a remote plasma ashing step which removes the photoresist and the polymer residuals. However, the type 1 approach has an oxygen RIE plasma step included before the ash step. This additional step breaks down the hardened photoresist which in turn assists the photoresist removal in the ashing step [101], [102]. Each photoresist removal process is immediately followed by a standard neutral oxide etch (NOE) step which removes the polymeric and ionic impurities [98].

### 3.2 Dedicated test structures of memristor module

The performance of the MEMRES module fabricated using the three different approaches is evaluated by using two different test structures, namely class A and class B. Figure 4.3 shows the schematic illustration of both test structures. Class A consists of device dimensions of  $75\ \mu\text{m} \times 75\ \mu\text{m}$  with four corners and  $300\ \mu\text{m}$  of device perimeter. It is mainly used to evaluate the quality of the layers within the MEMRES module [98].

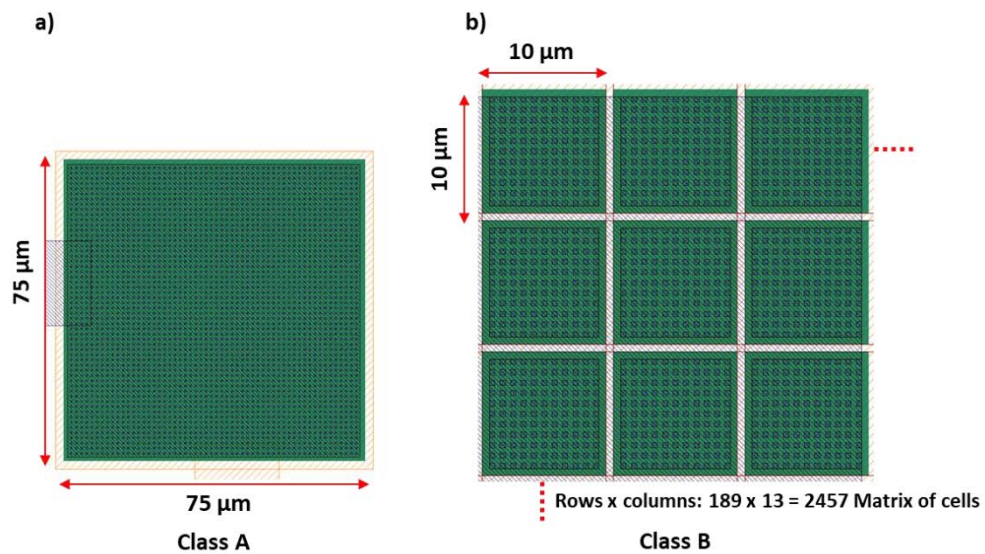


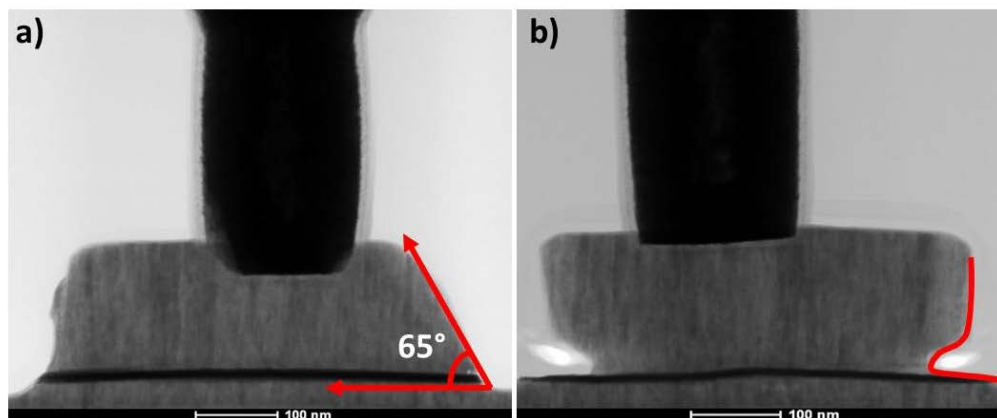
Figure 3.3: Schematic illustration of (a) class A and (b) class B test structures [98].

Class B consists of a device layout of 2457 small square structures, each of dimensions  $10\ \mu\text{m} \times 10\ \mu\text{m}$ . Together they form a matrix of cells which are in total composed of 9828 corners and a  $98,280\ \mu\text{m}$  device perimeter [98]. Due to the larger perimeter and number of corners compared to class A, the device is used to evaluate the sensitive memristor module topographies in the course of integrating the module into a SiGe BiCMOS baseline technology. Further, the class B devices demonstrate the

quality of the processes involved in the wafer-scale fabrication of the module, especially the patterning and cleaning steps. The high density of devices connected in parallel makes them suitable structures to investigate the influence of the fabrication process variations [98]. Both, class A and class B devices are PrCM test structures dedicated to the MEMRES module. The quality of the MEMRES module can be monitored and controlled by performing in-line electrical measurements on these dedicated PrCM test structures.

### 3.3 Anisotropic dry etching of MIM stack

RIE is a conventionally used plasma process in semiconductor manufacturing for the anisotropic removal of materials. It involves both chemical and physical etching mechanisms [103]. This anisotropic etching process is used for patterning the MIM stack of the memristive device. As shown in Figure 3.4, two different etching profiles were observed. The tapered profiles in Figure 3.4 (a) are formed due to the deposition of the sputtered photoresist on the sidewalls of the TiN electrode during the etching process. As a result, the TiN sidewalls are passivated and the width of the photoresist mask increases as the etching proceeds [104]. Further, this etch profile was observed in an Applied Materials research tool with a capacitively coupled plasma (CCP) source.



**Figure 3.4: Etching profiles of the MIM stack: (a) tapered etch profile; (b) notching at the metal-insulator interface. The etch profiles are highlighted in red.**

Notching at the TiN top electrode and the memristive switching layer interface is shown in Figure 3.4 (b). Normally the etch process is carried out in two steps. Initially a main etch step takes place with high bias power which removes the desired material comparatively faster. To achieve an etch uniformity over the entire wafer, an “overetch” step with low bias power is successively carried out after the main etch

[105]. In the standard MIM patterning, the overetch step removes the remaining top electrode with minimum damage to the dielectric layer underneath [103]. However, this step could be eliminated in patterning the memristive devices since the etch selectively of TiN to memristive switching layers is relatively good. This could avoid the deflection of the positive ions towards the TiN sidewall thus preventing the notching effect [106]. Further, this etch profile was observed in an Applied Materials production tool with an inductively coupled plasma (ICP) source. The MIM patterning was performed using the research tool for the MEMRES process development, while the production tool was used for the fabrication of the memristive arrays.

### 3.4 Electrical characterization of the PrCM test structures

The electrical measurements on the PrCM test structures were performed at a wafer-scale level on a 200 mm wafer consisting of 87 dies. The characterization aims to investigate the influence of the fabrication processes and the layout designs on the memristor module.

Initially, the class A and B devices fabricated using the three different approaches are electrically characterized for their leakage currents. The leakage currents are measured by DC sweeping the voltage from 0 to 1 V in steps of 50 mV. As shown in Figure 3.5, the class A devices fabricated using the type 1 method show comparatively higher leakage currents in their pristine state. Large variability is also observed in the current values in the range of nano- to milliamperes at 0.1 V as illustrated in Figure 3.5 (d). In contrast, the devices fabricated using type 2 and type 3 methods exhibited relatively lower pristine state currents with strongly reduced variability in the order of nanoampere at 0.1 V.

The leakage current behavior of class B devices is as shown in Figure 3.6. It can be seen that the devices fabricated using the type 1 method exhibit very high leakage currents and mostly in the order of milliamperes. Unlike class A devices, 25% of class B devices fabricated using the type 2 method still exhibit high leakage currents in their pristine state. Whereas, all the class B devices fabricated using the type 3 method exhibited pristine state currents in the range of nanoampere. The pristine state current of the well-insulated memristive devices in class B is one order of magnitude higher compared to class A. The difference in leakage current levels observed between them is due to the sensitive device topography of class B devices. The higher density of class B devices compared to class A devices plays a significant role in patterning the MIM stack. Because, the RIE process step used for patterning the MIM stack varies based on the density of the features [107].

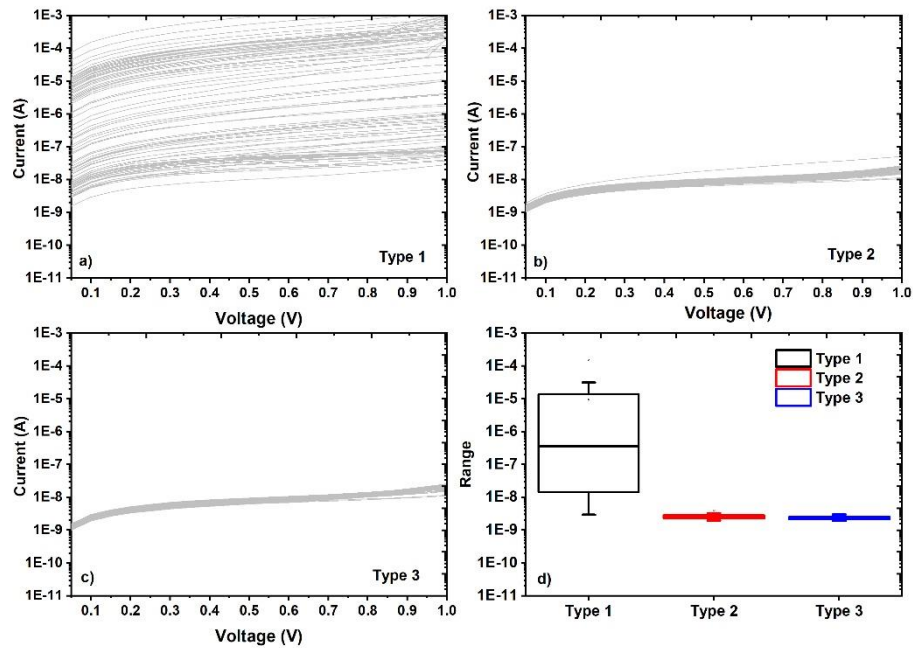


Figure 3.5: Wafer-scale leakage current measurement of class A devices which are fabricated by using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of the leakage currents of type 1, 2 and 3 devices measured at 0.1 V [98].

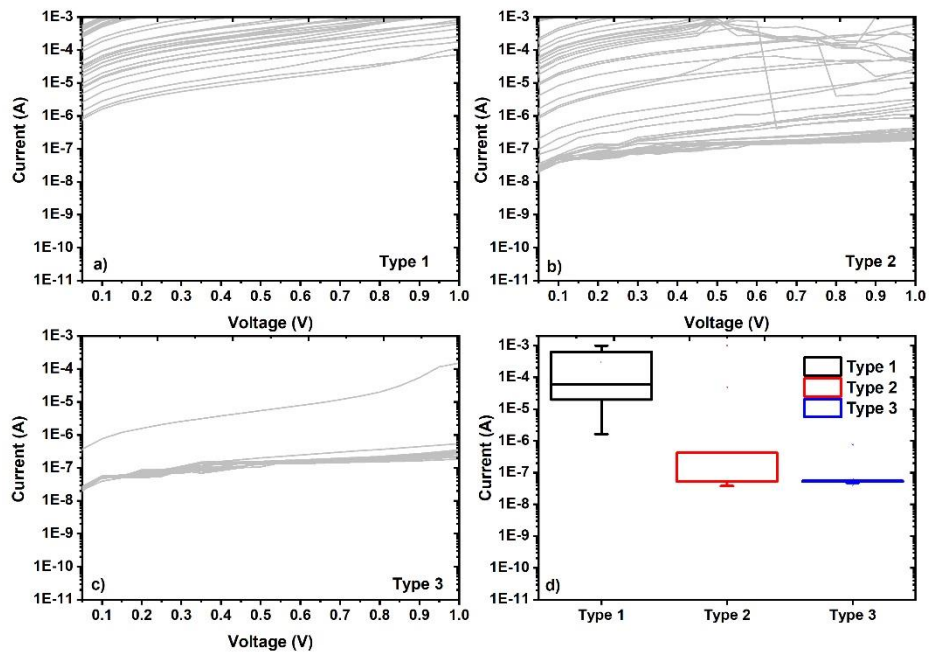
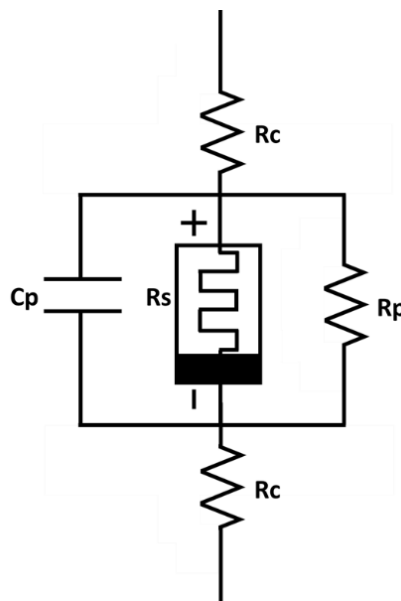


Figure 3.6: Wafer-scale leakage current measurement of class B devices which are fabricated by using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of the leakage currents of type 1, 2 and 3 devices measured at 0.1 V [98].

Further, an equivalent circuit of a memristive device ( $R_s$ ) is represented by a resistor ( $R_p$ ) and a capacitor ( $C_p$ ) in parallel as shown in Figure 3.7 [108]. The resistance of the TE, BE, and the contact resistance between the CF and the electrodes



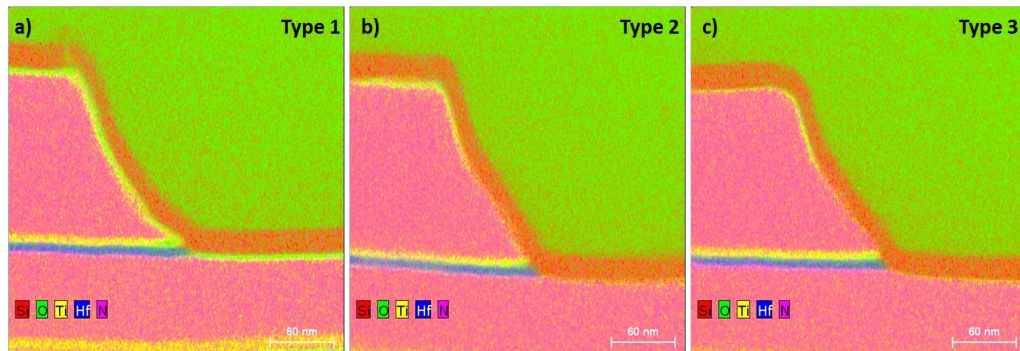
is represented by  $R_c$ .  $R_p$  accounts for the insulation and the dielectric losses between the top and bottom electrodes.  $C_p$  comes into the picture due to the MIM structure of the memristive device [108], [109]. The patterning of the MEMRES module using the type 1 and type 2 methods does not involve side wall protection of the TE before etching the memristive switching layer. These unprotected sidewalls can give rise to Ti-containing polymer residuals to reside on the dielectric sidewalls during its RIE process step. This leads to large pristine state currents through  $R_p$  in the equivalent circuit shown in Figure 3.7, which in the real case scenario are the sidewalls of the memristive devices. However, the type 3 fabrication approach uses enhanced sidewall protection of the devices through  $\text{Si}_x\text{N}_y\text{O}_z$  spacers which keeps the  $R_p$  value high and avoids leakage currents through the sidewalls.



**Figure 3.7: The schematic illustration of an equivalent circuit of a memristive device with its parasitic components. Adapted from [108].**

The impact of three different ways of the MEMRES module patterning is investigated in class A devices through TEM cross-section with energy dispersive X-ray (EDX) analysis as shown in Figure 3.8. In type 1 approach, the top and bottom electrodes of the memristive device are oxidized to form  $\text{Ti}_x\text{O}_y\text{N}_z$ . The  $\text{O}_2$  RIE ash step, which is only used in type 1, could be responsible for the oxidation of TiN electrode. Further, the Ti layer is also oxidized at the side wall. In the case of type 2 and 3, the  $\text{Ti}_x\text{O}_y\text{N}_z$  formation is slightly visible on the TEs and their tapered side walls. Even though the claimed Ti impurities which are responsible for the leakage current behavior are difficult to notice, the obtained electrical results in Figure 3.5 and Figure 3.6 support the statement. Further, the clear distinction between the  $\text{Si}_x\text{N}_y\text{O}_z$  spacer

and encapsulation in the type 3 approach is difficult to notice due to the usage of the same material and the low sidewall angles created by the RIE process step [98].



**Figure 3.8:** EDX-based chemical compositional maps of the side wall of the class A test structures fabricated using type (a) 1, (b) 2, and (c) 3 approaches [98].

Further, the class A and B devices are electrically characterized for their breakdown voltages at two different current densities  $1 \text{ pA}/\mu\text{m}^2$  and  $10 \text{ pA}/\mu\text{m}^2$ . The breakdown voltages are measured using a non-destructive measurement technique where the current is forced through the device and the voltage is measured across it. The class A and B devices fabricated using the type 1 method show comparatively lower breakdown voltages which corresponds with their higher leakage current behavior depicted in Figure 3.5 and Figure 3.6. Whereas, the class A devices fabricated using type 2 and type 3 methods exhibit very little difference in terms of their breakdown voltages. However, their breakdown voltage values are higher compared to type 1 devices as shown in Figure 3.9 (a) and (b). The improved breakdown voltage behavior in type 2 and 3 approaches (b) of class A devices demonstrate the quality of the layers fabricated within the MIM stack in the MEMRES module [98].

The influence of type 2 and type 3 fabrication approaches on the electrical performance of the devices is demonstrated in class B test structures. The mean value of the breakdown voltages of the class B devices fabricated using the type 2 approach is lower compared to the type 3 approach devices as shown in Figure 3.9 (c) and (d). However, the type 2 approach devices show large variability compared to type 3 approach devices. The devices with high pristine state currents exhibit low breakdown voltages. A comparatively higher breakdown voltage with less variability in high-density structures of class B devices patterned using the type 3 method implies reliable devices with improved fabrication quality [98].

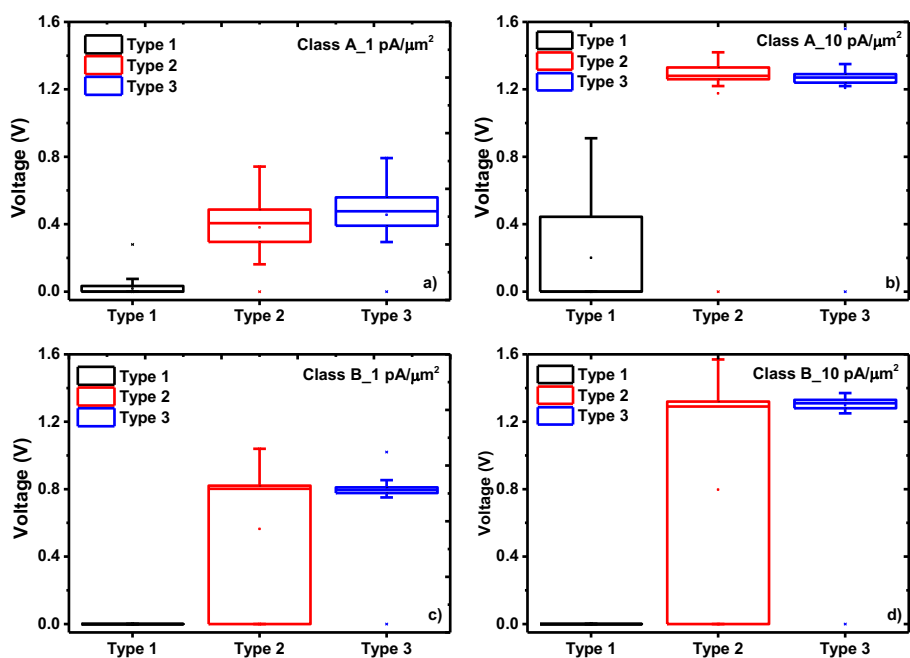


Figure 3.9: Wafer-scale breakdown voltage measurements of class A devices at (a)  $1 \text{ pA}/\mu\text{m}^2$ , (b)  $10 \text{ pA}/\mu\text{m}^2$  and class B devices at (c)  $1 \text{ pA}/\mu\text{m}^2$ , (d)  $10 \text{ pA}/\mu\text{m}^2$  fabricated using three different approaches [98].

Finally, the class A and class B PrCM test structures are measured for their capacitance values at 0 V and 100 kHz frequency. As shown in Figure 3.10 (a), the capacitance of class A devices fabricated using three different approaches is in a similar range. Whereas, the capacitance of class B devices fabricated using the type 1 approach was unable to measure due to the high leakage currents in their pristine state. The type 2 and type 3 approach fabricated class B devices demonstrate capacitance in the same range as shown in Figure 3.10 (b).

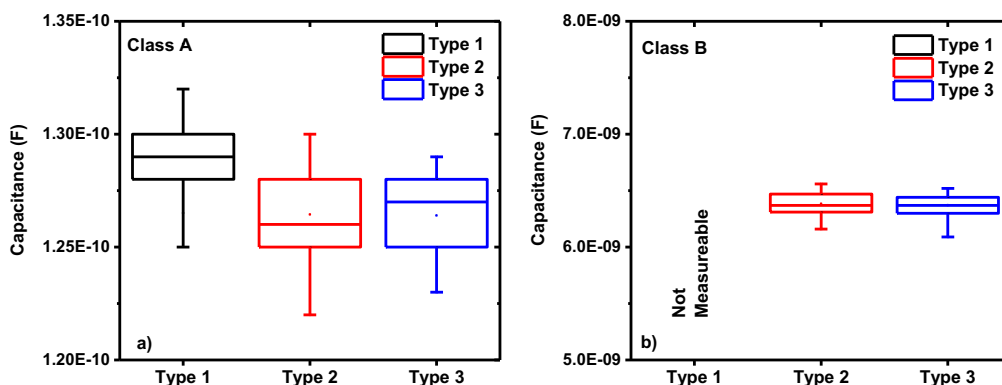


Figure 3.10: Capacitance measurements of (a) class A, (b) class B devices fabricated using three different approaches [98].

### 3.5 Electrical characterization of the memristive devices

The characterization aims to analyze the influence of three different fabrication approaches on the resistive switching operation of the memristive devices. The investigation is performed on class A test structures. The electrical measurements are carried out using a SPA connected to a Cascade Microtech probe station. The forming, reset and set operations are performed on the memristive devices using DC voltage sweeps. The current compliance (CC) required for the switching operations is provided through the SPA. The electrical parameters with their corresponding ranges are illustrated in Table 3. The forming operation is performed with a current compliance of 100  $\mu\text{A}$ . Owing to the fact that the forming operation determines the number of oxygen vacancies available for the subsequent memristive switching operations, a relatively lower CC value is chosen to achieve a higher value of resistance in the off-state of the memristive device [110]. A higher CC value of 100 mA is imposed for the reset operation to obtain a proper reset of the device. The 10 mA CC for the set operation is chosen based on the maximum value of the currents observed during the reset operation. After the initial operations, the devices are cycled for 50 subsequent reset and set operations. A typical illustration of an electrical measurement performed on the memristive device is shown in Figure 3.11. Similar to previous electrical measurements, the characterizations are performed on a wafer-scale level.

Table 3: DC sweep parameters for resistive switching operation of the memristive devices

Operation	Top Electrode (TE)	Bottom Electrode (BE)	Current Compliance
Forming	0 – 3 V	Gnd	100 $\mu\text{A}$
Reset	Gnd	0 – 0.8 V	100 mA
Set	0 – 0.8 V	Gnd	10 mA

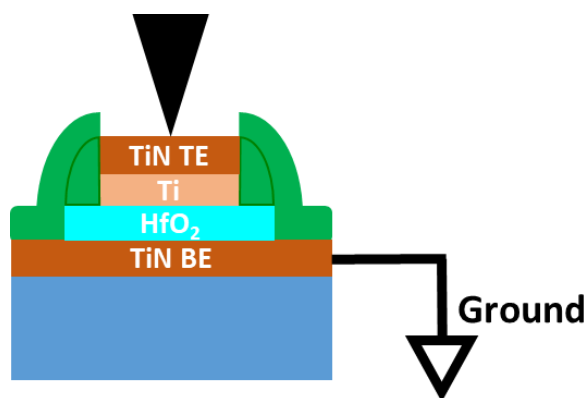
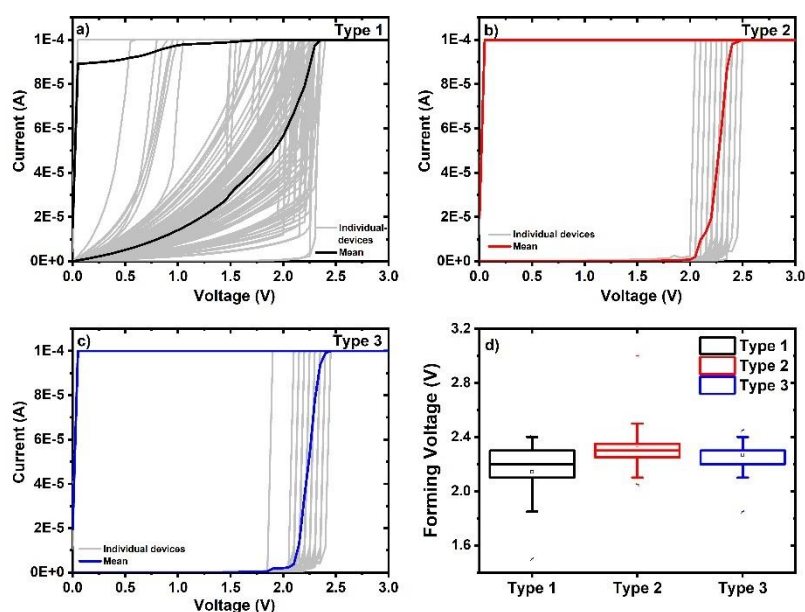


Figure 3.11: Schematic illustration of an electrical measurement of a memristive device.

The resistive switching operation of the memristive device begins with a forming step which creates a CF in the dielectric layer through its soft breakdown. Figure 3.12 shows the forming operation performed on the memristive devices which are fabricated using the three different fabrication approaches. It can be observed from the mean values of the forming I-V curves shown in Figure 3.12 (b) and (c) that the change in current during the forming operation is abrupt in type 2 and type 3 devices, while it is mostly continuous in case of type 1 approach devices. This D2D variability observed in the forming operation of type 1 devices is due their large pristine state currents. It can be directly related to their pristine state currents shown in Figure 3.5, since the initial leakage current level of the memristive devices determines the voltage at which the CF will be formed [111]. Although type 2 and 3 fabricated devices exhibit D2D variability, it is much lower compared to type 1 approach devices. Further, the distribution of the forming voltages is represented as box plots shown in Figure 3.12 (d). The forming voltages of type 1, 2, and 3 devices exhibit similar levels. However, type 1 devices demonstrate high variability, while type 3 devices demonstrate the lowest variability.



**Figure 3.12:** Wafer-scale forming I-V curves of class A devices fabricated using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of forming voltages of class A devices [98].

The forming step is followed by a reset operation which partially destroys the CF at the BE interface, thereby switching the device to the off state. Figure 3.13 shows the I-V curves of the first reset operation in type 1, 2, and 3 approach devices. Further, the memristive devices fabricated using three different approaches exhibit D2D variability. It is hard to notice any trend between the I-V curves and the fabrication approaches. Whereas, the parameters extracted from the I-V curves such as reset

voltage and the maximum currents in the first reset operation show a trend concerning the fabrication approaches. As shown in Figure 3.13 (d), the memristive devices fabricated using the type 3 method demonstrate slightly reduced variability compared to type 1 and 2 approach devices. The maximum currents of the first reset operation are quite often studied in literature due to the role played by the metal-oxide interface at the BE in strengthening the CF of the memristive device [112], [113]. These values are extracted from the I-V curves shown in Figure 3.13. The distribution of the maximum reset current values is represented in the form of box plots shown in Figure 3.14. Type 1 approach devices demonstrate large variability, while type 3 demonstrates the lowest variability.

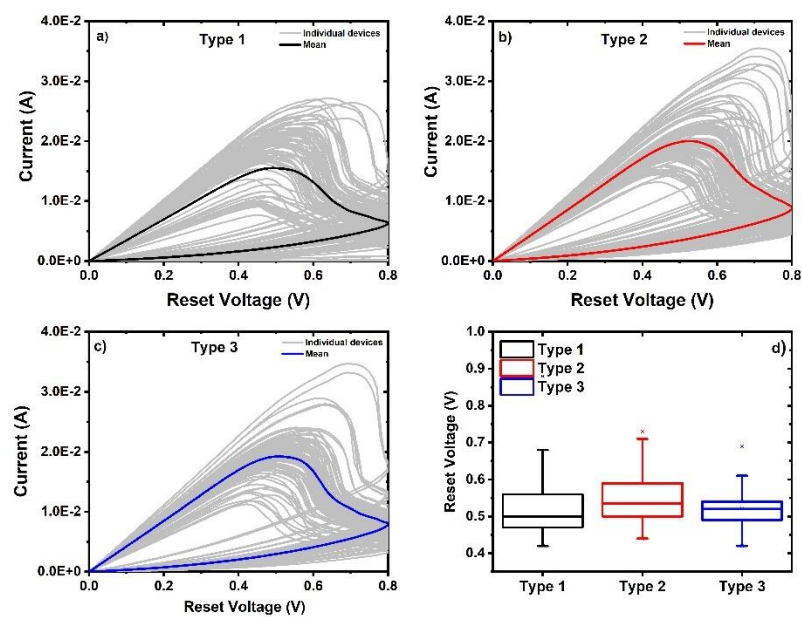


Figure 3.13: Wafer-scale first-reset I-V curves of class A devices fabricated using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of first-reset voltages of class A devices [98].

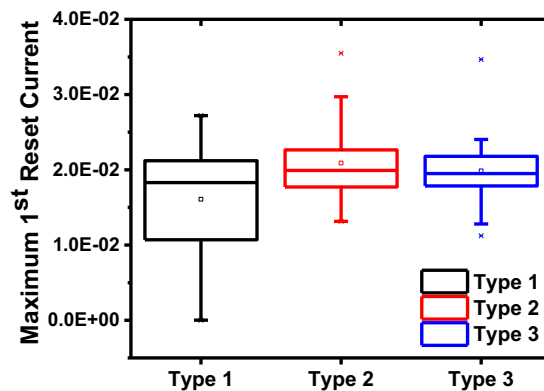
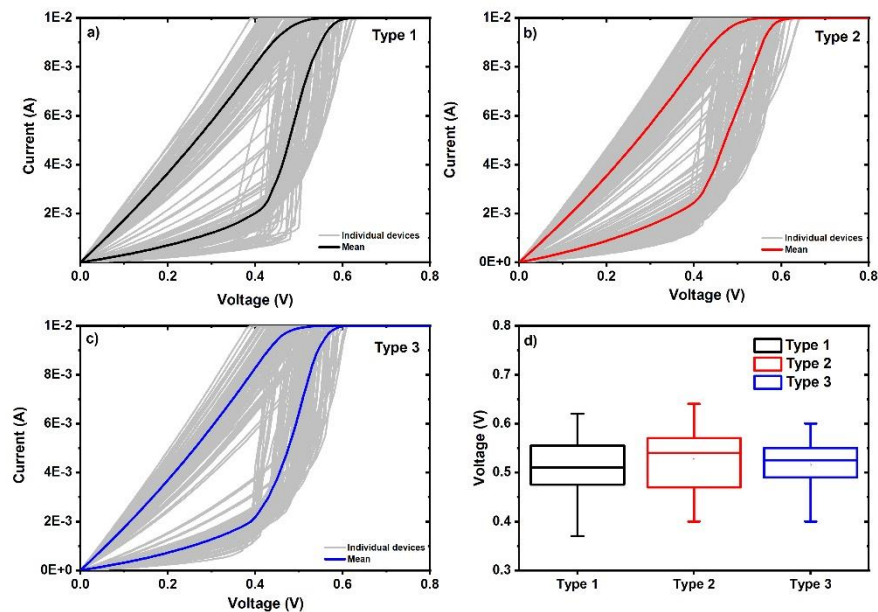


Figure 3.14: The distribution of maximum currents in the first reset operation of the memristive devices fabricated using the three different approaches [98].

The set operation of a memristive device recreates the partially broken CF. Figure 3.15 shows the I-V curves of the first set operation of the individual memristive devices fabricated using the three different approaches. Irrespective of the fabrication approach used all the memristive devices exhibited D2D variability as shown in Figure 3.15. Further, the set voltage values extracted from the I-V curves shown in Figure 3.15 (a), (b) and (c) are represented in the form of box plots in Figure 3.15 (d). Type 3 fabricated devices exhibited less variability in set voltage values compared to type 1 and 2 approach devices. The D2D variability observed in the I-V curves of the individual memristive devices in forming, reset and set operations is the consequence of the stochastic nature of the CF [114]. Additionally, the variations involved in the parameters of the memristive device fabrication processes such as memristive layer thickness, Ti layer thickness and the BEOL thermal budget which activates the oxygen scavenging properties of the Ti layer etc. also contribute to the D2D variability [10]. One of the ways to reduce this variability is to implement tighter process control in the fabrication [115] and use pulsed-based algorithms with a verify scheme to characterize the memristive devices [48].



**Figure 3.15: Wafer-scale set I-V curves of class A devices fabricated using type (a) 1, (b) 2 and (c) 3 approaches. (d) Distribution of set voltages of class A devices [98].**

Finally, the switching performance of the memristive devices is evaluated through the MW extracted from the set cycles. The distribution of the MW extracted after the 1<sup>st</sup> and 50<sup>th</sup> set cycles is shown in Figure 3.16. The MW decreases with cycling for type 1 approach devices, whereas it increases for the type 2 and type 3 approach devices. A relatively higher MW with slightly reduced variability is observed in type 3 approach devices [98].

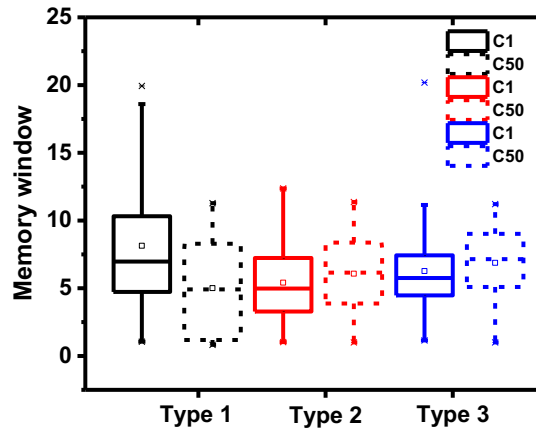


Figure 3.16: The memory window after 1<sup>st</sup> and 50<sup>th</sup> cycle, of the class A devices fabricated using three different approaches [98].

### 3.6 Conclusion

The memristive devices are fabricated in the BEOL of the 250 nm BiCMOS technology of IHP. The MEMRES module is patterned using three different fabrication approaches namely type1, type 2, and type 3. The quality of the layers in the MEMRES module and the process variations in the fabrication are evaluated through two different PrCM test structures, namely class A and class B. Both the test structures are characterized for their leakage currents, breakdown voltages, and capacitance values, where the type 3 approach devices demonstrated improved performance. Further, class A PrCM test structures are characterized to investigate the resistive switching performance. The memristive devices fabricated using the three different approaches still exhibited D2D variabilities in their switching operations. However, the reduced variability in switching parameters and the increase in the MW of type 3 approach devices after 50 subsequent cycles demonstrate the advantage of using this approach for the MEMRES module integration. This makes the type 3 approach devices with sidewall protection using spacer and encapsulation techniques suitable candidates for the emerging non-volatile memory applications.



## Chapter IV

---

# 4 Process Integration of Memristive Devices

---

Integrating the memristive devices into a CMOS baseline technology is essential to improve their performance [116]. Additionally, the CMOS compatibility of the memristive devices makes it feasible to fabricate them in the BEOL interconnects [117]. Integrating the memristive devices with the CMOS platform not only reduces the parasitic RC but also has the potential to realize hardware architectures beyond Von Neumann [118]. In-memory computing is one such architecture that can be realized by monolithic integration of FEOL CMOS logic and BEOL memristive devices [66]. As the name specifies, the computation takes place within the memory array which reduces the on-chip data transfer [119]. Such computing architectures demonstrate enhanced throughput and energy efficiency.

One of the main goals of this work is to integrate the MEMRES module developed in chapter 4 into a 130 nm BiCMOS baseline technology. This is achieved by adapting the MEMRES module fabrication process steps of a 250 nm baseline technology to 130 nm. In this regard, it is essential to understand the basic features of both baseline technologies. This chapter starts with discussing the main features and the basic differences between the 250 nm and 130 nm baseline technologies. The standard layout design rules for the MEMRES module integrated into the 250 nm technology are discussed and taken into consideration for the module integration in the 130 nm technology.

Selecting the metal level in the BEOL interconnects on which the MEMRES module needs to be fabricated is one of the important decisions. The metal level selection is addressed in this chapter with respect to the layout design rules and technology constraints.

Choosing the dimension of a select transistor in 1T-1R integrated devices is limited by the forming voltage requirement of memristive devices. This factor is considered and addressed in the corresponding layout design in this chapter.

Stable yield and uniform process parameters are the basic requirements of any standard CMOS baseline technology [120]. In this regard, it is essential to monitor and control the process parameters related to the MEMRES module. This factor is addressed in this chapter by designing the dedicated process control monitor (PrCM) test structures. Finally, the chapter discusses the impact of the MEMRES module integration on the 130 nm baseline technology.

*“The integration of memristive devices into the BEOL of the 130 nm BiCMOS technology of IHP was conceptualized by Professor Christian Wenger and Professor Andreas Mai. I implemented the conceptualized idea by considering the various factors necessary for integration and by designing the layouts for 1T-1R test structures and process control monitor (PrCM) test structures. Further, I developed the process fabrication flow for the MEMRES module integration under the supervision of Dr. Marco Lisker. Parts of this chapter have been published in Ref. [122].”*

## 4.1 IHP’s BiCMOS Baseline Technologies

SGB25V and SG13S are the two basic BiCMOS baseline technologies of IHP. SGB25V is the 250 nm technology node, while SG13S is the 130 nm technology node. The technology node here refers to the smallest feature size achievable using the lithography process which is also the smallest gate length of the transistor achievable in that process technology. Both SGB25V and SG13S technologies consist of industry standard front-end-of-line (FEOL) and back-end-of-line (BEOL) silicon processes. Figure 4.1 and Figure 4.2 show the cross-sectional schematic view of various device components in SGB25V and SG13S baseline technologies, respectively. The FEOL mainly consists of active devices such as low and high voltage CMOS transistors, and high-performance heterojunction bipolar transistors (HBTs) which are capable of operating at high frequencies. The BEOL part consists of the metal interconnects, and the passive electrical components like resistors, capacitors and inductors.

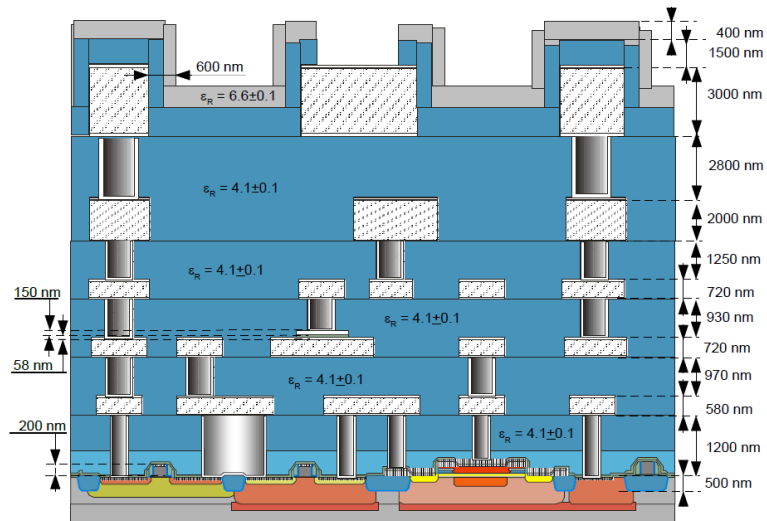


Figure 4.1: Schematic cross-section of the components in the SGB25V BiCMOS process technology of IHP.

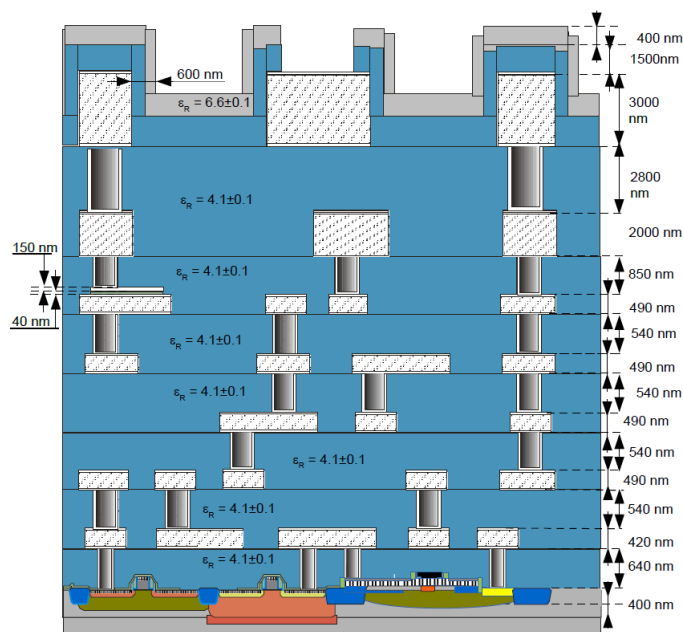
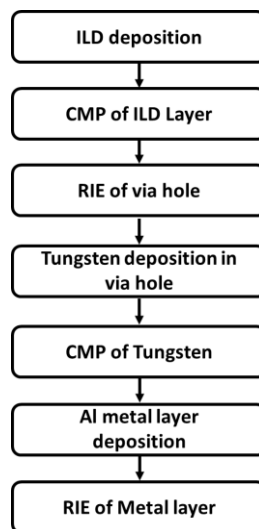


Figure 4.2: Schematic cross-section of the components in the SG13S BiCMOS process technology of IHP.

The BEOL interconnects are mainly made up of tungsten-based vias and aluminium-based metal lines. The process flow of the BEOL interconnects is illustrated in Figure 4.3. The fabrication of interconnects starts with the deposition of interlayer dielectric (ILD) silicon oxide ( $\text{SiO}_2$ ) using chemical vapor deposition (CVD). The deposited ILD layers are planarized using chemical mechanical planarization (CMP). Subsequently, the via holes are dry etched into the ILD layers with reactive ion etching (RIE) using photoresist as a mask. After RIE and wet clean process steps, Ti

and TiN liner layers are deposited in the via holes, which serve as the adhesion and the barrier layer, respectively, for the tungsten plug [100]. The vias are then filled with tungsten using CVD. The deposited tungsten layer is planarized using CMP, until the tungsten on top of the ILD is removed. Next, a metal layer stack consisting of Ti/TiN/Al:Cu/Ti/TiN is deposited by physical vapor deposition (PVD). The metal layers are structured by RIE with photoresist as a mask. The process flow illustrated in Figure 4.3 is repeated sequentially in order to obtain multiple levels of BEOL interconnects. Finally, a passivation layer stack consisting of SiO<sub>2</sub> and silicon nitride (SiN) is deposited on the topmost metal layer. The passivation layers on top of the contact pads are removed by RIE.

The complete end-of-line (EOL) fabrication process for SGB25V and SG13S requires 21 and 34 lithographic masks, respectively. Further, SGB25V consists of three thin and two thick metal layers, while SG13S consists of five thin and two thick metal layers. The two thick top metallization levels (TM1 and TM2) in both the technologies are mainly used for the fabrication of passive electrical components. The lower thin metallization levels are mainly used for interconnecting the device components within the circuits.



**Figure 4.3:** Process flow chart of the fabrication of BEOL interconnects in the SGB25V and SG13S technologies.

The metal-insulator-metal (MIM) capacitors are one of the main passive components embedded in the BEOL interconnects. They are mainly used for energy storage, limiting the voltages, filtering the signals and many other high frequency applications. Both baseline technologies use SiN as the dielectric layer for the MIM capacitors. The measured capacitance densities are in the range of fF/ $\mu\text{m}^2$ . In SGB25V, the MIM capacitor is placed between metal 2 (M2) and metal 3 (M3), whereas it is

placed between metal 5 (M5) and top metal 1 (TM1) in SG13S. The fabrication of the MIM capacitor requires a single mask in both technologies. The MIM stack is patterned by aligning the MIM layer mask with the already patterned via layer beneath. The MIM mask is aligned with via 1 (V1) layer in SGB25V, while it is aligned with via 4 (V4) in SG13S.

Sintering is one of the important EOL fabrication processes. It is an annealing process step carried out in a diffusion furnace in the presence of H<sub>2</sub> and N<sub>2</sub> gas mixtures which is popularly known as forming gas. Sinter anneal takes place at 400 °C which is a temperature well suited for BEOL processes. It neutralizes the interface charges created due to the fabrication process steps, thereby reducing the contact resistance between the interconnects [121]. The first sinter annealing in both technologies is carried out for 30 minutes, after the patterning of the M1 layer. SGB25V consists of three additional sinter annealing steps one each after the patterning of TM1, top metal 2 (TM2) and the passivation layers. SG13S consists of two additional sinter annealing steps, one after the patterning of TM2 and the other after the patterning of the passivation layers. The total sinter annealing times of SGB25V and SG13S are divided into 4 and 3 steps of 30 minutes each, respectively.

## 4.2 Integrated MEMRES module in 250 nm BiCMOS technology

The MEMRES module in SGB25V is fabricated using the type 1 approach described in chapter 4. The module is integrated between M2 and M3 in the BEOL interconnects by using the standard MIM mask of the baseline technology. The optimal memristive device dimensions determined from previous studies is 0.6 μm x 0.6 μm [95]. The layout design rules related to the MEMRES module integration are illustrated in Figure 4.4. In spite of the tolerances in fabrication processes, the via has to be covered with the enclosure margin in order to guarantee the reliability of the contact [122]. The size of via 2 (V2) along with its enclosure margin which is fabricated on top of the memristive device has to be less than the aforementioned device size. The size of V2 fabricated on top of the memristive device is 0.42 μm with an enclosure margin of 0.18 μm. The memristive devices are integrated with the NMOS transistors with a gate length of 240 nm and a gate width 1.14 μm from the SGB25V technology. Figure 4.5 depicts a TEM cross section of an integrated 1T-1R device in the SGB25V technology. The bottom electrode of a memristive device is connected to the drain terminal of a transistor from M2, while its top electrode terminal is connected from M3 through V2. In principle, the NMOS transistor forms a series connection with a memristive device in an integrated 1T-1R structure.

Since the standard MIMs in SGB25V are based on a SiN dielectric and the memristive devices are based on a HfO<sub>2</sub> dielectric, the MEMRES module is processed using dedicated wafers. The integration approach of the MEMRES module in SGB25V saves the expense on an additional mask, but at the cost of sacrificing the usage of standard MIM capacitors in the circuit and fabricating dedicated wafers for the module. Further, the standard MIM capacitors are one of the essential passive components which are required for designing the peripheral circuitry for memristive arrays. They are used in the peripheral circuit of the memristive arrays as by-pass capacitors at supply voltage terminals, and as an integral component of analog to digital converters (ADC) and operational amplifiers (op-amp). The non-usability of standard MIM capacitors of the baseline technology hinders the usage of memristive devices for neuromorphic and embedded applications.

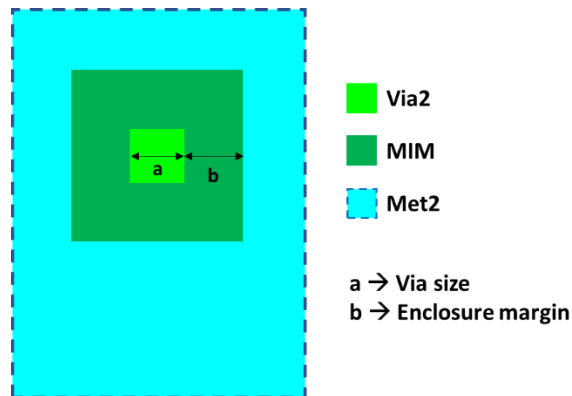


Figure 4.4: Layout design rules for the integration of the MEMRES module into the 250 nm BiCMOS baseline technology.

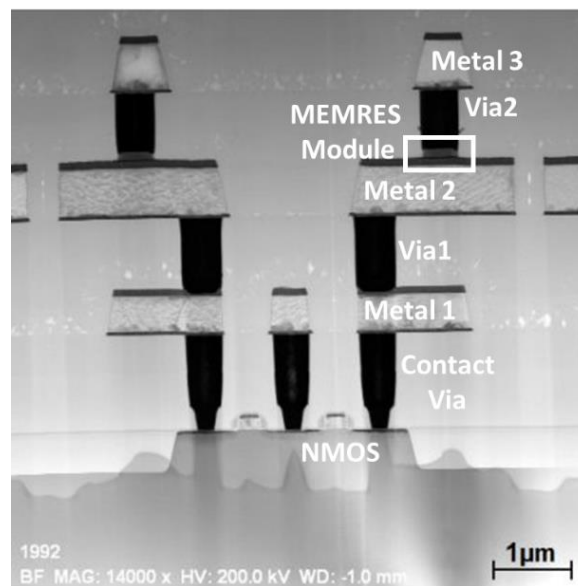


Figure 4.5: TEM cross section of a MEMRES module integrated between M2 and M3 in the BEOL interconnects of the 250 nm BiCMOS technology.

### 4.3 Integration of the MEMRES module into a 130 nm BiCMOS technology

In the system-level design based on memristive devices, the peripheral circuit is one of the most power consuming parts [123]. Hence, in order to obtain high performance with low power consumption, switching to smaller technology nodes is inevitable. Further, due to the challenges mentioned in the previous section with respect to the MEMRES module in SGB25V, it becomes essential to integrate the MEMRES module into SG13S and it is also one of the main goals of this work. The first step towards the integration process was to choose the BEOL metal levels between which the MEMRES module can be placed. A memristive device consists of two contact terminals, hence accordingly it can be integrated between any two metallization levels in the BEOL process [122].

The optimal device dimensions and the layout design rules of the MEMRES module in SGB25V were maintained similarly in SG13S. Instead of a standard MIM mask, an additional HFO mask is used for the fabrication of the module as shown in Figure 4.6. Table 4 illustrates the size of the vias in the BEOL interconnects of SG13S technology. The MEMRES module cannot be placed after M5 due to the violation of the previously mentioned layout design rule with respect to the via size and its enclosure margin. Further, placing the MEMRES module close to the CMOS transistor gives better control on the memristive device operations. However, placing it too close could affect the working of the CMOS transistor [122]. Hence, placing the MEMRES module on metal 1 (M1) is avoided. Placing the module on M2 was a good tradeoff with respect to the transistor and the memristive device operations. Additionally, the size of V2 with its enclosure margin of  $0.2\ \mu\text{m}$  comply with the aforementioned layout design rules of the MEMRES module.

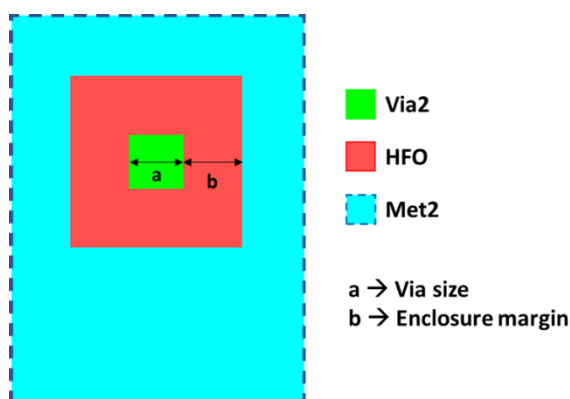
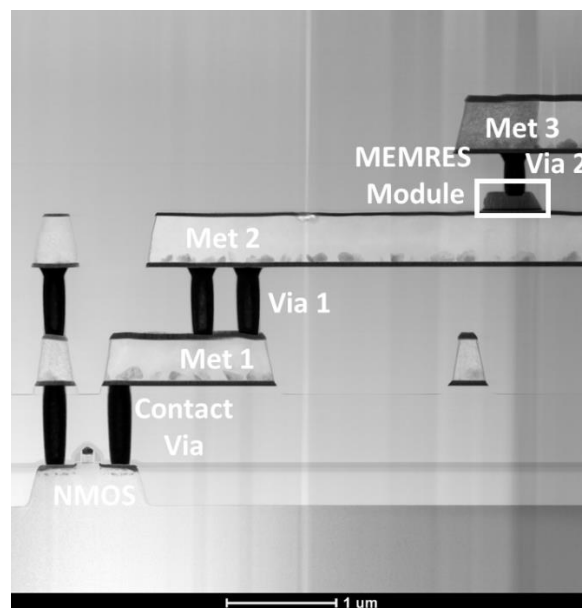


Figure 4.6: Layout design rules for the integration of the MEMRES module into the 130 nm BiCMOS baseline technology.

The BEOL thermal budget is another parameter on which the selection of metal levels depends upon. The BEOL fabrication processes such as ILD depositions, tungsten via depositions, AlCu metal layer depositions, sinter annealing etc. provide the thermal energy required for the activation of the oxygen scavenging layer (OSL) in the MEMRES module [124]. In order to keep the thermal energy encountered by the MEMRES module in SG13S and SGB25V technologies comparable, the choice of integrating the module between M2 and M3 was made. Further, this paves the way for utilizing the passive components of the baseline process for the peripheral circuitry of memristive arrays. Figure 4.7 shows a TEM image of a MEMRES module integrated between M2 and M3 in the SG13S technology.

**Table 4: Size of vias in the BEOL of the 130 nm BiCMOS technology.**

Sl.No	Description	Size ( $\mu\text{m}$ )
1	Via 1	0.19
2	Via 2	0.19
3	Via 3	0.19
4	Via 4	0.19
5	Top Via 1	0.42
6	Top Via 2	0.9



**Figure 4.7: TEM cross section of a MEMRES module integrated between M2 and M3 in the BEOL interconnects of the 130 nm BiCMOS technology.**

The BEOL thermal budget has a significant influence on the phase of the memristive switching layers. Hence, the phase of these layers is investigated by using X-ray diffraction (XRD) measurements. The XRD patterns of 8 nm of  $\text{HfO}_2$  layers and 6 nm  $\text{Hf}_{1-x}\text{Al}_x\text{O}_y$  layers which are deposited using TALD were analyzed. It can be



clearly seen from Figure 4.8 and Figure 4.9 that the as-deposited layers of both the oxides do not display any noticeable diffraction peaks. This clearly indicates that the layers are amorphous in nature. As illustrated in Figure 4.8 the  $\text{HfO}_2$  layers exhibit monoclinic diffractions after 30 minutes of sinter annealing. Whereas, the  $\text{Hf}_{1-x}\text{Al}_x\text{O}_y$  layers do not display any kind of diffractions even after 90 minutes of sinter annealing as shown in Figure 4.9 [122].

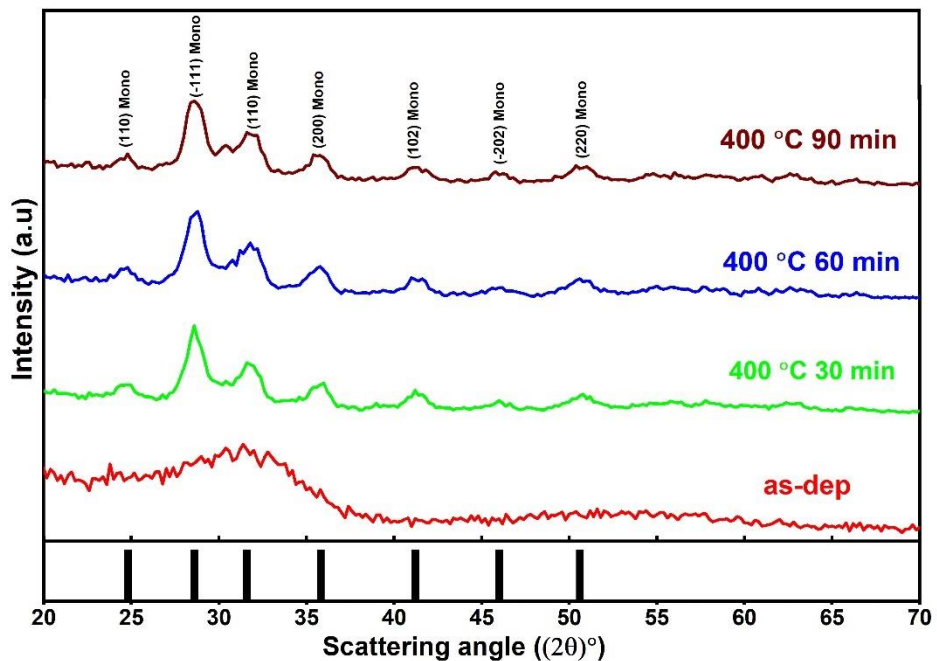


Figure 4.8: XRD patterns of as-deposited and sinter annealed 8 nm  $\text{HfO}_2$  layers deposited using TALD. Adapted from [122].

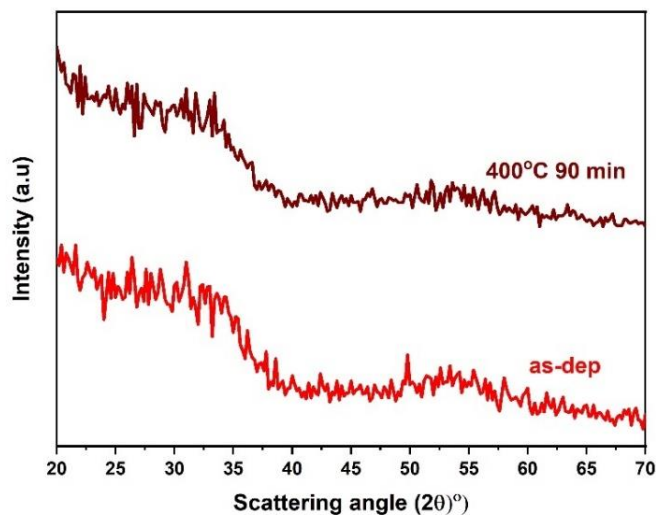


Figure 4.9: XRD patterns of as-deposited and sinter annealed 6 nm  $\text{Hf}_{1-x}\text{Al}_x\text{O}_y$  layers deposited using TALD. Adapted from [122].

One of the parameters which affect the performance of the memristive devices is the phase of their switching layers [47]. BEOL thermal budget including the sintering processes is inevitable due to the aforementioned reasons. The sintering process results in the formation of polycrystalline  $\text{HfO}_2$  layers which gives rise to high device-to-device (D2D) variability [96]. Hence,  $\text{HfO}_2$  is doped with  $\text{Al}_2\text{O}_3$  layers, which have higher crystallization temperatures, and thereby keep the memristive switching layers amorphous. Additionally, the  $\text{Hf}_{1-x}\text{Al}_x\text{O}_y$  layers result in uniform resistive switching of the devices with reduced D2D variability [48]. The memristive switching layers are discussed briefly in Chapter 1.

#### **4.4 Layout design of 1T-1R test structures in 130 nm technology**

The SG13S BiCMOS technology consists of two different types of transistors, namely low voltage and high voltage transistors. The low voltage transistors are fabricated with a gate oxide thickness of  $\sim 2.6$  nm, while the high voltage transistors are fabricated with  $\sim 7.5$  nm. Accordingly, the breakdown voltage of the low voltage transistors is  $\sim 2.5$  V, whereas it is  $\sim 6$  V for the high voltage transistors. The memristive device itself has various advantages as mentioned in Chapter 1. However, it is essential to integrate them with a CMOS transistor to avoid the sneak path currents in the arrays and to limit the current through the individual devices. Accordingly, the transistor avoids inaccurate measurements in the arrays and prevents the hard breakdown of the memristive devices, respectively [116].

Forming is a crucial and energy-consuming operation performed initially to activate the switching mechanism in the memristive devices [125]. The voltage levels required for this one-time operation are comparatively higher, making the usage of smaller select transistors challenging. However, smaller transistors consume less power and switch faster. Hence, the breakdown voltage of the transistor and the voltage required to perform the forming operation on a memristive device has to be carefully analyzed. The transistor suitable for the resistive switching operation in 1T-1R devices is investigated by integrating the memristive devices of aforementioned dimensions with twelve different transistor sizes in SG13S. The layout design of the 1T-1R test structures is shown in Figure 4.10. It consists of six different sized low and high voltage transistors, respectively. The low voltage transistors are named as DIGI01 and their dimensions are illustrated in Table 5, while the high voltage transistors are named as HIVT01 and their dimensions are illustrated in Table 6.

**Table 5: Dimensions of the six different low voltage transistors of layout section DIGI01.**

DIGI01					
Description	Gate Length (nm)	Gate Width (nm)	Description	Gate Length (nm)	Gate Width (nm)
T3	130	1000	T6	200	220
T2	130	500	T5	180	200
T1	130	150	T4	150	170

**Table 6: Dimensions of the six different high voltage transistors of layout section HIVT01.**

HIVT01					
Description	Gate Length (nm)	Gate Width (nm)	Description	Gate Length (nm)	Gate Width (nm)
T3	600	300	T6	600	600
T2	470	300	T5	470	600
T1	330	300	T4	330	600

The layout shown in Figure 4.10 consists of four columns of bond pads used for connecting the terminals of the 1T-1R test structures. The first and the second columns are dedicated to the low voltage transistors, and the third and fourth columns are dedicated to high-voltage transistors. The smallest transistor gate length in the SG13S technology is 130 nm. The transistors in the first column are designed with the 130 nm gate length but with gate widths of 150, 500, and 1000 nm. The drain current increases with the width of the transistor, which is essential for all device operations, especially the current read-out operation in memristive arrays. The transistors in the second column are designed to keep their W/L ratio constant. The high voltage transistors in the third and fourth columns are designed with increasing W/L ratio.

A layout design of a typical 1T-1R test structure is shown in Figure 4.11. The bond pads used in the layout are of dimensions  $80\ \mu\text{m} \times 80\ \mu\text{m}$  and are placed with a pitch of  $45\ \mu\text{m}$  in the Y direction. The contact terminals of the 1T-1R test structure illustrated in Figure 4.11 are described for a transistor with gate length of 130 nm and gate width of 150 nm. The gate and source terminals of a transistor are connected to individual bond pads from M1. The TE and BE of a memristive device are connected similar to SGB25V MEMRES module as described in section 4.2. M3 is used for connecting the TE contact terminal of a memristive device to an individual bond pad. Every column is provided with one bulk contact terminal connected to the last bond pad of a particular column from metal 1. Every transistor has an individual bulk contact that is placed close to it, and they are shorted to provide one bulk terminal.

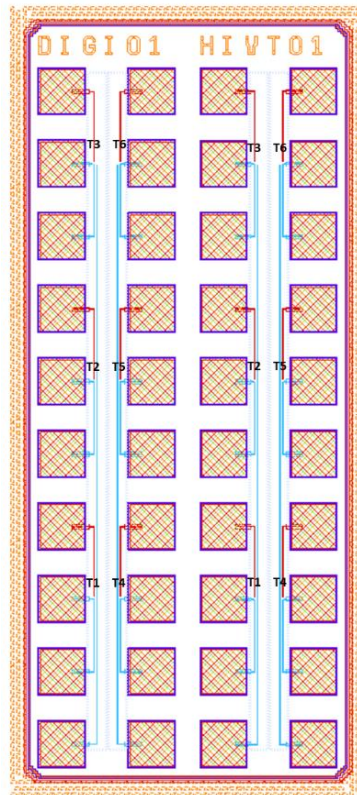


Figure 4.10: Schematic layout of CMOS integrated 1T-1R memristive test structures in the 130 nm BiCMOS technology.

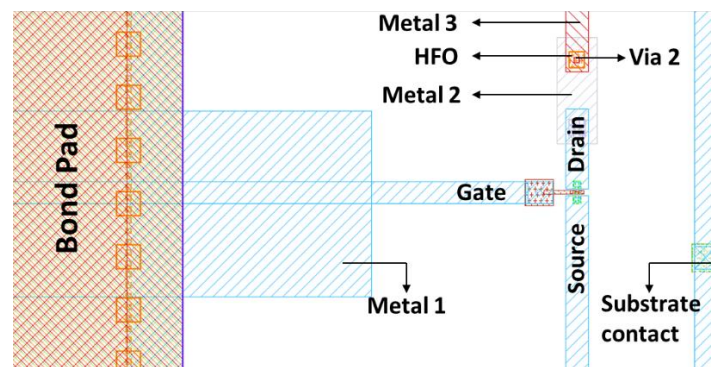


Figure 4.11: Schematic layout of a 1T-1R test structure with its contact terminals.

## 4.5 Layout design of PrCM test structures in 130 nm technology

After successfully integrating the MEMRES module into a 130 nm baseline technology, one of the important steps is to convert it into a stable technology. This is realized by designing monitor structures for the module. The process control monitor (PrCM) block with test structures allows monitoring and controlling the technology specifications of a module to obtain reliable and reproducible parameters of various

fabrication process steps. It is a part of the statistical process control (SPC) method to maintain the quality of the processes involved in the baseline technology [120]. Further, the electrical measurements can be performed on the monitor structures using inline testers inside the fab. If the obtained results are outside the defined specifications of the module, the process steps can be reworked.

Further, due to the specific bond-pad configurations, performing inline measurements on bigger layouts with memristive devices integrated on the circuit and system level is challenging. Hence, it is necessary to place a dedicated PrCM block for the MEMRES module close to the memristive device-based layouts. The layout design of the PrCM block dedicated for MEMRES module is shown in Figure 4.12. It allows to perform inline electrical measurements on the individual memristive devices and monitor structures. Six different kinds of monitor structures are designed for the MEMRES module, as shown in Figure 5.12. Soon after patterning the M3 layer, all PrCM structures are electrically characterized for their leakage current behavior. A leakage current value in the microampere range is set as the specification limit, and for values above this limit, the respective process steps have to be reworked.

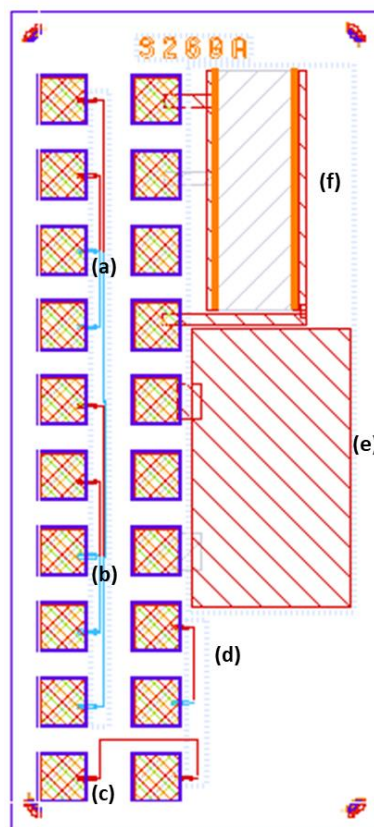


Figure 4.12: Schematic layout of PrCM test structures dedicated for the MEMRES module in the 130 nm BiCMOS technology.

As shown in Figure 4.13, test structures a and b of the PrCM block are 1T-1R devices with access to their drain terminals. These structures allow the electrical characterization of the select transistor and the memristive device separately. Consequently, the characteristics of the devices can be analyzed individually and together as in a typical 1T-1R structure. Further, in test structure b, the transistor is connected to the memristive device by a bridge structure as shown in Figure 4.13b. This structure connects the drain terminal of the transistor to the BE of the memristive device through V2 instead of V1, which is the case of the standard 1T-1R structure. This structure allows to investigate the effects of plasma-induced damage (PID) on memristive devices. As shown in Figure 4.13, test structures c and d of the PrCM block are individual memristive devices. In test structure c, the BE connection is made from the top of M2 using V2 interconnects. Whereas, the BE connection is made from the bottom of M2 through V1 interconnects for test structure d. These structures are also used for investigating the PID effects.

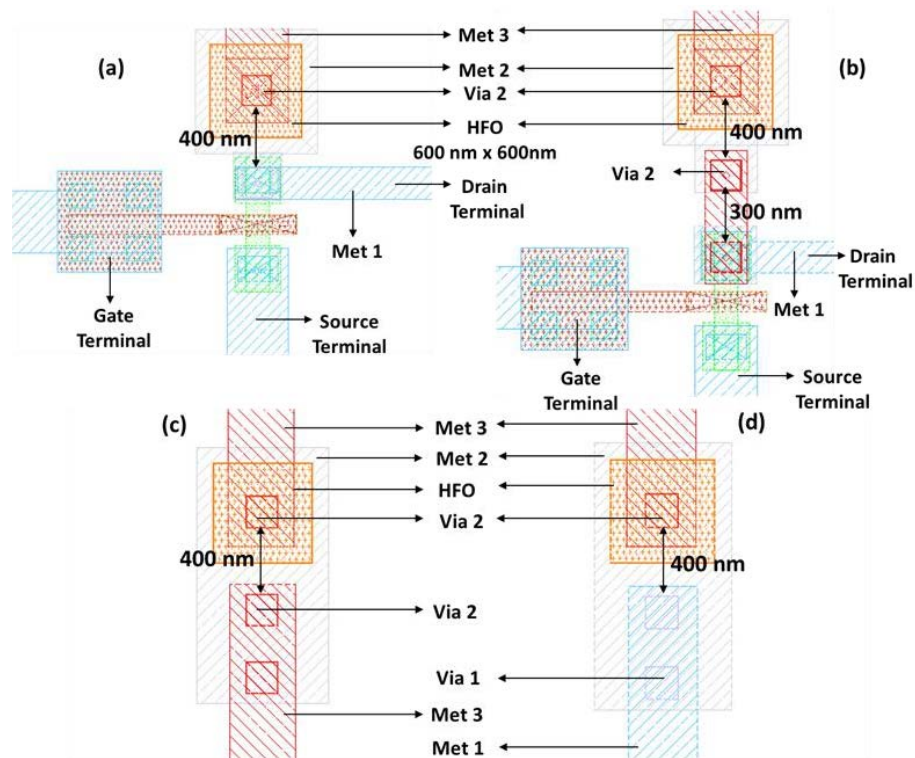


Figure 4.13: PrCM test structures (a) a, (b) b, (c) c, and (d) d of the MEMRES module.

Test structure e of the PrCM block consists of a corner MIM device, as shown in Figure 4.14. It is a matrix of memristive devices of dimensions  $0.6 \mu\text{m} \times 0.6 \mu\text{m}$  connected in parallel. The primary usage of these structures is explained in detail in chapter 3. Test structure f consists of two combs with 165 fingers each. Each comb structure behaves like an individual capacitor. These structures are mainly used to

check the top electrode etching process of MEMRES module. The two comb structures will be shorted if the top electrodes are not correctly etched. The electrical setup for measuring the comb PrCM structures is illustrated in Figure 4.15 b.

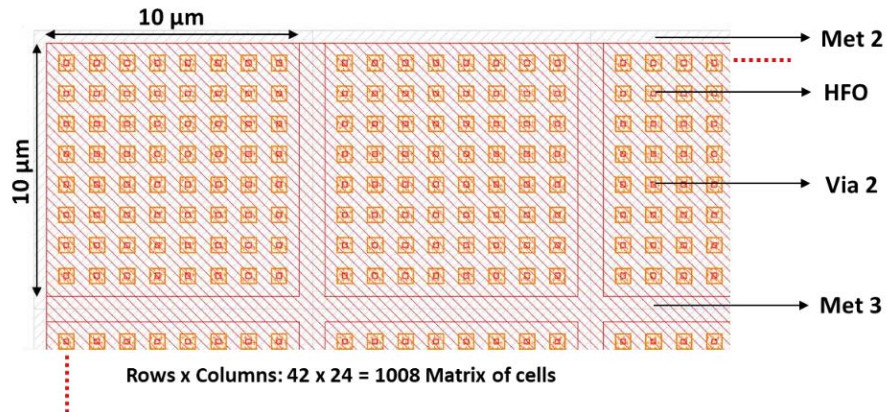


Figure 4.14: Matrix of MIM structures represented as structure e in the PrCM block dedicated for the MEMRES module.

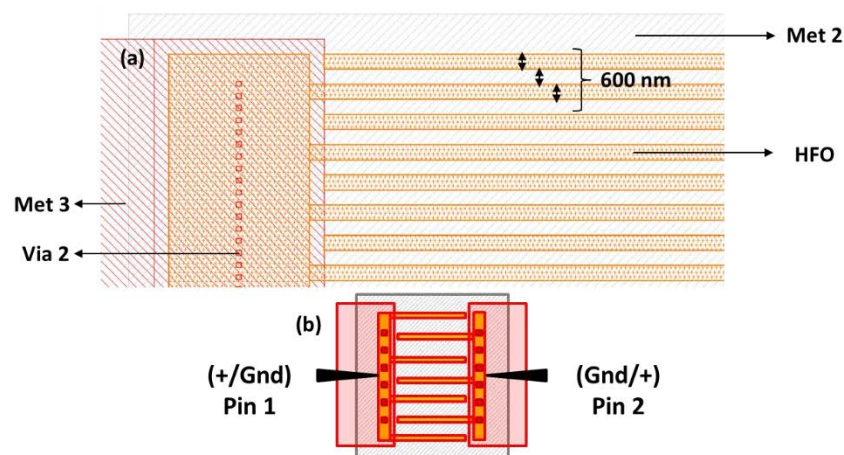


Figure 4.15: Structure f in the PrCM block dedicated for the MEMRES module (a) comprising of comb structures and (b) the schematic illustration of its electrical characterization.

## 4.6 Impact of MEMRES module integration on a 130 nm baseline technology

Integrating any new module into a baseline technology is a significant step. The new module has to be integrated in such a way that the active devices and the passive components of the standard baseline technology receive minimum to negligible impact. The impact of integrating the MEMRES module into the BEOL of a 130 nm baseline technology is investigated in this section. Since the module is integrated between M2 and M3, the fabrication process steps between them are analyzed for the

impact of integration. The process steps between M2 and M3 which could interfere with the baseline processes include M2 patterning, V2 fabrication and the thermal budget of the MEMRES module.

Initially, the leakage currents between the M2 lines are measured with the dedicated PrCM comb structures. The measured currents were in the range of picoampere signifying that the M2 lines were etched according to the technology specifications. The corresponding electrical results are shown in Appendix A. Hence, the impact of MEMRES module integration on M2 etching step was found to be negligible.

Via 2 in the MEMRES module integrated in 130 nm baseline technology has two different functionalities. Firstly, it is used for connecting the TE of the memristive device to M3 and secondly, as a standard interconnect that connects M2 and M3. The height of the standard V2 is  $\sim 0.6 \mu\text{m}$ , whereas it is reduced by  $\sim 165 \text{ nm}$  for the V2 fabricated on top of the memristive device. However, the standard baseline technology recipes are used for the fabrication of both the vias. The contact resistance of V2 is measured by four-point probe technique on via kelvin structures. The measured via resistance of  $\sim 12 \Omega$  is within the defined technological specifications indicating that the MEMRES module integration did not affect the contact resistance of standard V2 interconnects. The corresponding electrical results are shown in Appendix A.

The integrated MEMRES module also adds an extra thermal budget to the baseline process due to the deposition processes involved in the fabrication of the module. This includes the memristive switching layers deposited using TALD, the Ti/TiN TE layers deposited using PVD and the  $\text{Si}_x\text{N}_y\text{O}_z$  spacer and encapsulation layers deposited using CVD. However, all these layers are deposited at a BEOL compatible temperature. Further, the MEMRES module integration does not affect the passive devices in the BEOL as they are fabricated after M5. In order to investigate the impact of module's thermal budget on the FEOL devices, the important device characteristics of HBTs, nMOS and pMOS in the standard SG13S are compared with the MEMRES module integrated SG13S at the end-of-line measurement step. In case of HBTs, the device characteristics such as collector gain, collector currents and emitter resistances were compared. In case of nMOS and pMOS devices, drain leakage currents, drain saturation currents and threshold voltages were compared. The impact was found to be very minimum to negligible and the corresponding electrical results are shown in Appendix B.

The mechanical stress induced by the layers of the MEMRES module is considerably low. Firstly, the TALD layers are very thin (6 - 8 nm) and hence the stress



induced by them is negligible. The area coverage of Ti/TiN TE layers on the entire die is less than 1%, and their layer thickness is lower compared to the TE of the standard MIM. Therefore, the stress induced by them is also low. The  $\text{Si}_x\text{N}_y\text{O}_z$  layers are anyway used as an anti-reflection coating (ARC) material on the metal layers in the baseline technology, hence they do not impart any additional stress.

## 4.7 Conclusion

In this chapter, the necessity of integrating the MEMRES module into the 130 nm baseline technology is discussed. The module integration is explained with respect to the metal level selection in the BEOL interconnects, selection of FEOL transistor, thermal budget encountered by the module and the layout design rules. The module is successfully integrated into SG13S by using one additional HFO mask. The process parameters of the integrated module are monitored and controlled by using dedicated PrCM structures. Finally, the impact of integrating the MEMRES module into the 130 nm baseline technology is discussed with respect to the critical fabrication process steps.

## Chapter V

---

# 5 Variability in Memristive Devices

---

The variability in the memristive device operations is one of the major challenges which needs to be overcome to use them for commercial applications. The memristive devices exhibit mainly two different types of variabilities, namely device-to-device (D2D) and cycle-to-cycle (C2C). The main cause of D2D variability is the fabrication process, whereas the C2C variability occurs due to the stochasticity involved in the generation and recombination of oxygen vacancies within the memristive switching layer and at its interfaces [115]. Due to this, the memristive devices can be programmed with certain distribution and not to a nominal value [97]. This complicates the design of peripheral circuits for the memristive arrays. Additionally, it increases the overhead on the circuit and system level to design resilient and variation-tolerant circuits [10].

To reduce the variability in the memristive device operations, various methods are proposed in the literature regarding device engineering, electrical characterization techniques, and circuit design optimizations [114]. Under the device engineering domain, techniques such as doping the memristive switching layer with an alloy [126], bi-layer memristive stacks [49], nano-laminate memristive layer stacks [127], etc. have been reported. In electrical characterization, hot forming [128], pulse-based programming algorithms with a verify scheme [95], constant voltage programming [129], etc. among many other techniques have been reported, which exhibited reduced

variabilities. Under the optimized circuit design domain, techniques such as read-out from two parallel memristive devices, sensing circuits to tolerate the fluctuations in current read-out, and many other assist circuits at the periphery of the memristive arrays have been reported [114].

The mitigation of variability in memristive devices is addressed in this chapter under two different domains. Under the first domain, electrical characterization techniques are used. The variability in the forming operation of the memristive devices in a 4kbit array is investigated using two different pulse-based programming algorithms. The impact of the incremental step pulse and verify algorithm (ISPVA) is investigated at four different temperatures. Whereas, the impact of constant amplitude pulse and verify algorithm (CAPVA) is investigated at three different pulse amplitudes. Under the second domain, a device engineering technique was used. A thin Al<sub>2</sub>O<sub>3</sub> layer of two different thicknesses was added to single layer of HfO<sub>2</sub>-based memristive devices. The switching functionality of the single and bi-layer memristive devices is investigated along with the properties of the conduction filament (CF) in the on and off states of the device using the quantum point contact (QPC) model.

*“The investigated samples in this work were fabricated by the cleanroom staff at the IHP pilot line. The experimental work on the memristive arrays was conceptualized by Dr. Eduardo Perez and Professor Christian Wenger. I performed the automated electrical measurements and analyzed the data. The experimental work on individual memristive devices was conceptualized by Professor Christian Wenger and myself. The assessment of the QPC model was supported by Professor Christian Wenger, Professor Enrique Miranda and Dr. Eduardo Perez. The devices were fabricated under the supervision of Dr. Marco Lisker. I performed the electrical measurements on a manual probe station, analyzed the data, and performed the device modeling using QPC. Parts of this chapter have been published in Ref. [130], [131].”*

## 5.1 Variability in integrated 1T-1R memristive arrays

The variability of the integrated memristive devices in a memristive array is investigated concerning their forming operation. Forming is a one-time operation and a preparatory step required for the normal functioning of the memristive devices. It requires higher voltages as it has to create an initial CF by soft-breakdown of the memristive switching layer [114]. It is a crucial step which regulates the number of oxygen vacancies available for the subsequent memristive switching operations and thereby defines the device characteristics [51]. Forming operation on 1T-1R integrated memristive arrays is performed through a pulsed-based measurement technique. The two main parameters extracted from this one-time operation are the forming current

and the forming voltage. Unfortunately, these two parameters also exhibit D2D variability. Higher voltage requirement combined with the D2D variability makes this operation undesired for memory applications [10]. However, it is an essential process step, and it is vital to reduce the variability of the parameters involved in the forming operation. This work demonstrates the electrical characterization approaches used to reduce the variability in the forming operation of the 1T-1R memristive arrays.

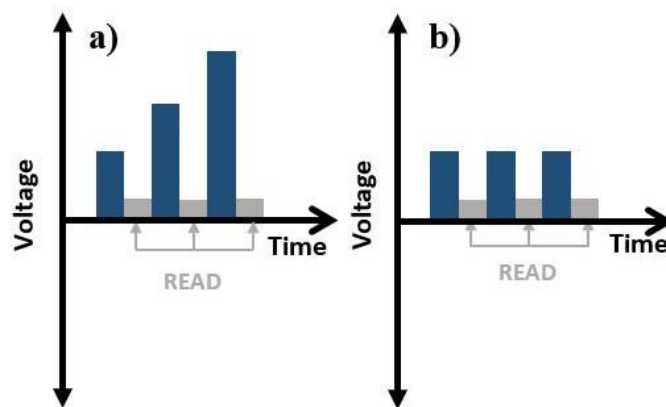
### 5.1.1 Programming algorithms and the memristive array operation

As described in Chapter 2, the memristive devices can be electrically characterized by using quasi-static DC or pulsed-based methods. The former method is time-consuming, and it is not the common way to work with memristive arrays [132]. Additionally, the compliance current continues to flow through the device even after the CF is formed until the voltage limit value. This causes undesired current stress on the memristive devices [133].

The pulsed-based technique causes less stress on the memristive devices due to the application of pulses with specific voltage amplitudes consisting of short pulse widths. However, applying pulses with higher voltage amplitudes and longer pulse widths results in larger forming current distributions [134]. This work utilizes two different sequences of the programming pulses along with the verify schemes, for forming the 1T-1R memristive devices in 4kbit arrays. The schematic illustration of their pulsing schemes is shown in Figure 5.1. The ISPVA consists of a sequence of incremental programming pulses in small voltage amplitude steps with a current read-out operation carried out after every pulse. The CAPVA consists of a sequence of programming pulses with constant voltage amplitude. Similar to ISPVA, a current read-out operation is performed after every pulse. The influence of the ISPVA is investigated at four different temperatures, while the CAPVA is investigated at three different voltage amplitudes.

The integrated 1T-1R memristive devices are fabricated in the SGB25V which is a 250 nm BiCMOS baseline technology at IHP. As shown in Figure 5.2, the architecture of 4 kbit memristive arrays in 250 nm technology consist of four main blocks: (I) an array of 4096 1T-1R memristive devices organized in  $64 \times 64$  rows and columns, respectively; (II) address decoder for the word line (WL); (III) address decoder for the bit line (BL); and (IV) the control circuitry for the operation of the memristive arrays (Mode) [135]. Each row of the memristive array consists of a dedicated WL which is connected to the gate terminals (G) of the transistors of that particular row. Similarly, each column of the memristive array consists of a dedicated BL which is connected to the drain terminals (D) of the transistors of that particular column. The source

terminals (S) of the transistor in every column are connected to a dedicated source line (SL). In principle, the architecture consists of 64 numbers of WL, BL, and SLs each, respectively. It allows the programming of an individual memristive device by selecting specific WL, BL, and SLs. However, the devices are programmed in this work row-wise. Each row is considered one page. The 4 kbit arrays consist of 64 pages in total. The 1T-1R memristive devices are fabricated as defined in section 5.2. The MEMRES module consists of a TiN/Al: HfO<sub>2</sub>/Ti/TiN stack. The thickness of the Ti layer and TiN top and bottom electrode layers are kept similar to the MEMRES module described in chapter 4. An Al-doped (~10 %) HfO<sub>2</sub> layer of 6 nm thickness is used as a memristive switching layer.



**Figure 5.1: Schematic illustration of (a) ISPVA and (b) CAPVA, pulsed based algorithms [130].**

The forming operation in 4kbit arrays is performed by applying 1.5 V gate bias on the WL. A sequence of increasing voltage pulse amplitudes from 2 V to 5 V in steps of 10 mV in case of ISPVA and constant voltage pulse amplitudes of 3, 4, and 5 V in case of CAPVA, are applied on the BL. The SL is grounded. Both programming algorithms use a pulse width of 10  $\mu$ s. A current read-out operation after every pulse is carried out at a BL voltage of 0.2 V while grounding the SL. Irrespective of the programming algorithm used, the write pulses are applied to the devices until they reach the defined target current value. Once a device crosses this threshold value, the application of the write programming pulse is stopped on it and only the current read-out operation takes place. Further, the write operation continues for the other devices until the defined value of the maximum voltage in ISPVA, and the defined value of the maximum number of pulses in CAPVA is achieved.

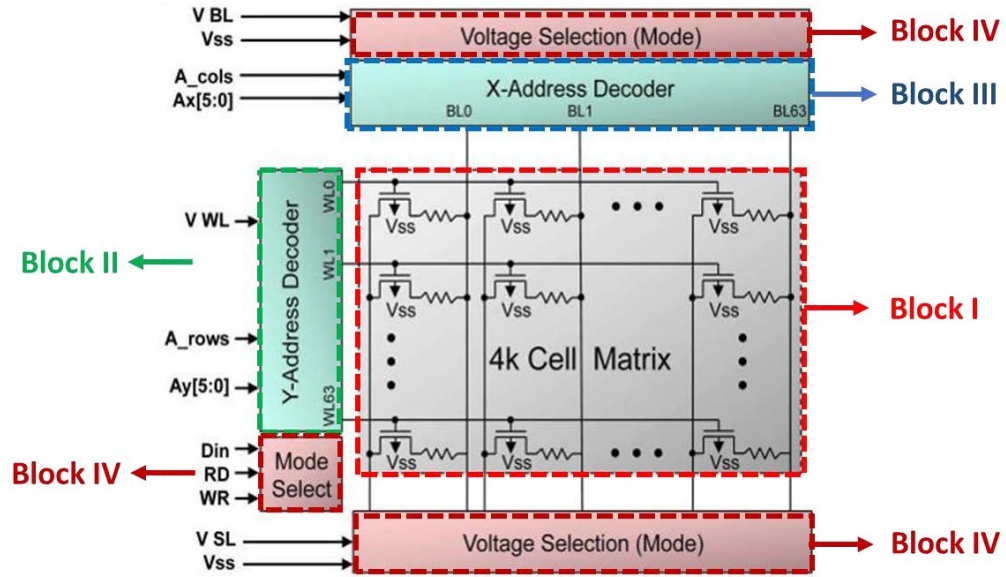


Figure 5.2: Block diagram of integrated 4kbit memristive arrays in 250 nm BiCMOS technology.

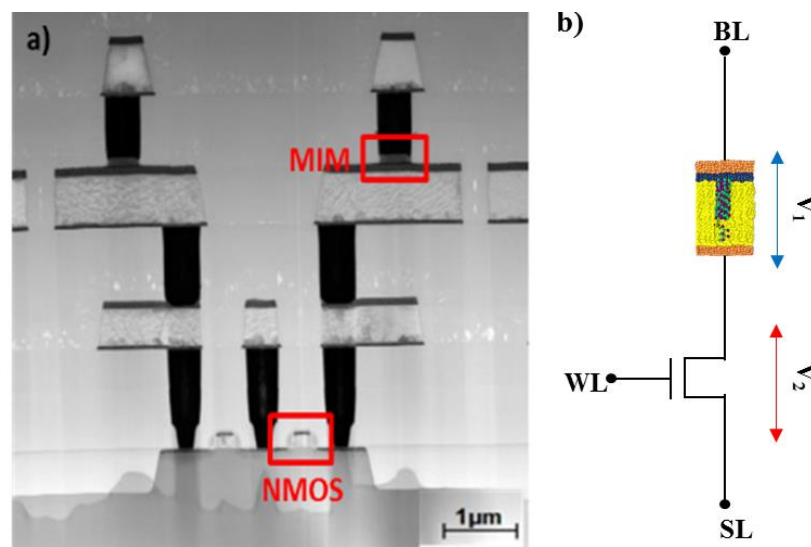
### 5.1.2 Variability in forming currents

The variability in the forming operation reduces the forming yield which in turn has a substantial effect on the memristive array and system performance [10]. For a given memristive device of specific dielectric thickness, the time required to form an initial CF depends on the voltage amplitude of the pulse and the operating temperature [134]. The forming operation on the memristive devices using ISPVA is performed at four different temperatures of  $-40^{\circ}\text{C}$ ,  $25^{\circ}\text{C}$ ,  $80^{\circ}\text{C}$ , and  $150^{\circ}\text{C}$ , while CAPVA is performed at three different constant voltage amplitudes of 3 V, 4 V, and 5 V. A target read-out current of  $30\ \mu\text{A}$  is defined for the forming operation. The required current compliance is provided through a 1.5 V gate bias of the nMOS transistor which is connected in series with the memristive device as shown in the cross section and the schematic of the 1T-1R cell in Figure 5.3.

Figure 5.4 shows the D2D cumulative distribution function (CDF) of the read-out currents of the forming operation performed using ISPVA and CAPVA. As shown in in Figure 5.4 (a), the memristive devices formed at  $-40^{\circ}\text{C}$  operating temperature using ISPVA show the highest current dispersion while the ones formed at  $150^{\circ}\text{C}$  exhibits the lowest current dispersion. This demonstrates that the dispersion of the forming currents reduces with an increase in temperature. Further, the formation of the CF is a stochastic process that needs to be tightly controlled to reduce the D2D variability [95]. The ISPVA programming algorithm provides the necessary energy required for filament formation through increasing pulse amplitudes in small steps of 10 mV. The

current levels of the devices are subsequently verified. This way, the ISPVA scheme makes sure that the devices reach the target current of  $30 \mu\text{A}$  with tighter control over the CF creation. As a consequence, the distribution in the forming currents is also reduced [130].

The forming currents distribution of CAPVA programmed devices shown in Figure 5.4 (b) exhibit no clear trend with change in pulse amplitudes levels, namely, 3 V, 4 V, and 5 V. The scheme consists of a current monitoring step similar to ISPVA. However, due to the inherent stochasticity of the devices, the applied pulse amplitudes are either on the lower or higher side of the actual voltage requirement. Around 44 % and 17 % of the devices remain unformed at 3 V and 4 V pulse amplitudes, respectively. At these amplitudes, the devices with higher forming voltage requirement does not form. Whereas, at higher pulse amplitudes, such as 5 V, the devices with low forming voltage requirements are compelled to form at high voltages causing unnecessary stress on the devices. The absence of the small voltage steps in CAPVA makes the tighter control of the CF formation difficult [130].



**Figure 5.3: (a) Cross-sectional TEM image and (b) the schematic of the integrated 1T-1R memristive device [130].**

Further, the forming yield of the memristive devices is determined by the number of devices reaching the target current value of  $30 \mu\text{A}$ . As shown in Figure 5.5 (a), the devices formed using ISPVA exhibit a forming yield of greater than 94 % at all four temperatures. Whereas in the case of the CAPVA formed devices, the forming yield increases with the voltage pulse amplitude as depicted in Figure 5.5 (b). A high forming yield of 94 % is reached only at 5 V pulse amplitude. These higher pulse amplitudes cause undesired stress on the select transistors eventually leading to their damage [130].

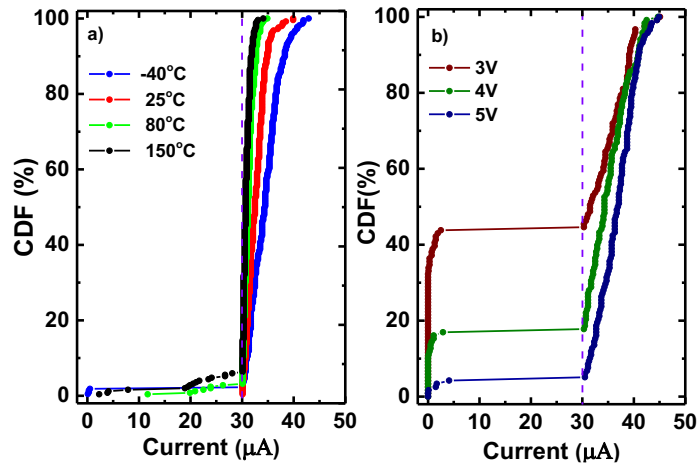


Figure 5.4: D2D CDFs of forming currents of the memristive devices (a) at different temperatures programmed using ISPVA and (b) at different voltage amplitudes programmed using CAPVA [130].

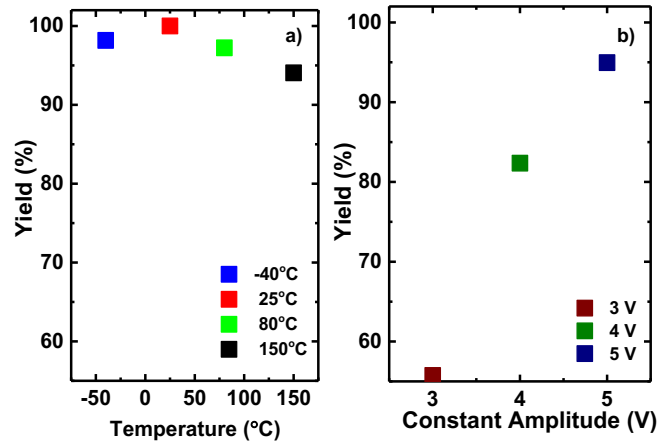


Figure 5.5: Yield of formed memristive devices after the forming operation as function (a) of temperature by using ISPVA and (b) of voltage amplitudes by using CAPVA [130].

The mean values of the read-out currents collected from the forming operations of the memristive devices which are programmed using ISPVA decreases with increase in operating temperature as shown in Figure 5.6 (a). Their dispersion values decrease as well and are illustrated as error bars. The dispersion relation is determined using the Equation 2.1.

$$Dispersion = \frac{\sigma^2}{\mu} \quad 5.1$$

Where  $\sigma$  is the standard deviation and  $\mu$  is the mean of the forming current data set.



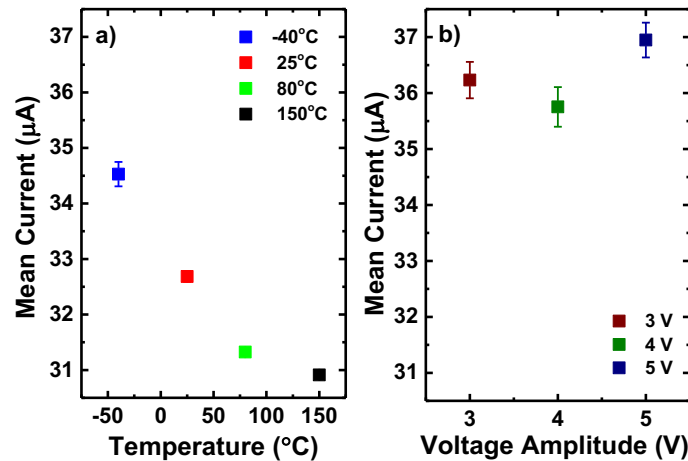


Figure 5.6: Mean values of the read-out currents with error bars measured after the forming procedure of memristive devices (a) at different temperatures by using ISPVA and, (b) at different voltage amplitudes by using CAPVA [130].

The decrease in the conduction of the memristive device with operating temperature is due to the metallic-like conductivity of the CF in its on-state [73], [136]. This behavior of the memristive devices agrees with the results from *Perez et al* [137]. Figure 5.7 demonstrates that the decrease in the conduction of the 1T-1R device is due to the effect of operating temperature on the memristive device alone. The measurements on the transistor and the memristive device are performed independently. The drain current of 22  $\mu\text{A}$  of the transistor remains unaffected by the operating temperature, while the conduction current of the memristive device decreases. Further, Figure 5.3 illustrates the manifestation of the 1T-1R structure as a voltage divider network with voltage drop  $V_1$  and  $V_2$  across the memristive device and transistor, respectively [136]. The decrease in conduction of the memristive device results in an increase in its device resistance which consequently increases the voltage drop  $V_1$  across the memristive device. This in turn reduces the voltage  $V_2$  experienced at the drain of the transistor which eventually decreases the compliance current level imposed by the select transistor. The above explanation also supports the narrow CDFs of forming read-out currents which are programmed using ISPVA at high operating temperatures.

The memristive devices programmed using CAPVA shown in Figure 5.6 (b) do not display any trend in the forming current levels for the change in the pulse amplitude. However, the distribution of their forming currents is found to be three times larger than the ISPVA programmed memristive devices at room temperature [130].

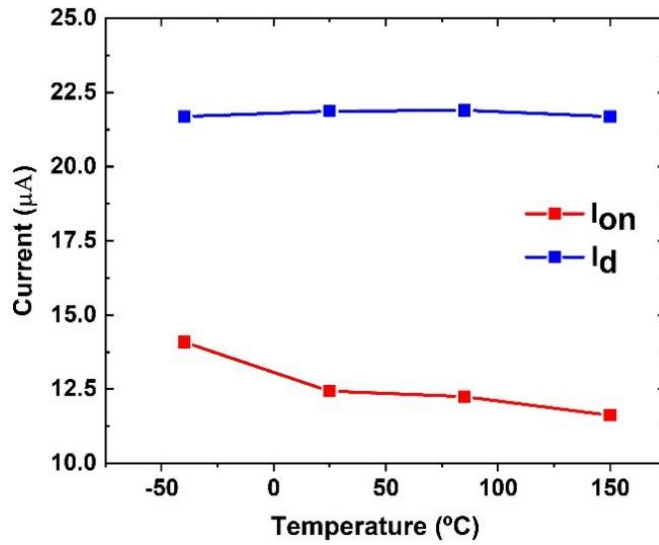


Figure 5.7: General illustration of the drain current ( $I_d$ ) of the transistor at  $V_G = 1.5$  V and  $V_{DS} = 1$  V, without the integrated memristive device and the on current ( $I_{on}$ ) of the memristive device without the transistor contribution at various operating temperatures.

### 5.1.3 Weibull distribution

Besides the forming current, the two other parameters of the forming operation which are variable are, the forming voltage in case of ISPVA programmed devices and forming time in case of CAPVA programmed devices. Instead of their nominal values, the data points of the forming voltage and the forming time are distributed over a range. A statistical distribution type that fits the data points accurately needs to be selected. The fit parameters extracted from the distribution can be used for physics-based modeling of the memristive devices. The material properties of the memristive devices can be strongly related to their electrical characteristics with the right kind of statistical distribution method selected to fit the data points [51].

The Weibull distribution is widely used for modeling the forming operation involving the soft breakdown of the dielectric layer in the memristive devices. It is defined using equation 5.2.

$$F(x) = 1 - \exp\left(-\left(\frac{x}{\eta}\right)^\beta\right) \quad 5.2$$

Where  $F(x)$  defines the CDF of the statistical variable  $x$ , which in this work refers to voltage or time,  $\beta$  the Weibull slope or shape parameter and  $\eta$  the scale parameter or in other words, the value of the statistical variable  $x$  at  $F \approx 63\%$ . The coefficient of variation and the mean values in the normal distribution are comparable to the shape and the scale parameters of the Weibull distribution, respectively [138].

In order to simplify the fitting process, the equation 5.2 can be reduced to a linear form:

$$y = mx + C \quad 5.3$$

After reducing the equation 5.2 to the form in equation 5.3 the final expression obtained is as given below:

$$\ln(-\ln(1 - F)) = \beta(\ln x) - \beta(\ln \eta) \quad 5.4$$

Where the variables  $y$  and  $x$ , the shape parameter  $\beta$  and the scale parameter  $\eta$  are deduced as:

$$y = \ln(-\ln(1 - F)) \quad 5.5$$

$$x = \ln(x) \quad 5.6$$

$$\beta = m \quad 5.7$$

$$\eta = e^{-\left(\frac{C}{\beta}\right)} \quad 5.8$$

### 5.1.3.1 Variability in forming voltages

Initially, the statistical data of the forming voltages obtained by programming the memristive devices using ISPVA are analyzed. Figure 5.8 (a) shows the CDF of the forming voltages at four different temperatures. The range of voltage distributions is similar for all the operating temperatures. However, 90 % of the memristive devices are formed between 2.5 V and 3.75 V at operating temperatures of 80 °C and 150 °C. Further, the mean value of the forming voltages reduces with an increase in operating temperature as depicted in Figure 5.8 (b), but its dispersion remains constant [130].

The breaking of Hf-O and Al-O bonds in the HfAlO-based memristive switching layer and the movement of oxygen anions ( $O^{2-}$ ) towards the top electrode results in the formation of a metal-rich region between the top and bottom electrodes [51]. This metal-rich region is termed CF which is the result of the redox reactions taking place under the influence of external factors such as electric field and temperature [33]. Further, the movement of oxygen anions is often explained in the literature by using oxygen vacancies. The movement of these oxygen vacancies is accelerated under the influence of operating temperatures which assist the formation of CF at lower forming

voltages [130]. This phenomenon explains the behavior of the forming voltage reduction illustrated in Figure 5.8 (b).

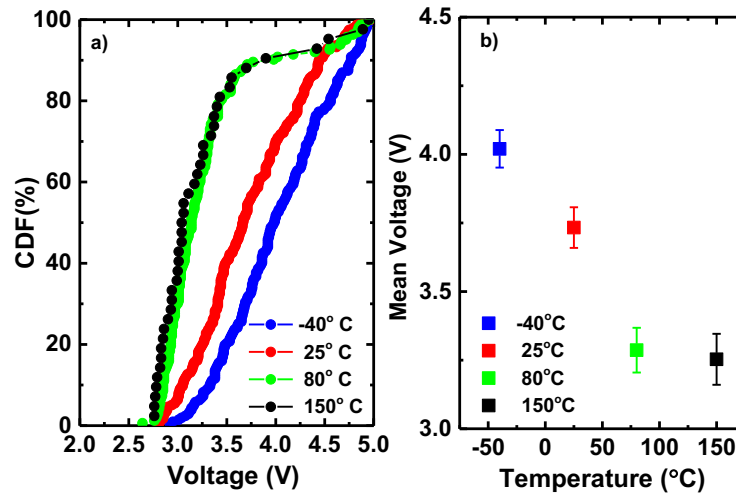


Figure 5.8: D2D (a) CDFs and (b) the mean values, of forming voltages with error bars of the memristive devices at four different temperatures programmed using ISPVA [130].

The statistical data of forming voltages at four different operating temperatures with their respective Weibull fit is depicted in Figure 5.9 using the  $\ln(-\ln(1-F))$  versus  $\ln(V-V_{in})$  plot. The voltage at which the forming takes place is represented by  $V$ . The CDF of the forming voltage values are fit between the initial value of the forming voltage ( $V_{in}$ ) which is 2.75 V and the final value of the forming voltage 5 V. An initial value of 2.75 V is chosen since the memristive devices began to exhibit switching at this voltage. The Weibull fit parameters extracted from the plot in Figure 5.9 are listed in Table 7. Both the shape and scale parameters of the fit decrease with an increase in operating temperature. This trend of fit parameters demonstrates the narrow distribution of the forming voltages with increasing temperatures. Further, the claim that increasing operating temperatures reduces the voltage required to form the initial CF is supported by the reduction of scale parameter [130].

The breakdown of the dielectric is often described in the literature using the percolation model [139], [140]. The voltage stress on the dielectric creates point defects at random sites. When the defects approach a critical level termed critical defect density (CDD), the breakdown event of the dielectric is triggered. This creates an uninterrupted low resistance path between the top and bottom electrodes [141]. Since a similar behavior is observed in the forming operation of the memristive devices, the forming event is often reported in the literature with the help of the percolation model [142], [143]. Consequently, the decrease in the slope parameters illustrated in Table 7

denotes that the CDD required to form the initial CF in the memristive devices reduces with increase in operating temperature [130].

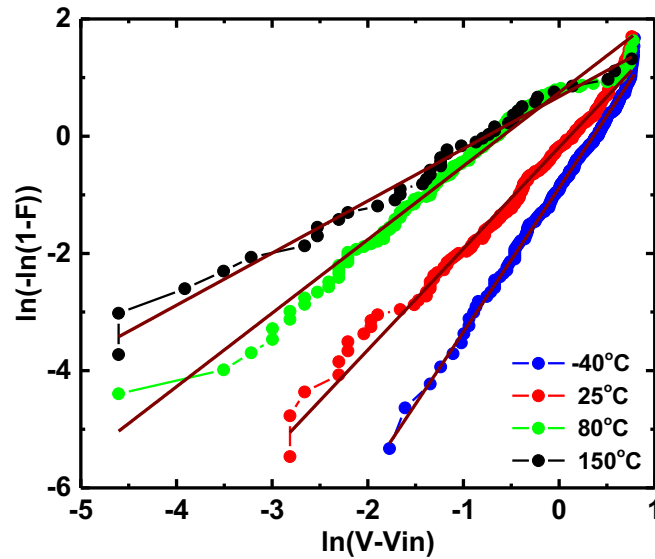


Figure 5.9: Weibull fitted CDFs of forming voltages of the memristive devices at four different operating temperatures programmed using ISPVA [130].

Table 7: Summary of Weibull fit parameters of the memristive devices programmed using ISPVA [130].

Temperature (°C)	Shape parameter ( $\beta$ )	Scale parameter ( $\eta$ )
-40	2.46	1.43
25	1.73	1.12
80	1.25	0.55
150	0.89	0.47

The memristive devices programmed using ISPVA at higher operating temperatures exhibit a narrow distribution of forming currents with reduced forming voltages. The presence of small incremental voltage steps along with current monitoring read-out operation allows the ISPVA programming algorithm to have good control over the creation of the CF. This makes the ISPVA programmed memristive devices suitable for applications related to multi-level resistive switching [130].

### 5.1.3.2 Variability in forming times

The statistical data of the forming times obtained by programming the memristive devices using CAPVA at 3, 4, and 5 V are shown in Figure 5.10. The distribution of the forming time reduces with an increase in voltage amplitude. Further, the forming

times of the ISPVA programmed devices at room temperature are found to be larger than CAPVA programmed devices as shown in Figure 5.10 (a). The time needed for 63 % of the memristive devices to complete the forming operation at three different voltage amplitudes is summarized in Table 8 and shown in Figure 5.10 (b) [130].

For a given memristive layer thickness, the increase in voltage amplitude increases the electric field inside the memristive switching layer. A higher electric field across the memristive layer reduces the energy required to break the Hf-O and Al-O bonds and thereby increases the rate of defect generation [139]. As a consequence, the number of oxygen vacancies generated increases which accelerates the creation of CFs in the memristive switching layers [51]. Hence, the time required to form the memristive devices reduces with an increase in voltage amplitude.

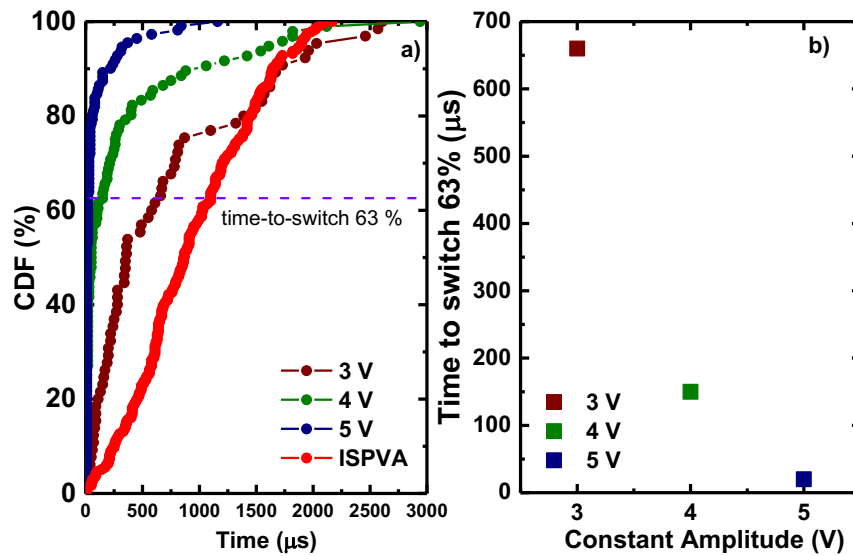


Figure 5.10: (a) The D2D CDFs of the forming times and (b) time to switch 63% of the memristive devices at different pulse amplitudes programmed using CAPVA [130].

The statistical data of forming times at three different constant voltage amplitudes are also fitted using the Weibull distribution as shown in Figure 5.11. Since the pulse width used for the forming operation was 10 μs, a saturation in the distribution was observed at this value. The memristive devices falling in this part of the distribution were excluded to avoid artifacts in the statistical analysis. The extracted fit parameters from the Weibull distribution are summarized in Table 8. Both the shape and scale parameters decrease with an increase in voltage amplitudes. Further, the decrease in the shape parameter supports the previously mentioned statement on the average time required to form a CF which decreases with an increase in voltage amplitude [130].

Table 8: Summary of Weibull fit parameters and forming time of the memristive devices programmed using CAPVA [130].

Constant Amplitude (V)	Shape parameter ( $\beta$ )	Scale parameter ( $\eta$ )	Time to form 63% of the devices ( $\mu\text{s}$ )
-3	0.95	640	660
4	0.45	129	150
5	0.41	23	20

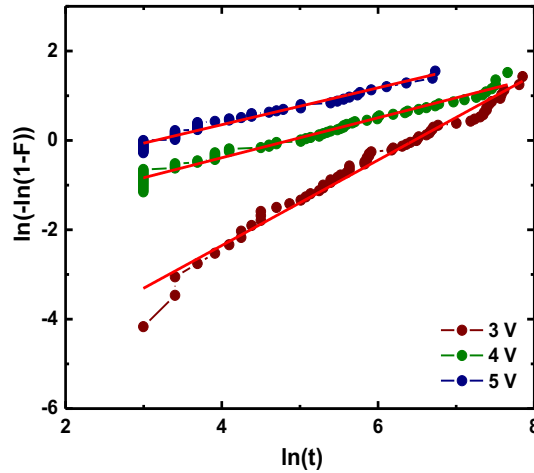


Figure 5.11: Weibull fitted CDFs of the forming times of the memristive devices at three different voltage amplitudes programmed by using CAPVA [130].

The CAPVA programmed memristive devices exhibit a reduced distribution of forming times for the increase in voltage amplitudes. Although this reduction in D2D variability is an advantage, the performance of the select transistors is affected by the higher voltage amplitudes. Further, the lack of adaptability in CAPVA through small voltage steps makes it hard for the programming algorithm to control the morphology of the CF during its creation [130].

## 5.2 Variability in individual 1T-1R memristive devices

Individual access to the memristive device and the transistor inside an integrated 1T-1R memristive array is limited. They can be accessed only together through the peripheral circuits. Further, obtaining the device characteristics in the form of I-V curves under DC conditions is difficult in the case of the memristive array architectures. The I-V characteristics of a physical memristive device are essential for its accurate device modeling. Simulating the functionality of the memristive device based circuits for various applications is made possible through its device modeling [144].

The complete stack of the memristive device starting from the TE, BE, OSL layer, and memristive switching layers can be engineered to enhance the device performance [10]. The electrodes are engineered by considering the factors such as their work function, inertness, layer thickness, etc [33]. The memristive switching layers are engineered as single, bi-, or multilayers of oxides, doping the oxide with an alloy or nanocrystals, among many others [145]–[148]. However, due to the reasons mentioned in the above paragraph, the device characteristics of an engineered memristive layer stack have to be initially tested on individual 1T-1R devices.

Utilizing a bi-layer for memristive switching is one of the approaches reported under the device engineering methods to mitigate variability in memristive devices [10]. Out of many bi-layer memristive stacks, the  $\text{Al}_2\text{O}_3|\text{HfO}_2$ -based memristive devices have grabbed attention due to their CMOS compatibility and BEOL suitable deposition temperatures [49], [149], [150]. In this section, the variability in the set and reset operations of the individual 1T-1R memristive devices consisting of three different memristive layer stacks are investigated.

### 5.2.1 Fabrication of memristive devices

The MEMRES module consisting of the single and bi-layer memristive stack is fabricated between M2 and M3 in the BEOL of the 130 nm BiCMOS technology of IHP. The type 3 approach described in chapter 3 is used for the MEMRES module fabrication and the integration of the module is carried out as explained in chapter 4. The memristive devices are classified into three groups, namely, V1, V2, and V3. The thickness of the layers within the three different variants is illustrated in Table 9. The  $\text{Al}_2\text{O}_3$  and  $\text{HfO}_2$  bi-layers are deposited using the TALD process in the same process chamber without vacuum breakage. Their deposition processes are explained in chapter 2. Figure 5.12 shows the TEM cross section of the memristive device variants with their elemental mapping performed using EDX.

**Table 9: Variants of memristive devices with respective layer thicknesses [131].**

Description	V1 (nm)	V2 (nm)	V3 (nm)
TiN TE	150	150	150
Ti	7	7	7
$\text{HfO}_2$	8	8	8
$\text{Al}_2\text{O}_3$	-	1	2
TiN	150	150	150



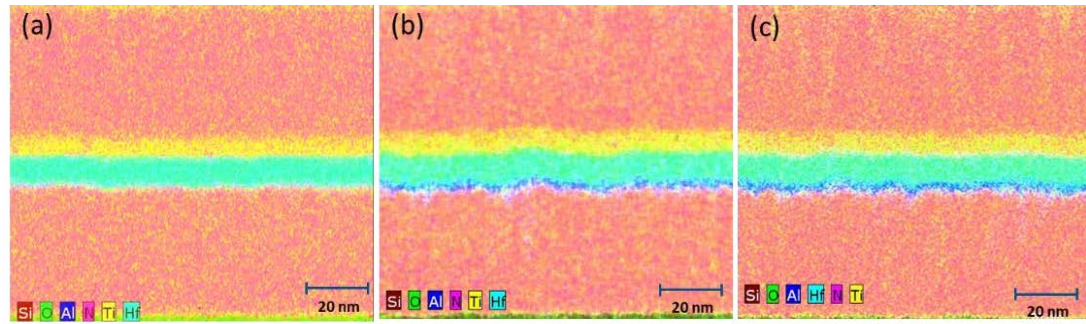


Figure 5.12: Cross-sectional TEM images with EDX elemental mapping of memristor modules in (a) V1 (b) V2 and (c) V3 variants [131].

The  $\text{Al}_2\text{O}_3$  layer is widely investigated as a memristive switching material due to its large band gap and high-k dielectric properties [151], [152]. It is often used in combination with other memristive switching layers in the form of doping [153], a buffer layer [154], [155], a barrier layer [151], and nano-laminate layers [146]. The  $\text{Al}_2\text{O}_3$  layers exhibit stable material properties at BEOL processing temperatures and require high temperatures to crystallize [93]. As illustrated in Figure 5.13 the 6 nm  $\text{Al}_2\text{O}_3$  layers rapid thermal annealed (RTA) at 1000 and 1100 °C for 60 s do not exhibit any diffraction peaks, which indicates their amorphous nature. Further, the possibility of depositing highly uniform and thin  $\text{Al}_2\text{O}_3$  layers using ALD serves as an advantage to use them in memristive layer stack. Figure 5.14 shows a wafer map of thin  $\text{Al}_2\text{O}_3$  layers of thickness ~1 and 2 nm which are deposited with a standard deviation of 1.6 % and 0.6 %, respectively. These thin  $\text{Al}_2\text{O}_3$  layers are used for the bi-layer memristive devices in variants V2 and V3.

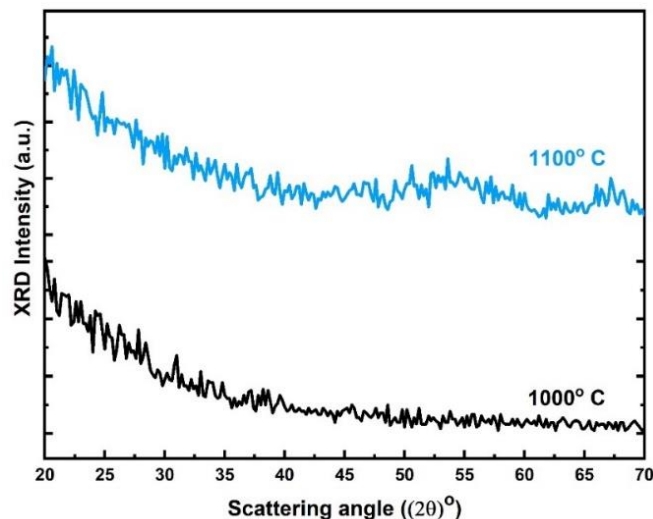


Figure 5.13: XRD patterns of rapid thermal annealed 6 nm  $\text{Al}_2\text{O}_3$  layer deposited using ALD.

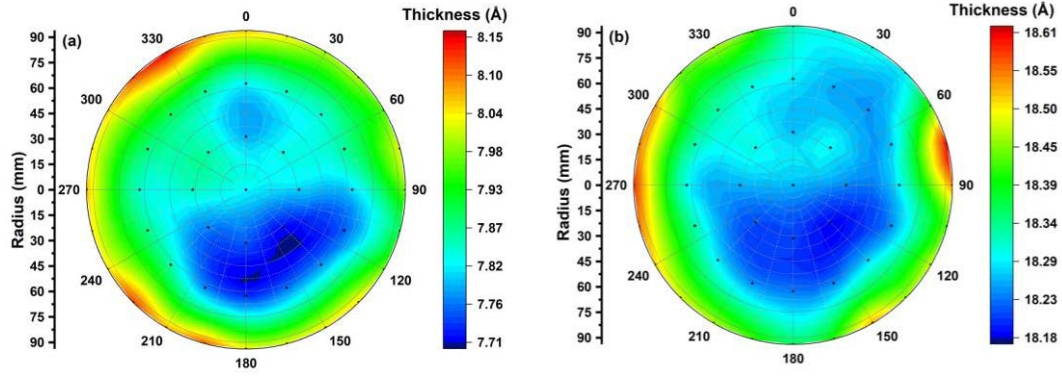


Figure 5.14: Wafermaps of TALD grown  $\text{Al}_2\text{O}_3$  layers of thickness (a) 1 nm and (b) 2 nm on a 200 mm Si-substrate.

## 5.2.2 Variability in memristive device operations

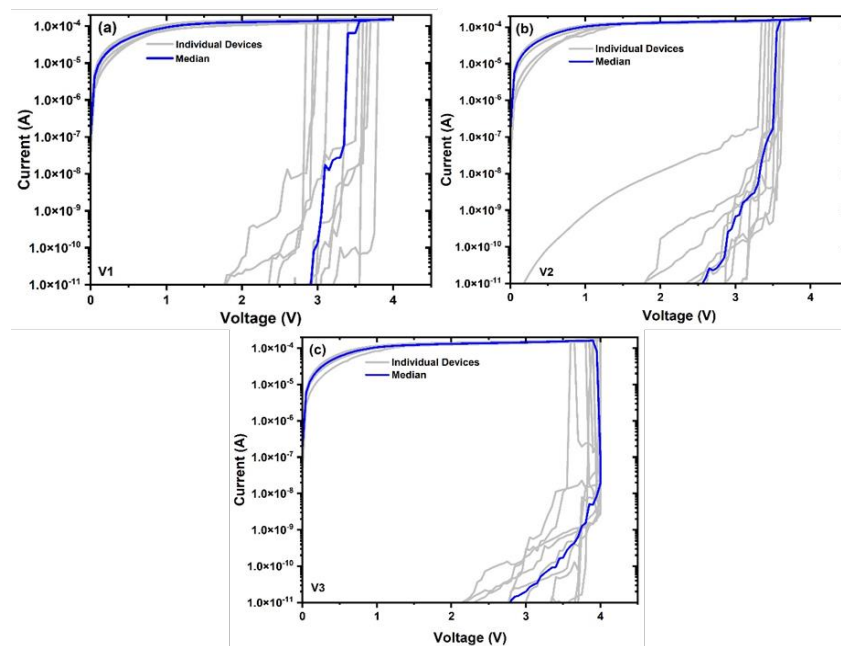
The individual 1T-1R memristive devices are electrically characterized for their memristive switching operations at room temperature. The forming, reset and set operations of the three memristive device variants are performed under identical DC conditions, and the parameters used for their characterization are illustrated in Table 10. The gate terminal is supplied with specific levels of DC bias for various memristive operations. Whereas, the voltages at the source and drain terminals are swept in dual sweep mode or grounded depending on the memristive operation. The three memristive device variants had 10 devices each, considered for their electrical characterization. After the initial forming step, every memristive device is subsequently programmed for 50 cycles of reset and set operations. Further details on the electrical setup used for the characterization of the individual memristive devices are discussed in chapter 2.

Table 10: Electrical parameters for operation of V1, V2, and V3 memristive device variants

Operation	$V_G$	$V_S$	$V_D$
Forming	1.5 V	Ground	0 – 4 V
Reset	2.9 V	0 – 2 V	Ground
Set	1.5 V	Ground	0 – 2 V

The electrical characterization of the memristive devices begin with a forming operation. The individual I-V curves of the forming operation of V1, V2, and V3 memristive devices are illustrated in Figure 5.15. The voltage at which an abrupt change in the current takes place during the I-V sweep is termed as forming voltage. These voltage values are extracted from the individual forming I-V curves, and are plotted in Figure 5.16. It can be seen that the mean value of the forming voltages increases with the total dielectric thickness. The forming voltage of the memristive

devices is mainly dependent on the thickness of the dielectric layer and its dielectric constant. The dielectric thickness varies directly, whereas the square root of the dielectric constant ( $\sqrt{k}$ ) varies inversely with the forming voltage [156], [157]. The total dielectric thickness of the memristive layers in V1, V2, and V3 variants are 8, 9 and 10 nm, respectively. The dielectric constants of Al<sub>2</sub>O<sub>3</sub> and HfO<sub>2</sub> layers are measured to be  $\sim 8.5$  and  $\sim 22$ , respectively. The lower dielectric constant of Al<sub>2</sub>O<sub>3</sub>, and the increase in the effective thickness of the memristive switching layer with its addition to the single layer of HfO<sub>2</sub> is mainly responsible for the increase in forming voltage behavior observed in Figure 5.16. Further, the single layer HfO<sub>2</sub> based memristive device variants (V1) exhibited larger dispersion in forming voltages compared to the bi-layer memristive devices (V2 and V3) [131].



**Figure 5.15: Forming I-V characteristics of (a) V1, (b) V2 and (c) V3 device variants. The characteristics of individual devices from each variant are represented in grey and the computed median curves are represented in blue [131].**

Additionally, memristive devices with a single layer of 6 nm Al<sub>2</sub>O<sub>3</sub>, with and without a 7 nm Ti layer, are electrically characterized to demonstrate their higher breakdown strengths. Their respective MIM stacks are schematically shown in the insets of Figure 5.17 (a), and the TEM cross section of the devices with EDX analysis is shown in Figure 5.17 (b). The devices are fabricated in a BEOL short flow without the FEOL CMOS transistors. The lower breakdown voltages of the memristive devices with a Ti layer compared to the ones without a Ti layer demonstrates the oxygen scavenging properties of Ti layer. Further, as shown in Figure 5.17 (a), the breakdown voltages of single layer Al<sub>2</sub>O<sub>3</sub> based memristive devices are higher compared to the three memristive device variants [131].

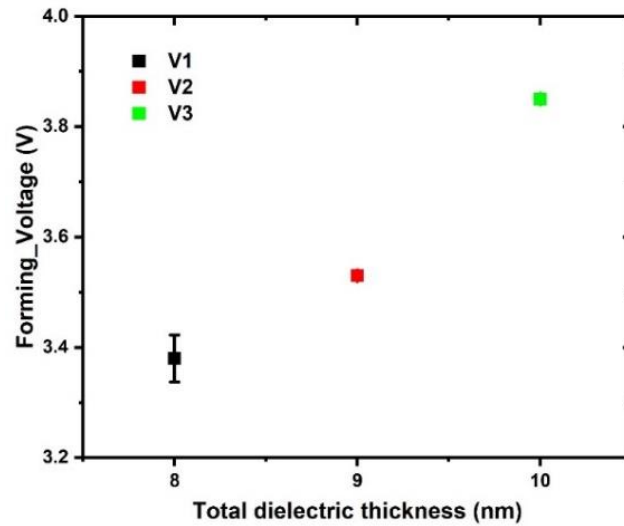


Figure 5.16: Mean values of forming voltages with error bars versus the total dielectric thickness of memristive device variants fabricated in 130 nm CMOS technology of IHP [131].

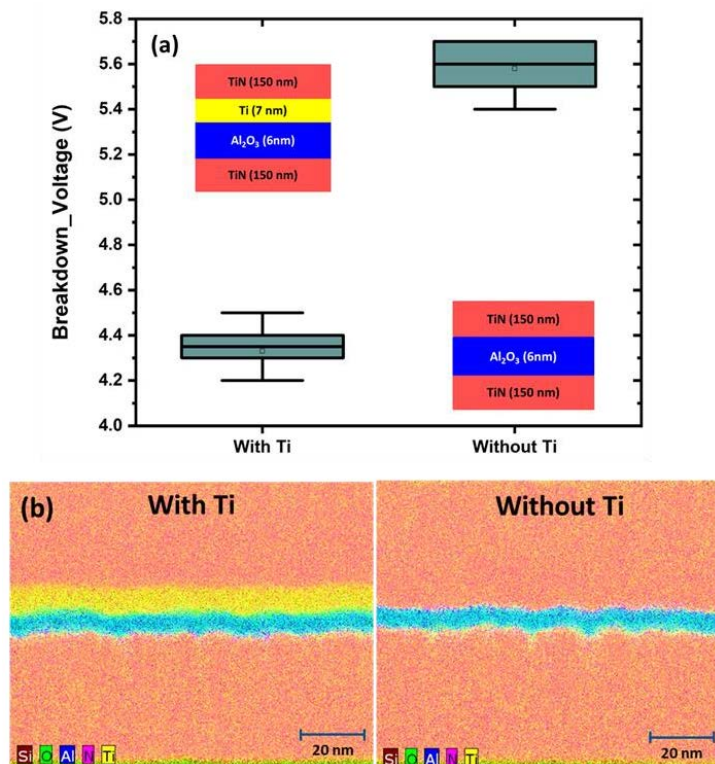
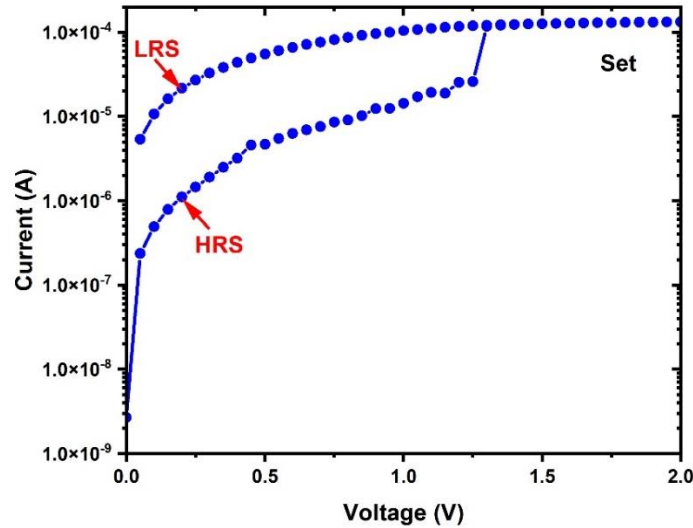


Figure 5.17: The single layer  $\text{Al}_2\text{O}_3$  based MIM devices with and without a Ti layer are (a) electrically characterized for their breakdown voltages (the inset images illustrate the schematic of the layer stack) and (b) material characterized by using cross-sectional TEM images with EDX elemental mapping [131].

The individual memristive devices of three different variants are characterized further for reset and set operations. The variability in the two different resistance states of the memristive devices, namely LRS and HRS are analyzed from the DC set

operations. As shown in Figure 5.18, the HRS and LRS current values are extracted from the forward and reverse bias sweeping of voltage, respectively at 0.2 V ( $V_D$ ) from the DC set operation.



**Figure 5.18: I-V characteristic of a set operation carried out under DC conditions.**

The values from 50 set operations applied on 10 devices of each memristive variant (V1, V2, and V3) are extracted and plotted in the form of box plots as shown in Figure 5.19 and Figure 5.20. The box plots represent the distribution of currents in the two resistance states. The mean value of the LRS and HRS currents increase, whereas their dispersion reduces with the addition of the  $Al_2O_3$  layer to single-layer  $HfO_2$  memristive devices. The increase in the current levels is discussed in the next section under the framework of quantum point contact (QPC) modeling. The reduction in the dispersion of currents in LRS and HRS states of V2 and V3 devices could be due to the oxygen vacancy formation energies being lower in  $Al_2O_3$  layers [153]. As a result, the oxygen vacancies can gather around the Al atoms easily improving the memristive switching properties [131],[154].

The memory window (MW) of the individual memristive devices which are characterized under DC conditions is determined by the ratio of LRS to HRS current values. The mean values of the LRS and HRS currents and the determined MWs are illustrated in Table 11. Although V3 memristive devices exhibited lower MW and higher forming voltages, the reduction in the variabilities of LRS and HRS currents makes them good candidates for multi-level resistive switching operations. Further, the multi-level operation is one of the features sought out for utilizing the memristive devices for neuromorphic computing applications [131].

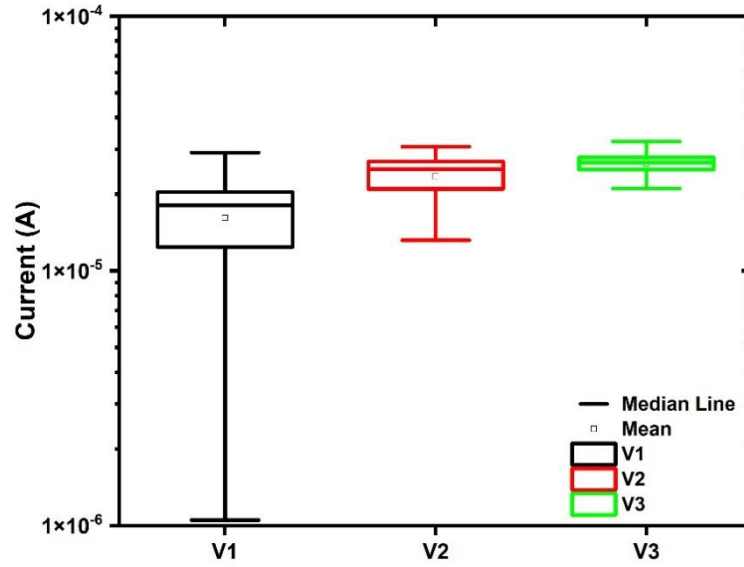


Figure 5.19: LRS currents extracted from DC set operations of V1, V2 and V3 memristive devices [131].

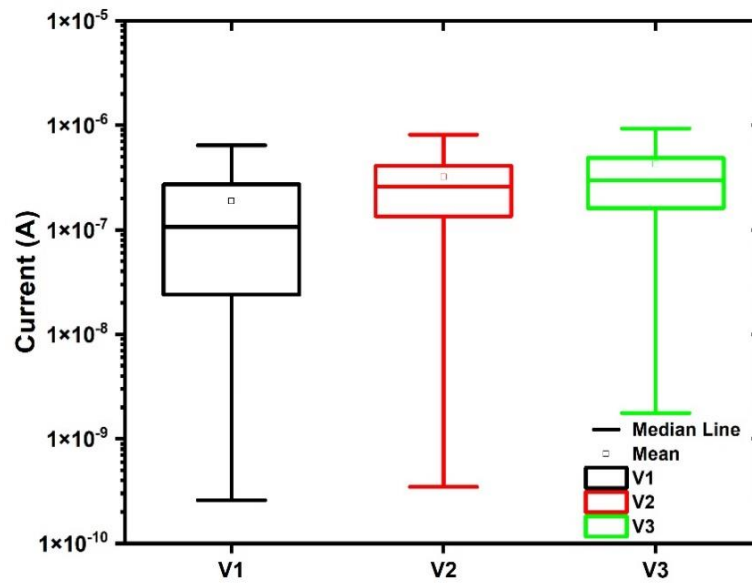


Figure 5.20: HRS currents extracted from DC set operations of V1, V2 and V3 memristive devices [131].

Table 11: The mean values of memory window (MW) of V1, V2 and V3 devices determined from the respective mean values of LRS and HRS currents [131].

Description	V1	V2	V3
Mean LRS	$1.61 \times 10^{-5}$	$2.35 \times 10^{-5}$	$2.63 \times 10^{-5}$
Mean HRS	$1.88 \times 10^{-7}$	$3.2 \times 10^{-7}$	$4.23 \times 10^{-7}$
MW (Mean LRS/Mean HRS)	~86	~73	~62

### 5.2.3 Modeling of conduction filament properties

A memristive device behavior has to be modeled to understand the mechanism behind its switching behavior. The modeling of memristive devices is mainly classified into two types, namely physical models and analytical models [158]. The memristive switching behavior in physical models is explained concerning first principle calculations which involve fundamental device physics. Whereas, an analytical model explains the switching behavior with the help of a mathematical equation which is formulated using the experimental device characteristics of a real memristive device [159]. The physical models are complex and incur an overhead of computational cost. Whereas, the analytical models, which are physics-based, are compact compared to the physical models and also provide an understanding of the underlying physical mechanism behind the switching behavior of the memristive devices [160].

The conduction mechanism of memristive devices is reported in the literature using various physical models such as tunnel barrier [161], Schottky emission [162], thermal activated hopping conduction [160], and Pool-Frenkel conduction [163] among many others. However, these models fail to describe the conduction behavior of the memristive devices in the on-state where the CF exhibit metallic-like conductivity. A physics-based analytical model such as quantum point contact (QPC) can model the CF properties of memristive devices in both on and off states [57]. Hence, the conduction properties of V1, V2, and V3 device variants in this work are explained under the framework of the QPC model.

The QPC model was initially proposed to describe the conduction mechanism in metal-oxide-semiconductor (MOS) transistors after the breakdown of the gate dielectric. Both, the soft and hard breakdown conduction modes are successfully modeled using QPC in SiO<sub>2</sub>-based gate dielectrics [164]. The model was later adapted to describe the properties of the CF in LRS and HRS of the memristive devices [165]. The QPC model assumes that irrespective of the memristive state of the device, the current flows through the CF formed between the two metal electrodes [166]. According to Miranda et al., the current in the LRS flow through a bundle of N conducting channels, whereas the current in the HRS flows through a single channel with a potential barrier [165]. Accordingly, the flow of current in the HRS of the device is given by equation 5.9 [165].

$$I = \frac{2e}{h} \frac{G}{G_0} \left[ eV + \frac{1}{\alpha} \ln \left\{ \frac{1 + e^{\alpha[\phi - \beta eV]}}{1 + e^{\alpha[\phi + (1-\beta)eV]}} \right\} \right] \quad 5.9$$

Where  $I$  is the current flowing through the CF,  $V$  is the applied voltage, and  $\beta$  represents the voltage drop fractions at cathode and anode interfaces [57]. Due to the asymmetry of the CF constriction, almost the entire applied voltage drops close to the oxygen scavenging Ti layer and hence,  $\beta$  was estimated to be 1 [95]. Further,  $e$  is the electron charge,  $h$  is the Planck's constant, and  $G/G_0$  is the conductance parameter which corresponds to the number of conducting channels  $N$  at low voltages, which is equal to 1 for HRS. The potential barrier is created due to the re-oxidation of the vacancies in a single conduction channel [57]. The height of the potential barrier is defined by the parameter  $\phi$  and the thickness of the barrier is defined by  $T_B$ . The parameter  $\alpha$  is related to the thickness of the potential barrier. As described by Lian et al. [167], at high enough potential barriers and low voltages, equation 2.1 converges to equation 2.1 .

$$I = \frac{2e}{h} N e^{(-\alpha\phi)} \left[ V + \frac{\alpha\beta}{2} V^2 \right] \quad 5.10$$

$R$  is the resistance value in series, external to the CF constriction. In the case of the memristive device without a select transistor, the external resistance could be due to the cathode and anode metal electrodes. This external resistance is smaller compared to the resistance due to the potential barrier in HRS and hence neglected while evaluating the HRS current of the memristive device. In the case of the 1T-1R memristive devices, the external resistance is due to the select transistor which is connected in series with the memristive device. Irrespective of the nMOS terminals (source/drain) used for biasing the transistor during the set and reset operations, the transistor is operated in its linear or saturation region. The transistor resistance in these regions is negligible in comparison to the resistance of the memristive device in HRS and hence not considered for QPC evaluation.

The potential barrier width ( $T_B$ ) in the HRS of a memristive device is defined as below [165].

$$T_B = \frac{h\alpha}{2\pi^2} \sqrt{\frac{2\phi}{m^*}} \quad 5.11$$

Where  $m^*$  is the effective electron mass within the CF.

The radius of the CF constriction ( $R_B$ ) in the HRS of a memristive device is defined as below [165].

$$R_B = \frac{hz_0}{2\pi\sqrt{2\phi m^*}} \quad 5.12$$



Where  $z_0$  is the first order Bessel function, whose value is equal to 2.404.

In the LRS of the memristive device, the CF exhibits a metallic-like conductivity [73]. As a result, the potential barrier collapses and its value is considered as zero [167]. Further, a linear I-V characteristic is observed in the LRS of the memristive device. The external series resistance in the LRS of 1R or 1T-1R devices is comparable to the memristive device and hence cannot be neglected. The value of R for the on resistance of the transistor in this work is determined from the cadence simulations and the electrical characterization of a physical nMOS transistor. The obtained value of R was 3 k $\Omega$ . The conduction of the LRS of a memristive device is expressed using the equation 5.13 [166].

$$I = \frac{NG_0V}{1 + NG_0R} \quad 5.13$$

Where  $G_0 = 2e^2/h$  is the quantum conductance unit. Its value is determined to be one over (12.9 k $\Omega$ )<sup>-1</sup>.

The experimental I-V characteristics of V1, V2, and V3 memristive devices in LRS and HRS are fitted using the QPC model defined in equation 5.13 and equation 5.10, respectively. The mean values of the I-V curves from the set operation under DC conditions in the  $V_D$  range from 0 to 0.5 V are used for the QPC fit of LRS shown in Figure 5.21. The parameter N which is the number of conduction channels or filaments is extracted from the fit and the values are illustrated in Table 12. The fit parameter N increases with the addition of Al<sub>2</sub>O<sub>3</sub> layers. Hence, the conductivity of the devices increases accordingly. This signifies the formation of stronger conduction paths with the addition of Al<sub>2</sub>O<sub>3</sub> layers into a bi-layer memristive device stack [131].

**Table 12: The fitting parameter N extracted from the QPC model fit for LRS curves [131].**

Description	N
V1	1.43
V2	2.33
V3	2.73

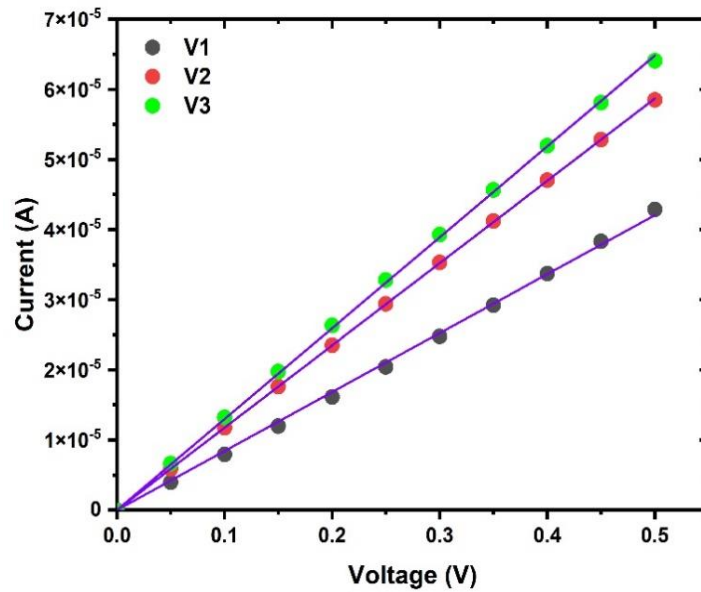


Figure 5.21: Experimental I–V characteristics (symbols) associated with the mean LRS currents extracted from DC set operations of V1, V2 and V3 memristive devices. The solid lines are simulated characteristics using the QPC model for LRS [131].

The mean values of the I-V curves from the reset operation under DC conditions in the  $V_s$  range from 0 to 0.5 V are used for the QPC fit of HRS in the memristive devices as shown in Figure 5.22. The parameters  $\alpha$  and  $\phi$  are extracted from the fit and the values are illustrated in Table 13. The fit parameters obtained for the  $\text{HfO}_2$ -based memristive devices (V1) are comparable with the results from Grossi et al. [168].

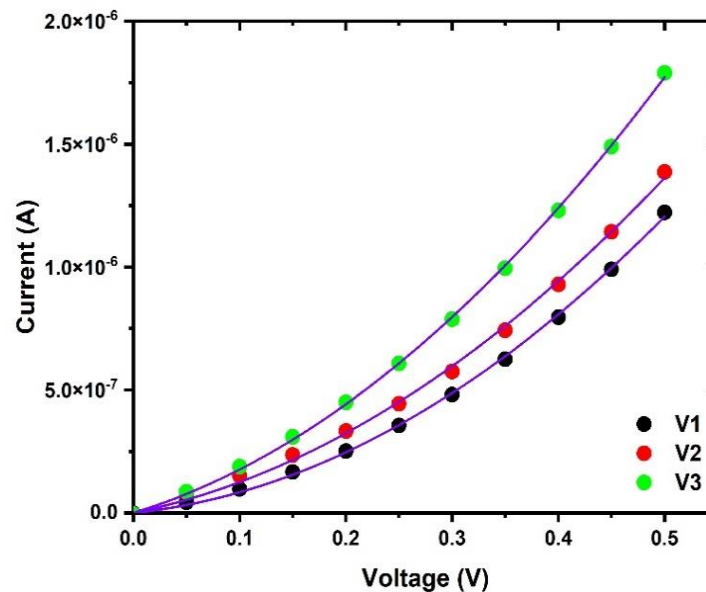
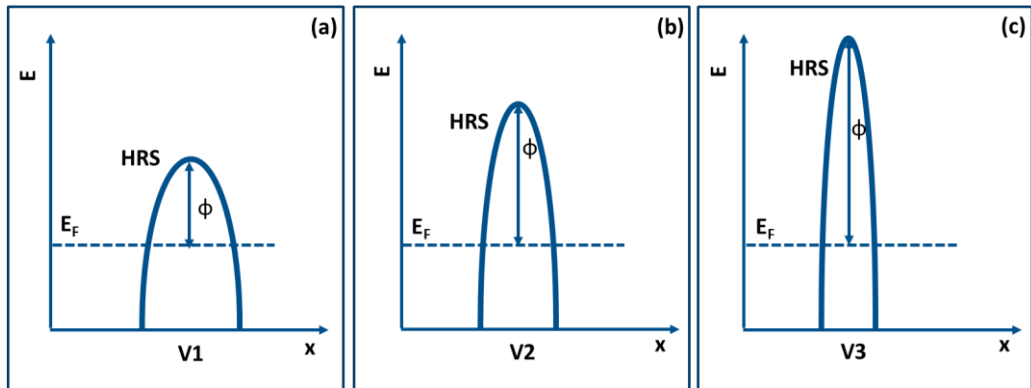


Figure 5.22: Experimental I–V characteristics (circles) associated with the mean HRS currents extracted from DC reset operations of V1, V2 and V3 memristive devices. The solid lines are simulated characteristics using the QPC model for HRS [131].

**Table 13: The fitting parameters  $\alpha$  and  $\phi$  extracted from the QPC model fit for HRS curves and the determined ratios of  $T_B/R_B$  [131].**

Description	$\alpha$	$\phi$	$T_B/R_B$
V1	17.61	0.29	1.35
V2	8.37	0.53	1.17
V3	6.75	0.6	1.07

The addition of  $\text{Al}_2\text{O}_3$  layers into the bilayer memristive stack increases the potential barrier height ( $\phi$ ) in the HRS of the memristive devices. The schematic of the potential barrier height for the three memristive device variants is illustrated in Figure 5.23. The increase in  $\phi$  should result in the decrease of the HRS currents but an opposite behavior is observed i.e. the HRS currents increase. Further, the movement of oxygen vacancies in  $\text{Al}_2\text{O}_3$  layers is lower compared to the  $\text{HfO}_2$  layers [149]. This reduction in the mobility of oxygen vacancies limits the re-oxidation of the CF near the BE during the reset operation [57]. Hence, the addition of  $\text{Al}_2\text{O}_3$  layers results in more localized constriction of a CF, increasing the HRS current levels. Additionally, the decrease in the shape parameter  $\alpha$  and the ratios of  $T_B/R_B$  determined from equations 5.11 and 5.12 support the claim. The values  $\alpha$  and  $T_B/R_B$  are illustrated in Table 13. To avoid the complex estimations and calculations concerning the effective electron mass within the CF of a bi-layer memristive device stack, the  $T_B/R_B$  ratios are calculated instead of their individual values [131].



**Figure 5.23: Schematic representation of the energy band diagram of the conductive filament potential barrier in (a) V1, (b) V2 and (c) V3 memristive devices.  $E$  is the energy of electrons,  $x$  is the direction of current flow in the filament,  $E_F$  is the Fermi level and  $\phi$  is the potential barrier height with respect to the Fermi level [131].**

## 5.2.4 Conclusion

The variability issue in memristive devices are discussed in this chapter regarding the electrical characterization approach in the memristive arrays, and the device

engineering approach in the individual memristive devices. In the former approach, the variability in the forming operation of the memristive arrays is investigated by using the ISPVA and CAPVA programming algorithms. The forming operation performed on the memristive devices at higher temperatures using ISPVA exhibited reduced distribution of forming currents and lower values of forming voltages. The small incremental pulse amplitude steps together with the current verify operation enable ISPVA to gain control over the formation of the CF. Thus, ISPVA enables the multi-level programming of memristive devices. The forming operation performed on the memristive devices using CAPVA exhibited a reduction in the forming times and forming currents distribution at high voltage amplitudes. The variability in CAPVA is reduced at the cost of high voltage amplitudes which could affect the performance of the select transistors. Further, the control over the CF morphology was difficult to achieve in CAPVA due to the absence of small voltage steps.

In the latter approach, the variability in the memristive switching operation is discussed in the single and bi-layer individual memristive devices. Adding the TALD-grown thin  $\text{Al}_2\text{O}_3$  layers to  $\text{HfO}_2$ -based memristive devices increased the forming voltage and reduced the variability in the on and off-state currents. The properties of the CF in the on and off state of the memristive devices are discussed under the framework of the QPC model. The experimental I-V curves from the set and reset operations of the three memristive device variants, namely V1, V2, and V3 fit accurately with the QPC model for LRS and HRS. The parameter obtained from the LRS fit signified the increase in the number of CFs with the addition of  $\text{Al}_2\text{O}_3$  layers. Whereas, the HRS fit parameters signified the increase in potential barrier height and decrease in potential barrier thickness. As a result, the constriction of the CF became more localized and increased the HRS current levels in V2 and V3 memristive devices. Finally, a combination of device engineering and electrical characterization approaches could strongly reduce the variability in memristive devices.

## Chapter VI

---

# 6 Towards Neuromorphic Computing by Exploiting Switching Variabilities in HfO<sub>2</sub>- based Memristive Devices

---

Neuromorphic computing is a human brain-inspired emerging computation technique that aims to improve computing efficiency while lowering power consumption compared to conventional Von-Neumann-based architectures [169]. Implementing this computation technique requires artificial neural networks consisting of artificial neurons and synapses which can emulate the functionalities of biological counterparts [170].

An artificial neural network (ANN) can be built on software, hardware or both platforms [171]. The software ANNs use Von-Neumann-based hardware for their computation which limits their throughput due to the latency delay between the memory and logic units [171]. Implementing the ANNs on microelectronic hardware has gained momentum due to the advent of in-memory computing [172]. Further, utilizing the memristive devices to emulate the behavior of the biological synapses has significant advantages. Firstly, the memristive devices are CMOS compatible and

fabricated in the BEOL of a CMOS baseline technology. Secondly, it is feasible to monolithically integrate the memristive devices with the FEOL CMOS circuitry which can emulate the performance of artificial neurons [66].

The ANN architectures are broadly classified into two types: (a) Neuro-science focused and (b) Deep learning-focused [171]. Neuroscience-focused architectures are named spiking neural networks (SNNs) [173]. They meant to understand the fundamental operations of biological neuronal networks and mimic their functionality on the microelectronic hardware or software [171]. Deep learning-focused architectures are named deep neural networks (DNNs) [174]. They illustrate the primitivities of the biological neuronal networks and executed on the conventional Von-Neumann-based architectures such as graphics processing units (GPUs) and CPUs [171]. In the case of SNNs, the memristive device requirements are relaxed and the sensitivity of the network to the device variability is also lower [175]. Whereas, DNNs demand very stringent performance from the memristive devices regarding the symmetry of switching, switching variabilities, number of switching states, etc., [175]–[177].

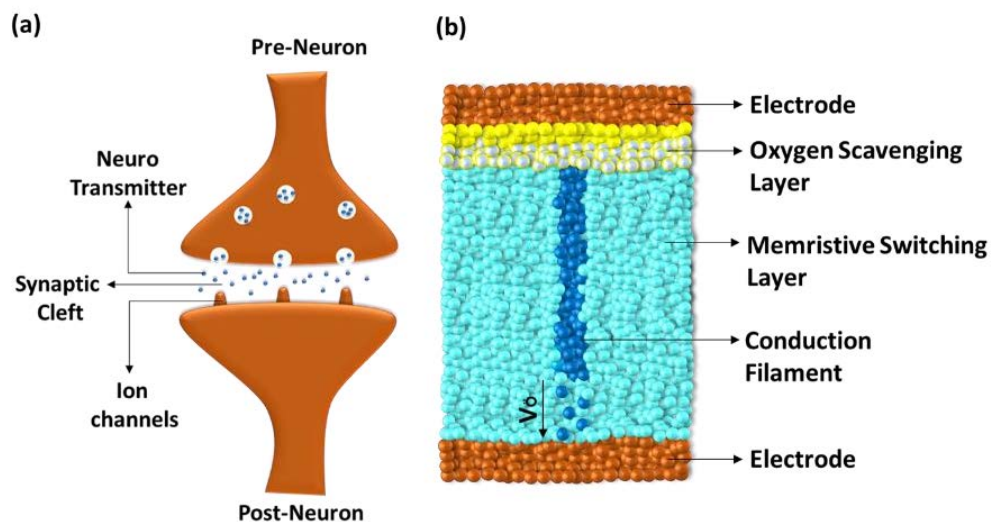
As mentioned in chapter 5, the formation and rupture of a CF in the memristive devices is a stochastic process that causes switching variabilities. The stochasticity involved in the switching operations can be exploited in ANNs [178]. Further, stochasticity can be introduced into an ANN either through neurons or through synapses, and the network is termed as stochastic artificial neural network (StochANN) [179]. Such stochastic computing systems are particularly tolerant to noise and errors [180].

This chapter demonstrates how utilizing the switching variability of the HfO<sub>2</sub>-based filamentary memristive devices, the analog synaptic functionality can be mimicked. In this regard, the memristive switching layers are grown using atomic vapor deposition (AVD) at 400 °C targeting the variability. Further, the fabricated memristive devices are electrical characterized to investigate the switching variabilities in LRS and HRS. Finally, the reliability of the memristive devices as electronic synapses is discussed with the help of cyclic endurance and retention measurements.

*“The investigated samples in this work were fabricated by the cleanroom staff at the IHP pilot line. The experimental work on the memristive arrays was conceptualized by Professor Christian Wenger and Professor Martin Ziegler. The devices were electrically characterized by Dr. Eduardo Perez. I performed the processing of the devices and supported the electrical measurements and evaluation of data. Parts of this chapter have been published in Ref. [181].”*

## 6.1 Memristive devices as synapses

Neurons and synapses are the fundamental units of the human brain. The neurons are connected to each other through the synapses [182]. They communicate with one another through an electrical impulse called action potential [183]. When an action potential from a pre-neuron reaches the pre-synaptic terminal, the neurotransmitters are released into the synaptic cleft as shown in Figure 6.1 (a). Further, the binding of these neurotransmitters with specific receptors causes the ion channels to open or close [184]. This results in the conductance change of a synaptic cell and the strength of the synapse is altered before the action potential reaches the post-neuron [183]. This mechanism of the change in conductance, based on the neuronal activities is called synaptic plasticity. It is a stochastic process responsible for the creation of memory and learning in a human brain [183].



**Figure 6.1:** (a) Schematic representation of a biological synapse. (b) Schematic representation of a memristive device which can be used as an electronic synapse.

As discussed in chapter 3, the memristive device consists of a MIM stack. The OSL layer of the stack scavenges the oxygen ions from the memristive switching layer under the influence of the electrical stimuli. This results in a sub-stoichiometric memristive switching layer which leads to the formation of a CF [124]. As mentioned in section 5.1.3, the movement of the oxygen anions is often described using oxygen vacancies. The generation and migration of oxygen vacancies inside a memristive switching layer (as shown in Figure 6.1(b)) leads to a change in its resistance [183]. Depending on the direction of the applied electrical stimuli, a CF can be annihilated and recreated multiple times by sending the device to off-and-on states, respectively. This functionality of a memristive device can be used to mimic the behavior of the

biological synapses in microelectronic hardware for neuromorphic computing applications [178].

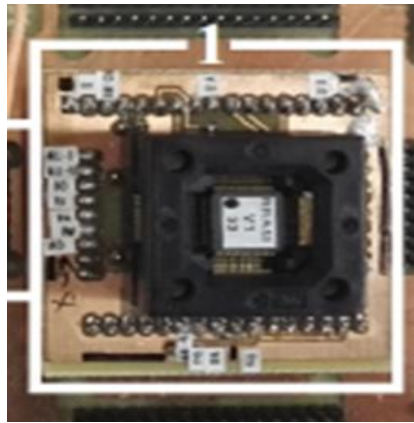
In order to realize the artificial synapses with memristive devices, the devices must exhibit multiple conductance states [169]. This property is realized through multi-level programming (MLP) or analog switching [67][45]. The MLP is achieved by varying the gate voltage or fixing the current compliance in the memristive devices with and without select transistors, respectively. The analog behavior can be achieved through the combinations of different memristive switching layers [45]. Further, the high variability in the switching operations of the memristive devices can also be used for emulating the performance of biological synapses. Rather than implementing the discrete resistance states of the memristive devices, their switching probabilities could be employed into the learning algorithms of the StochANNs [181].

## 6.2 Fabrication and electrical characterization

The 1T-1R memristive arrays utilized in this chapter are fabricated using the SGB25V technology at IHP. A detailed description of the fabrication process is presented in section 4.2. The memristor module consists of a TiN/HfO<sub>2</sub>/Ti/TiN stack. The thickness of the TE, BE, and OSL mentioned in Chapter 3 are maintained here as well. Unlike the memristive switching layers in Chapter 3, the layers used here are grown using atomic vapor deposition (AVD) instead of ALD, in order to target the maximum variability. Accordingly, an 8 nm thick HfO<sub>2</sub> layer is grown at 400 °C in a polycrystalline state. A polycrystalline phase of an oxide consists of grain boundaries (GBs) which are the lattice mismatch points [185]. These GBs serve as the defect sites and are the preferential locations for the creation of percolation paths which eventually become a CF [50]. This conduction mechanism at the GBs results in large device-to-device variability in the memristive devices [168].

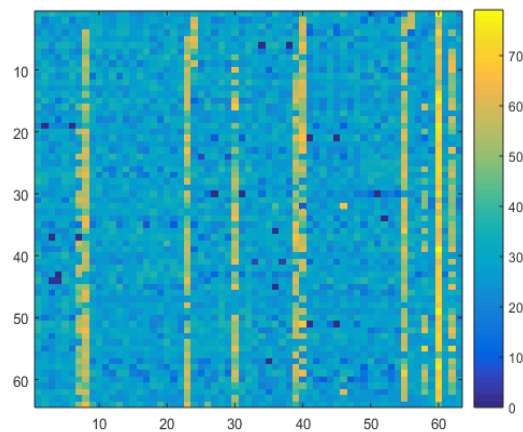
The integrated circuit (IC) of a 4 kbit memristive array is packaged before its electrical characterization as shown in Figure 6.2. The memristive devices are investigated to evaluate their D2D variability and reliability. Irrespective of the purpose of their investigation, a forming operation with a target current of 20  $\mu$ A is initially performed on all the memristive devices at room temperature. Due to the advantages of the ISPVA mentioned in Chapter 5, it was used for forming the memristive devices. The read-out current operation is carried out at 0.2 V pulse amplitude, and the pulse width of 10  $\mu$ s is kept constant for all the operations.



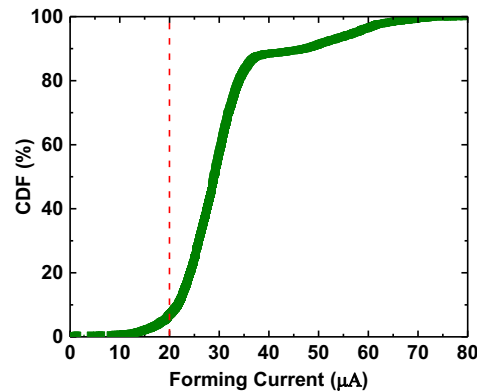


**Figure 6.2:** Photograph of the packaged 4 kbit memristive array placed in the socket of the experimental set-up [181].

The current map of the forming currents in 4 kbit memristive arrays is depicted in Figure 6.3. It can be seen that the devices are formed with minor defect cells. Further, the forming yield of the memristive devices is determined by the number of devices crossing the target current of the forming operation. As depicted in the CDF plot of the forming currents in Figure 6.4, a forming yield of 95 % is achieved [181]. Despite using ISPVA for the forming operation, a wide distribution of forming current values from 20  $\mu$ A to 80  $\mu$ A is observed. The stochasticity involved in the CF formation along the grain boundaries in the polycrystalline phase of the HfO<sub>2</sub> layer is responsible for the large D2D variabilities observed [168].

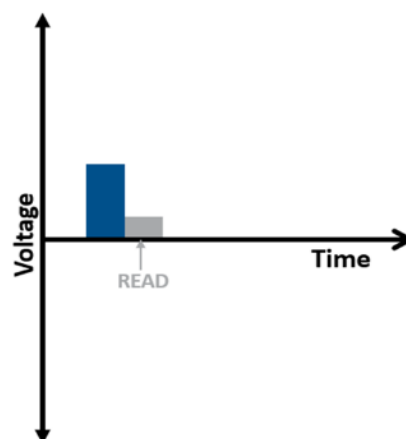


**Figure 6.3:** Current map of the 4,096 memristive devices after forming, in a packed 64  $\times$  64 array [181].



**Figure 6.4:** CDF of forming currents of the devices programmed by using ISPVA [181].

After the completion of the forming operation, the memristive devices used for evaluating the D2D variability are divided into two groups. The devices in the first group are programmed for the LRS through set operations, while the second group devices are programmed for the HRS through reset operations. The ISPVA is used for the programming, and the target current values of 20  $\mu\text{A}$  and 5  $\mu\text{A}$  are used for the set and reset operations, respectively. The D2D variability of the memristive devices in the LRS and HRS evaluated independently. This evaluation is performed using trains of 100 single programming pulses (one-single-pulse mode) at three different voltage pulse amplitudes. The schematic of a single programming pulse is shown in Figure 6.5. It is worth mentioning that the memristive devices used for the variability evaluation are not cycled. Further, a threshold current criterion is set for LRS and HRS, and the number of devices crossing that criterion is noted. The switching probability of the devices is determined from the aforementioned criterion [181].



**Figure 6.5:** Schematic illustration of one-single-pulse mode operation. Adapted from [181].

Finally, the memristive devices are characterized to evaluate their performance and reliability as electronic synapses through retention and endurance measurements. The ISPVA is used for programming the memristive devices. It is important to mention that the memristive devices are cycled in endurance measurements. The devices are cycled up to 10,000 cycles. The endurance measurements were carried out on four different sets of devices with four different cycle numbers, namely 1, 100, 1000, and 10,000. At the end of the aforementioned endurance cycles, one-half of the devices are programmed for LRS, and the remaining half are programmed for HRS using ISPVA. Further, the retention measurements are carried out on the LRS and the HRS programmed memristive devices at different time intervals of 1, 10, and 100 hours at 125 °C.

### 6.3 Evaluation of variability in memristive devices

The variability in the memristive devices is evaluated with respect to voltage pulse amplitude and pulse numbers. As mentioned in the previous section, after the initial forming operation, the memristive devices are divided into two groups and programmed to LRS and HRS, respectively. The threshold criteria for the LRS and HRS transition are set as 20 and 10  $\mu$ A, respectively. The D2D variability concerning the transition of the memristive devices from LRS to HRS is determined at voltage pulse amplitudes of -0.6, -0.9, and -1.2 V. Whereas, for the HRS to LRS transition, the D2D variability is determined at voltage pulse amplitudes of 0.6, 0.9, and 1.2 V. Further, the memristive devices are characterized by applying 100 single pulses on each device at aforementioned voltage pulse amplitudes. The pulse number dependent variability is evaluated at pulse series of 1, 10, and 100 for both from LRS to HRS and from HRS to LRS transitions [181].

In the case of the LRS to HRS transition, the impact of the voltage pulse amplitude and the pulse number are found to be negligible at -0.6 V as illustrated in Figure 6.6 (a). Whereas at -0.9 V, the number of devices crossing the threshold current criteria for HRS increases from 3 % for the first pulse to 33% for the 100<sup>th</sup> pulse, as illustrated in Figure 6.6 (b). Finally, at -1.2 V, the device transition to HRS increase from 75 % for the first pulse to 95% for the 100<sup>th</sup> pulse as shown in Figure 6.6 (c) [181].

Similar to the previous case, the transition of the memristive devices from HRS to LRS at 0.6 V is found to have negligible impact regardless of the number of pulses applied, as shown in Figure 6.7 (a). At 0.9 V, the HRS to LRS transition of the devices increased strongly from 20 % to 65 % for the series of pulses from 1<sup>st</sup> to 100<sup>th</sup>,

respectively, as illustrated in Figure 6.7 (b). The impact of the pulse number vanishes at a voltage pulse amplitude of 1.2 V, as illustrated in Figure 6.7 (c). It can be observed from Figure 6.6 and Figure 6.7 that the switching characteristics of the memristive devices demonstrates a strong dependency on the voltage pulse amplitude and the pulse numbers used for their programming [181].

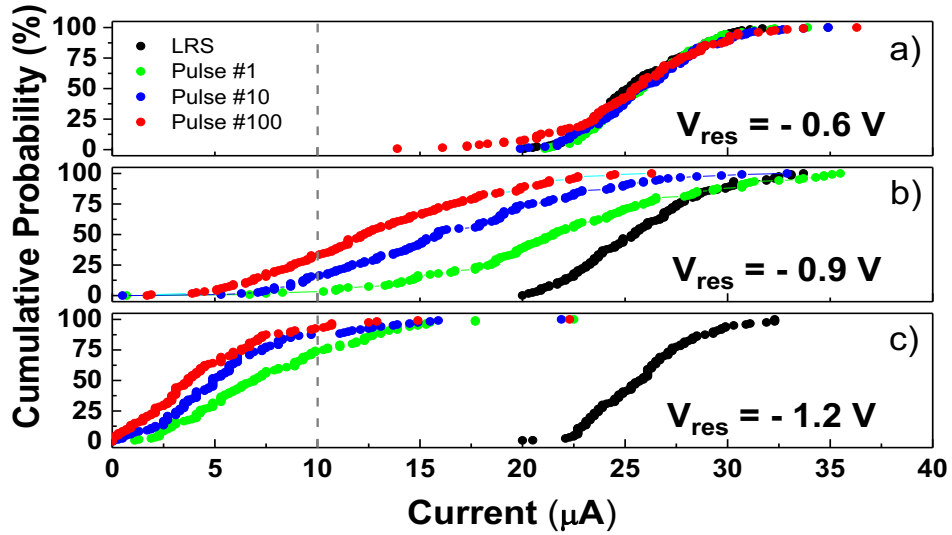


Figure 6.6: Evolution of the D2D distributions of the 1T-1R cells after applying a series of 100 pulses with amplitudes of  $-0.6 V$  (a),  $-0.9 V$  (b) and  $-1.2 V$  (c). The dotted lines represent the HRS threshold current value of  $10 \mu A$  [181].

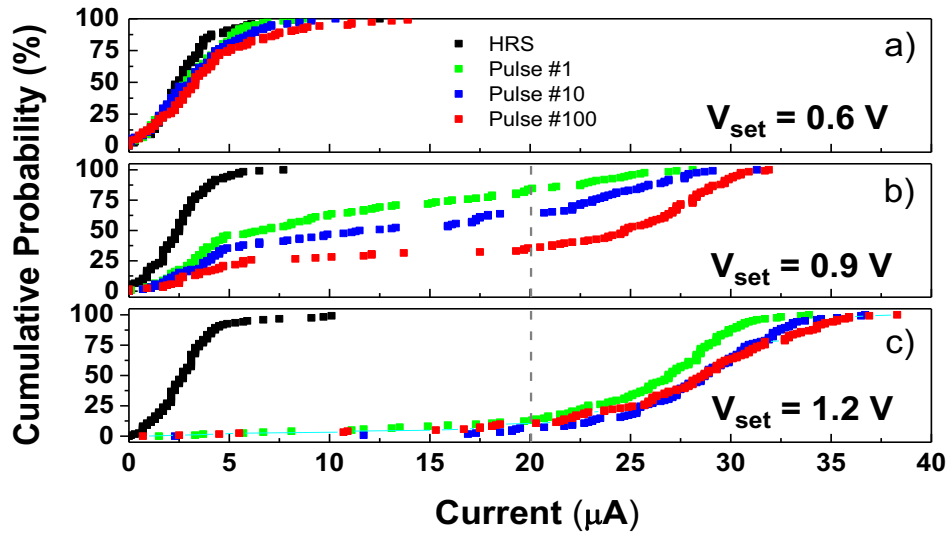


Figure 6.7: Evolution of the D2D distributions of the 1T-1R cells after applying a series of 100 pulses with amplitudes of  $0.6 V$  (a),  $0.9 V$  (b) and  $1.2 V$  (c). The dotted lines represent the LRS threshold current value of  $20 \mu A$  [181].

The D2D variabilities are evaluated further at voltage pulse amplitudes ranging from -0.5 to -1.8 V and 0.4 to 1.4 V for reset and set operations, respectively, over a pulse series of 1, 10, and 100. Further, as described in Wenger et al., the resulting read-out currents of the memristive devices after applying a defined voltage pulse amplitude are gaussian distributed [178].

The switching probability of the memristive devices is determined using equation 6.1.

$$\text{Switching probability} = \frac{\text{No. of devices crossing the threshold current}}{\text{No. of working devices}} \quad 6.1$$

The switching probability of memristive devices can be described by using a sigmoid function illustrated in equation 6.2 [186].

$$f_N = \left[ \frac{1}{1 + e^{-d(V-V_0)}} \right] \quad 6.2$$

Where,  $f_N$  describes the distribution function of 1T-1R memristive devices for a set of N voltage pulses at a voltage pulse amplitude V,  $V_0$  is the threshold voltage determined under the condition that  $f_N = 0.5$ , and d is the constant determined from the slope of the reset and set plots shown in Figure 6.8. Thus, the experimental data points of switching probability of the memristive devices at different voltage pulse amplitudes and pulse numbers can be fitted using the sigmoid function described in equation 6.2. The continuous lines in Figure 6.8 represent the sigmoid fits, and the switching probability of the memristive devices increases with the voltage amplitude of the applied pulse and the pulse numbers for both reset and set operations [186]. The slope parameter d is different for both reset and set transitions. As shown in Figure 6.8, the set transitions are steeper compared to the reset transitions which gives rise to a larger slope values of the sigmoid function. Further, the threshold voltage  $V_0$  decrease with an increase in pulse numbers for both reset and set transitions as shown in Figure 6.8 (a) and (b) [181].

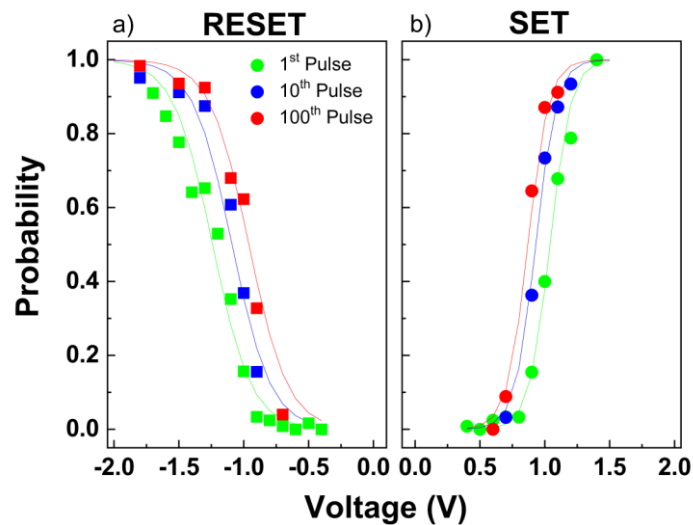


Figure 6.8: Switching probability plots for the reset (a) and set (b) operations as function of voltage amplitudes and number of pulses [181].

The synaptic efficacy in biological neural networks is defined by the level of neuronal activity [187]. In other words, it defines the presynaptic input strength to impact the postsynaptic output [182]. This mechanism could be emulated in microelectronic hardware through the probability of the resistance change in memristive devices [178]. The activation function for the synaptic efficacy is defined in equation 2.1 which is evaluated with the knowledge gained from D2D and pulse number dependent variability of the memristive devices. The neural networks reported in the literature implemented the discrete resistance states of binary memristive devices into the learning algorithms as a standard practice [47], [177]. On the contrary, the switching probability of the memristive devices can be implemented into the activation function to process the synaptic information [181]. Hence, the binary HfO<sub>2</sub> based 1T-1R memristive devices can be utilized to emulate the performance of analog synapses. The discussed D2D variability dependent on the pulse number in the HfO<sub>2</sub>-based polycrystalline memristive devices can be used for low-power neuromorphic computing applications such as pattern recognition [175], [178].

## 6.4 Evaluation of reliability in memristive devices

The reliability of the HfO<sub>2</sub>-based 1T-1R memristive devices is assessed through endurance and retention measurements. The endurance measurements evaluate the ability to write and erase the data in the memristive device multiple times by cycling

the devices. The retention measurements evaluate the ability to retrieve the data stored in a memristive device after an adequate period of time.

Temperature acceleration tests are performed on the memristive devices to evaluate the data retention just after the first reset and set cycle (not shown). The CDF of the HRS and LRS currents just after the first endurance cycle and their corresponding retention measurements at time intervals of 1, 10, and 100 hours are illustrated in Figure 6.9 and Figure 6.10, respectively. The impact of the temperature stress on both HRS and LRS currents is evident immediately after the first endurance cycle of the memristive devices [181]. In the case of HRS, the temperature stress results in the drift of its currents towards larger values, as shown in Figure 6.9. The distribution of the LRS currents increases as a result of the temperature stress, as shown in Figure 6.10.

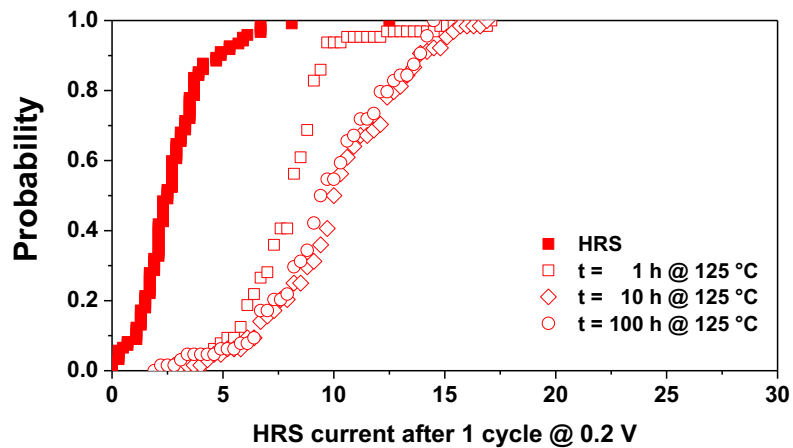


Figure 6.9: HRS cumulative distributions after 1 set/reset cycle at different retention times [181].

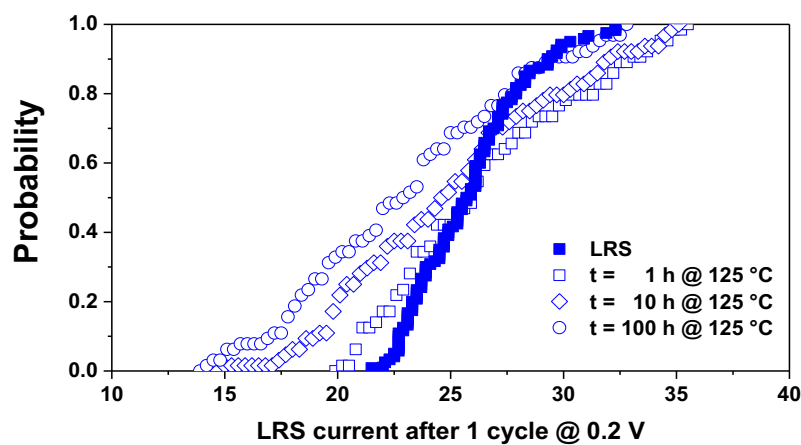


Figure 6.10: LRS cumulative distributions after 1 set/reset cycle at different retention times [181].

Further, the summary of the whole retention test performed on the four sets of memristive devices after the endurance test is illustrated in Figure 6.11 and Figure 6.12. The mean values of the HRS currents and their dispersion coefficients mainly increase with device cycling as shown in Figure 6.11. Irrespective of the cycling endurance, the HRS currents increase slightly with the temperature stress in the beginning but the current levels remain mostly constant after 10 hours of baking time [181].

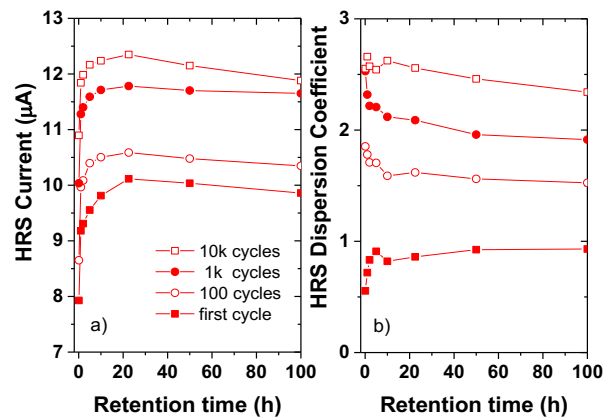
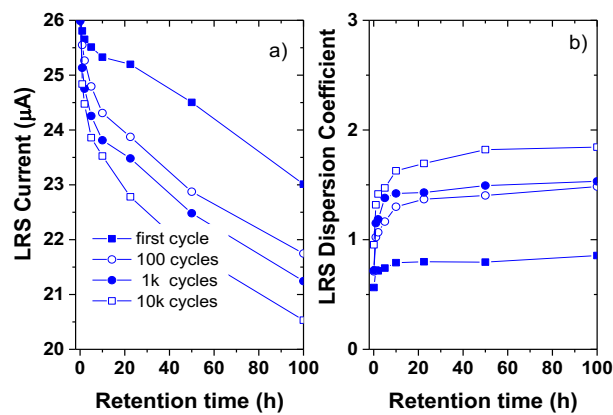


Figure 6.11: HRS mean (a) and dispersion (b) read currents as a function of retention time and number of endurance cycles [181].

In contrast to HRS current values, the LRS current values are affected by both cycling endurance and temperature stress. As shown in Figure 6.12 (a), the mean values of the LRS currents are reduced by 10% after the first programming cycle and 100 hours of baking time. Meanwhile, the dispersion coefficients of the LRS currents increase with the number of endurance cycles, but they stabilize with the baking time as shown in Figure 6.12 (b) [181]. The degradation observed in the LRS and HRS currents due to the cycling endurance can be related to the carbon impurities in the metal-organic precursor used for the HfO<sub>2</sub> deposition in the AVD process [95], [188]. Further, the degradation of the LRS currents observed during the retention measurements is due to the diffusion of the oxygen vacancies being activated under the influence of external temperature [189].





**Figure 6.12: LRS mean (a) and dispersion (b) read currents as a function of retention time and number of endurance cycles [181].**

The degradation in the cycling endurance and retention behavior of the memristive devices could impact the pattern recognition rate in neuromorphic computing systems [176]. In case of the stochastic neural networks such as the one described in Zahari et al., the learning takes place in a few iteration steps. Hence, the endurance and retention factors are considered less significant [175]. However, once the training is completed, the retention degradation will affect the performance of the network.

## 6.5 Conclusion

The functionality of a synapse in a biological neural network, and using a memristive device to emulate its performance in an artificial neural network is briefly discussed. The advantages of utilizing the memristive devices as artificial synapses, fabricated in the BEOL of a CMOS baseline technology are explained. Further, the HfO<sub>2</sub>-based binary memristive devices utilized for emulating the analog synaptic performance are presented. The variability in the memristive devices is targeted. Accordingly, the memristive switching layers are grown in a polycrystalline phase. The switching variabilities of the memristive devices in LRS and HRS are demonstrated as a function of the voltage pulse amplitudes and pulse numbers. The switching probabilities of the memristive devices are extracted from the previous measurements and they are plotted against the voltage pulse amplitudes and pulse numbers. The parameters extracted from this plot are plugged into a synaptic activation function to obtain the necessary conversion required for the synaptic information processing. Finally, a set of endurance and retention measurements demonstrate the reliability of the memristive devices. The cycling endurance resulted

in the degradation of both LRS and HRS. Whereas, degradation due to retention is mainly observed in the LRS of the memristive devices. These factors are given less significance while training the ANNs such as stochastic neural networks where the memristive devices can be successfully used as artificial synapses.

## Chapter VII

---

### 7 Conclusion and Future Outlook

---

The work undertaken in this thesis presents the process development and electrical characterization approaches to address the main challenges associated with memristive devices. Throughout the thesis, various aspects of the memristive devices such as reliability, performance, forming voltages, and mitigating and harnessing the variabilities were discussed. The solutions were implemented at different levels starting from the fabrication of memristive devices, their integration into the BiCMOS platform, device-engineering methods, and various pulse programming schemes to electrically characterize them.

The process development approach in Chapter 3 presented a stable and reliable technique to fabricate memristive devices with improved performance. Accordingly, the devices were fabricated in three different approaches. The reliability of the devices was assessed, based on the pristine state current measurements performed on the PrCM structures. Type 3 devices fabricated with spacer and encapsulation layers exhibited lower value pristine-state currents in comparison with the other two approaches. These low initial currents directly impacted the forming process of the memristive devices, resulting in a reliable forming operation with reduced forming voltage variabilities. Further, improved memory windows of the devices and reduced variability in resistive switching parameters were observed for the type 3 fabricated

devices. To further improve the performance of the memristive devices and reduce the parasitic RC, they were integrated into the 130 nm BiCMOS technology of IHP. The integration had a negligible impact on the FEOL and BEOL baseline processes. It allows the memristive device-based layouts to be fabricate in multi-project-wafer processes with minimum impact on the other layouts. Hence, it is cost-effective and avoids dedicated wafer requirement for memristive device-based layouts.

Further, this thesis addressed the forming voltage issues through electrical characterization methods. The forming current and voltage values, as well as their distributions were reduced by performing the forming operation at high temperatures by using the ISPVA programming scheme. The formation of the CF was controlled through small incremental programming pulses in ISPVA along with current read-out verification steps. Thus, enabling the programming of memristive devices for multi-level conduction.

Additionally, a device engineering method was presented to reduce the variability in memristive devices. Thin  $\text{Al}_2\text{O}_3$  layers of variable thickness are added to the  $\text{HfO}_2$ -based memristive devices, and the conduction properties of the devices were compared. The CF properties in the bi-layer memristive device stack, which is reported seldomly in literature, was investigated and presented. The experimental I-V curves of the LRS and HRS of the memristive devices were fitted using the QPC model. The addition of  $\text{Al}_2\text{O}_3$  layers resulted in the formation of a strong CF in the LRS and localized constriction of the CF in the HRS. Further, the variability in the on and off-state currents was reduced by adding  $\text{Al}_2\text{O}_3$  layers. Thus, enabling the devices for multi-level programming.

Although mitigating the variability was aimed throughout this thesis, the final chapter focuses on harnessing it for neuromorphic computing applications. The memristive devices are intentionally grown in a polycrystalline phase to target variability. The goal was to obtain the analog synaptic functionality of artificial synapses in memristive devices. They were programmed using train of programming pulses, and the switching variabilities of the devices were shown as a function of voltage pulse amplitudes and pulse numbers. The stochasticity involved in the memristive device switching was extracted in the form of switching probabilities. It can be implemented into the sigmoidal activation function for synaptic information processing. Further, the reliability of the memristive devices was demonstrated through a set of endurance and retention measurements. The degradations were observed in LRS and HRS of the memristive devices during reliability measurements. However, the devices find their application in training of stochastic neural networks where these characteristics are considered less significant.

The different techniques used in this work to improve the device performance can be combined and applied on the memristive devices to enhance the performance further. Additionally, as a part of the future outlook of this work, some of the improvements are proposed below.

The memristive devices integrated into the Bi-CMOS platform have the potential to enable wide range of applications. Various improvements are possible from the technology standpoint. The BiCMOS platform at IHP uses the chemical vapor deposited (CVD) tungsten for the vias in the BEOL interconnects. This tungsten could be used as the bottom electrode for the memristive devices. The high thermal conductivity of tungsten assists in further improving the performance of memristive devices. The MEMRES module integration steps on the tungsten vias need to be developed. The roughness of the tungsten vias after the chemical mechanical polishing poses a hurdle and has to be mitigated through novel integration approaches. Additionally, aluminum metal layers in the BEOL can be used as the bottom electrode for memristive devices, to explore the multi-level and analog resistive switching properties. However, an integration scheme needs to be developed here as well. Further, the type 3 MEMRES module developed under this work, can be integrated between different metal levels in the BEOL interconnects. The effect of the thermal budget and other process-related parameters on memristive devices, can be evaluated through performance and reliability measurements.

Further, the  $\text{Al}_2\text{O}_3$  layers used in the bi-layer memristive device stack of this work are deposited by using  $\text{H}_2\text{O}$  as the oxygen precursor. Changing the oxygen precursor to ozone results in high-density  $\text{Al}_2\text{O}_3$  layers. They can be used in bi-layer memristive device stacks for improved performance. Further, if aluminum is used as the bottom electrode, it can be oxidized to obtain thin  $\text{Al}_2\text{O}_3$  layers. A comparative study can be made on the conduction filament properties of the memristive devices consisting of different qualities of  $\text{Al}_2\text{O}_3$  layers in a bi-layer stack. Further, the device engineering techniques can be combined with pulse-programming schemes, to mitigate the variability issues in memristive devices further.

---

## 8 Bibliography

---

- [1] Intel Corporation, "Intel's First Product – the 3101," Nov. 28, 2022.
- [2] M. Marinella, "The future of memory," in *2013 IEEE Aerospace Conference*, 2013, pp. 1–11.
- [3] D. S. Jeong *et al.*, "Emerging memories: resistive switching mechanisms and current status," *Reports on progress in physics*, vol. 75, no. 7, p. 76502, 2012.
- [4] R. Bez, E. Camerlenghi, A. Modelli, and A. Visconti, "Introduction to flash memory," *Proceedings of the IEEE*, vol. 91, no. 4, pp. 489–502, 2003.
- [5] J. S. Meena, S. M. Sze, U. Chand, and T.-Y. Tseng, "Overview of emerging nonvolatile memory technologies," *Nanoscale Res Lett*, vol. 9, no. 1, pp. 1–33, 2014.
- [6] Yole Development, "Emerging Non-Volatile Memory (NVM) Technologies & Markets 2017," Grenoble, Jun. 2017.
- [7] O. Krestinskaya, A. P. James, and L. O. Chua, "Neuromemristive circuits for edge computing: A review," *IEEE Trans Neural Netw Learn Syst*, vol. 31, no. 1, pp. 4–23, 2019.
- [8] ITRS, "International Technology Roadmap for Semiconductors 2011," 2011.

- 
- [9] D. Ielmini, "Resistive switching memories based on metal oxides: mechanisms, reliability and scaling," *Semicond Sci Technol*, vol. 31, no. 6, p. 63002, 2016.
- [10] W. Banerjee, "Challenges and applications of emerging nonvolatile memory devices," *Electronics (Basel)*, vol. 9, no. 6, p. 1029, 2020.
- [11] T. Mikolajick *et al.*, "Doped Hafnium oxide—an enabler for ferroelectric field effect transistors," *Advances in Science and Technology*, vol. 95, pp. 136–145, 2014.
- [12] D. J. Wouters, R. Waser, and M. Wuttig, "Phase-change and redox-based resistive switching memories," *Proceedings of the IEEE*, vol. 103, no. 8, pp. 1274–1288, 2015.
- [13] Y. Huai and others, "Spin-transfer torque MRAM (STT-MRAM): Challenges and prospects," *AAPPS bulletin*, vol. 18, no. 6, pp. 33–40, 2008.
- [14] H. S. P. Wong *et al.*, "Metal-oxide RRAM," in *Proceedings of the IEEE*, 2012, vol. 100, no. 6, pp. 1951–1970. doi: 10.1109/JPROC.2012.2190369.
- [15] T. Rueckes, K. Kim, E. Joselevich, G. Y. Tseng, C.-L. Cheung, and C. M. Lieber, "Carbon nanotube-based nonvolatile random-access memory for molecular computing," *Science (1979)*, vol. 289, no. 5476, pp. 94–97, 2000.
- [16] Y. Chen *et al.*, "Nanoscale molecular-switch crossbar circuits," *Nanotechnology*, vol. 14, no. 4, p. 462, 2003.
- [17] J. Fu *et al.*, "Si-nanowire based gate-all-around nonvolatile SONOS memory cell," *IEEE Electron Device Letters*, vol. 29, no. 5, pp. 518–521, 2008.
- [18] B. Mohammad *et al.*, "State-of-the-art of metal-oxide memristor devices," *Nanotechnol Rev*, vol. 5, no. 3, pp. 311–329, 2016.
- [19] A. Mehonic, A. Sebastian, B. Rajendran, O. Simeone, E. Vasilaki, and A. J. Kenyon, "Memristors—From in-memory computing, deep learning acceleration, and spiking neural networks to the future of neuromorphic and bio-inspired computing," *Advanced Intelligent Systems*, vol. 2, no. 11, p. 2000085, 2020.
- [20] A. Sebastian, M. le Gallo, G. W. Burr, S. Kim, M. BrightSky, and E. Eleftheriou, "Tutorial: Brain-inspired computing using phase-change memory devices," *J Appl Phys*, vol. 124, no. 11, p. 111101, 2018.
- [21] A. D. Kent and D. C. Worledge, "A new spin on magnetic memories," *Nat Nanotechnol*, vol. 10, no. 3, pp. 187–191, 2015.

- 
- [22] F. Pellizzer and A. Redaelli, "3D-Xpoint fundamentals," in *Semiconductor Memories and Systems*, Elsevier, 2022, pp. 253–276.
- [23] EVERSPIN TECHNOLOGIES, "Everspin 64Mb DDR3 Spin-Torque MRAM," 2016.
- [24] J. R. Jameson *et al.*, "Conductive-bridge memory (CBRAM) with excellent high-temperature retention," in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 30.1.1-30.1.4. doi: 10.1109/IEDM.2013.6724721.
- [25] D. Chen *et al.*, "Single-Event Effect Performance of a Commercial Embedded ReRAM," *IEEE Trans Nucl Sci*, vol. 61, no. 6, pp. 3088–3094, 2014, doi: 10.1109/TNS.2014.2361488.
- [26] S. Bertolazzi and E. Jolivet, "Emerging NonVolatile Memory 2020," 2020. Accessed: Nov. 30, 2022. [Online]. Available: [https://s3.i-micronews.com/uploads/2020/02/Yole\\_YDR20066\\_Emerging-Non-Volatile-Memory-2020\\_sample.pdf](https://s3.i-micronews.com/uploads/2020/02/Yole_YDR20066_Emerging-Non-Volatile-Memory-2020_sample.pdf)
- [27] F. Clermidy *et al.*, "Resistive memories: Which applications?," in *2014 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2014, pp. 1–6.
- [28] S. Senni *et al.*, "Embedded systems to high performance computing using STT-MRAM," in *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2017, 2017, pp. 536–541.
- [29] G. Molas and E. Nowak, "Advances in Emerging Memory Technologies: From Data Storage to Artificial Intelligence," *Applied Sciences*, vol. 11, no. 23, p. 11254, 2021.
- [30] S. Bertolazzi and E. Jolivet, "EMERGING NON-VOLATILE MEMORY 2021," 2021. Accessed: Dec. 13, 2022. [Online]. Available: <https://s3.i-micronews.com/uploads/2021/02/YINTR21218-emerging-Non-Volatil-Memory-2021-flyer-web.pdf>
- [31] A. K. Kamath, L. Monis, A. T. Karthik, and B. Talawar, "Storage class memory: principles, problems, and possibilities," *arXiv preprint arXiv:1909.12221*, 2019.
- [32] B. Li, B. Yan, and H. Li, "An overview of in-memory processing with emerging non-volatile memory for data-intensive applications," in *Proceedings of the 2019 on Great Lakes Symposium on VLSI*, 2019, pp. 381–386.



- 
- [33] R. Waser, *Nanoelectronics and information technology: advanced electronic materials and novel devices*. John Wiley & Sons, 2012.
- [34] R. Waser, R. Dittmann, G. Staikov, and K. Szot, “Redox-based resistive switching memories—nanoionic mechanisms, prospects, and challenges,” *Advanced materials*, vol. 21, no. 25–26, pp. 2632–2663, 2009.
- [35] S. Yu, “Resistive random-access memory (RRAM),” *Synthesis lectures on emerging engineering technologies*, vol. 2, no. 5, pp. 1–79, 2016.
- [36] D. Ielmini, R. Bruchhaus, and R. Waser, “Thermochemical resistive switching: materials, mechanisms, and scaling projections,” *Phase Transitions*, vol. 84, no. 7, pp. 570–602, 2011.
- [37] R. Waser, “Electrochemical and thermochemical memories,” in *2008 IEEE International Electron Devices Meeting*, 2008, pp. 1–4.
- [38] D. Kumar, R. Aluguri, U. Chand, and T.-Y. Tseng, “Metal-oxide resistive switching memory: materials, properties and switching mechanisms,” *Ceram Int*, vol. 43, pp. S547–S556, 2017.
- [39] L. Goux and I. Valov, “Electrochemical processes and device improvement in conductive bridge RAM cells,” *physica status solidi (a)*, vol. 213, no. 2, pp. 274–288, 2016.
- [40] H. Abbas, J. Li, and D. S. Ang, “Conductive Bridge Random Access Memory (CBRAM): Challenges and Opportunities for Memory and Neuromorphic Computing Applications,” *Micromachines (Basel)*, vol. 13, no. 5, p. 725, 2022.
- [41] I. Valov, R. Waser, J. R. Jameson, and M. N. Kozicki, “Electrochemical metallization memories—fundamentals, applications, prospects,” *Nanotechnology*, vol. 22, no. 25, p. 254003, 2011.
- [42] H. Kim, H.-B.-R. Lee, and W.-J. Maeng, “Applications of atomic layer deposition to nanofabrication and emerging nanodevices,” *Thin Solid Films*, vol. 517, no. 8, pp. 2563–2580, 2009, doi: <https://doi.org/10.1016/j.tsf.2008.09.007>.
- [43] T.-Y. Wang *et al.*, “Forming-free flexible memristor with multilevel storage for neuromorphic computing by full PVD technique,” *J Mater Sci Technol*, vol. 60, pp. 21–26, 2021.
- [44] S. Bagdzevicius, K. Maas, M. Boudard, and M. Burriel, “Interface-type resistive switching in perovskite materials,” *J Electroceram*, vol. 39, no. 1, pp. 157–184, 2017.

- 
- [45] J. Woo and S. Yu, "Resistive memory-based analog synapse: The pursuit for linear and symmetric weight update," *IEEE Nanotechnol Mag*, vol. 12, no. 3, pp. 36–44, 2018.
- [46] B. Gao *et al.*, "Modeling disorder effect of the oxygen vacancy distribution in filamentary analog RRAM for neuromorphic computing," in *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 4.4.1-4.4.4. doi: 10.1109/IEDM.2017.8268326.
- [47] V. Milo *et al.*, "Multilevel HfO<sub>2</sub>-based RRAM devices for low-power neuromorphic networks," *APL Mater*, vol. 7, no. 8, p. 81120, 2019.
- [48] E. Pérez, A. Grossi, C. Zambelli, P. Olivo, R. Roelofs, and C. Wenger, "Reduction of the cell-to-cell variability in Hf<sub>1-x</sub>Al<sub>x</sub>O<sub>y</sub> based RRAM arrays by using program algorithms," *IEEE Electron Device Letters*, vol. 38, no. 2, pp. 175–178, 2016.
- [49] M. Azzaz *et al.*, "Benefit of Al<sub>2</sub>O<sub>3</sub>/HfO<sub>2</sub> bilayer for BEOL RRAM integration through 16 kb memory cut characterization," in *European Solid-State Device Research Conference*, Nov. 2015, vol. 2015-November, pp. 266–269. doi: 10.1109/ESSDERC.2015.7324765.
- [50] X. Guan, S. Yu, and H.-S. P. Wong, "On the switching parameter variation of metal-oxide RRAM—Part I: Physical modeling and simulation methodology," *IEEE Trans Electron Devices*, vol. 59, no. 4, pp. 1172–1182, 2012.
- [51] A. Padovani, L. Larcher, O. Pirrotta, L. Vandelli, and G. Bersuker, "Microscopic Modeling of HfO<sub>x</sub> RRAM Operations: From Forming to Switching," *IEEE Trans Electron Devices*, vol. 62, no. 6, pp. 1998–2006, 2015, doi: 10.1109/TED.2015.2418114.
- [52] S. Aldana *et al.*, "Resistive switching in HfO<sub>2</sub> based valence change memories, a comprehensive 3D kinetic Monte Carlo approach," *J Phys D Appl Phys*, vol. 53, no. 22, p. 225106, 2020.
- [53] S. Dirkmann, J. Kaiser, C. Wenger, and T. Mussenbrock, "Filament growth and resistive switching in hafnium oxide memristive devices," *ACS Appl Mater Interfaces*, vol. 10, no. 17, pp. 14857–14868, 2018.
- [54] S. Yu, X. Guan, and H.-S. P. Wong, "Conduction mechanism of TiN/HfO<sub>x</sub>/Pt resistive switching memory: A trap-assisted-tunneling model," *Appl Phys Lett*, vol. 99, no. 6, p. 63507, 2011.

- 
- [55] C.-Y. Lin *et al.*, "Attaining resistive switching characteristics and selector properties by varying forming polarities in a single HfO<sub>2</sub>-based RRAM device with a vanadium electrode," *Nanoscale*, vol. 9, no. 25, pp. 8586–8590, 2017.
- [56] K.-M. Chang, W.-H. Tzeng, K.-C. Liu, and W.-R. Lai, "Modulation of the electrical characteristics of HfO<sub>x</sub>/TiN RRAM devices through the top electrode metal," *ECS Trans*, vol. 28, no. 2, p. 119, 2010.
- [57] X. Lian *et al.*, "Multi-scale quantum point contact model for filamentary conduction in resistive random-access memories devices," *J Appl Phys*, vol. 115, no. 24, Jun. 2014, doi: 10.1063/1.4885419.
- [58] Z. Wang *et al.*, "Engineering incremental resistive switching in TaO<sub>x</sub> based memristors for brain-inspired computing," *Nanoscale*, vol. 8, no. 29, pp. 14015–14022, 2016.
- [59] W. Wu, H. Wu, B. Gao, N. Deng, S. Yu, and H. Qian, "Improving analog switching in HfO<sub>x</sub>-based resistive memory with a thermal enhanced layer," *IEEE Electron Device Letters*, vol. 38, no. 8, pp. 1019–1022, 2017.
- [60] S. Mandal, A. El-Amin, K. Alexander, B. Rajendran, and R. Jha, "Novel synaptic memory device for neuromorphic computing," *Sci Rep*, vol. 4, no. 1, pp. 1–10, 2014.
- [61] Y. Y. Liauw, Z. Zhang, W. Kim, A. el Gamal, and S. S. Wong, "Nonvolatile 3D-FPGA with monolithically stacked RRAM-based configuration memory," in *2012 IEEE International Solid-State Circuits Conference*, 2012, pp. 406–408.
- [62] E. Vianello *et al.*, "Resistive memories for ultra-low-power embedded computing design," in *IEEE International Electron Devices Meeting (IEDM)*, 2014, vol. 6, pp. 1–6.
- [63] M. Chang *et al.*, "A 40 nm 60.64 TOPS/W ECC-Capable Compute-in-Memory/Digital 2.25 MB/768 kB RRAM/SRAM System with Embedded Cortex M3 Microprocessor for Edge Recommendation Systems," in *2022 IEEE International Solid-State Circuits Conference (ISSCC)*, 2022, vol. 65, pp. 1–3.
- [64] D. L. Lewis and H.-H. S. Lee, "Architectural evaluation of 3D stacked RRAM caches," in *2009 IEEE International Conference on 3D System Integration*, 2009, pp. 1–4.

- 
- [65] S. B. Furber, "Brain-inspired computing," *IET Comput Digit Tech*, vol. 10, no. 6, pp. 299–305, 2016.
- [66] S. Yin *et al.*, "Monolithically integrated RRAM-and CMOS-based in-memory computing optimizations for efficient deep learning," *IEEE Micro*, vol. 39, no. 6, pp. 54–63, 2019.
- [67] V. Milo *et al.*, "Optimized programming algorithms for multilevel RRAM in hardware neural networks," in *2021 IEEE International Reliability Physics Symposium (IRPS)*, 2021, pp. 1–6.
- [68] J. Cai, X. Han, X. Wang, and X. Meng, "Atomic Layer Deposition of Two-Dimensional Layered Materials: Processes, Growth Mechanisms, and Characteristics," *Matter*, vol. 2, no. 3, pp. 587–630, 2020, doi: <https://doi.org/10.1016/j.matt.2019.12.026>.
- [69] H. B. Profijt, S. E. Potts, M. C. M. de Sanden, and W. M. M. Kessels, "Plasma-assisted atomic layer deposition: basics, opportunities, and challenges," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 29, no. 5, p. 50801, 2011.
- [70] H. Kim and I.-K. Oh, "Review of plasma-enhanced atomic layer deposition: Technical enabler of nanoscale device fabrication," *Jpn J Appl Phys*, vol. 53, no. 3S2, p. 03DA01, 2014.
- [71] H. C. M. Knoop, S. E. Potts, A. A. Bol, and W. M. M. Kessels, "Atomic Layer Deposition. Handbook of Crystal Growth." Elsevier, 2015.
- [72] R. D. Clark, "Emerging Applications for High K Materials in VLSI Technology," *Materials*, vol. 7, no. 4, pp. 2913–2944, 2014, doi: 10.3390/ma7042913.
- [73] C. Walczyk *et al.*, "Impact of Temperature on the Resistive Switching Behavior of Embedded HfO<sub>2</sub>-Based RRAM Devices," *IEEE Trans Electron Devices*, vol. 58, no. 9, pp. 3124–3131, 2011, doi: 10.1109/TED.2011.2160265.
- [74] S. M. Rosnagel, R. Wisnieff, D. Edelstein, and T. S. Kuan, "Interconnect issues post 45nm," in *IEEE International Electron Devices Meeting, 2005. IEDM Technical Digest.*, 2005, pp. 89–91.
- [75] M. Häyrynen *et al.*, "Low-loss titanium dioxide strip waveguides fabricated by atomic layer deposition," *Journal of Lightwave Technology*, vol. 32, no. 2, pp. 208–212, 2013.

- 
- [76] J. Liu and X. Sun, "Elegant design of electrode and electrode/electrolyte interface in lithium-ion batteries by atomic layer deposition," *Nanotechnology*, vol. 26, no. 2, p. 24001, 2014.
- [77] O. Tiurin and Y. Ein-Eli, "A critical review: The impact of the battery electrode material substrate on the composition and properties of atomic layer deposition (ALD) coatings," *Adv Mater Interfaces*, vol. 6, no. 24, p. 1901455, 2019.
- [78] B. Vermang *et al.*, "Blistering in ALD Al<sub>2</sub>O<sub>3</sub> passivation layers as rear contacting for local Al BSF Si solar cells," *Solar energy materials and solar cells*, vol. 101, pp. 204–209, 2012.
- [79] L. Wang *et al.*, "Enhanced moisture barrier performance for ALD-encapsulated OLEDs by introducing an organic protective layer," *J Mater Chem C Mater*, vol. 5, no. 16, pp. 4017–4024, 2017.
- [80] M. Staszuk *et al.*, "Structure and properties of ZnO coatings obtained by atomic layer deposition (ALD) method on a Cr-Ni-Mo steel substrate type," *Materials*, vol. 13, no. 19, p. 4223, 2020.
- [81] J. Jeong *et al.*, "Conformal hermetic sealing of wireless microelectronic implantable chiplets by multilayered atomic layer deposition (ALD)," *Adv Funct Mater*, vol. 29, no. 5, p. 1806440, 2019.
- [82] H. C. M. Knoop *et al.*, "Deposition of TiN and TaN by remote plasma ALD for Cu and Li diffusion barrier applications," *J Electrochem Soc*, vol. 155, no. 12, p. G287, 2008.
- [83] R. W. Johnson, A. Hultqvist, and S. F. Bent, "A brief review of atomic layer deposition: from fundamentals to applications," *Materials Today*, vol. 17, no. 5, pp. 236–246, 2014, doi: <https://doi.org/10.1016/j.mattod.2014.04.026>.
- [84] Erwin Kessels, "ALD from an application perspective," Jan. 19, 2022. <https://www.atomiclimits.com/2022/01/19/ald-from-an-application-perspective-an-upcoming-avs-webinar/> (accessed May 11, 2022).
- [85] G. Niu *et al.*, "Material insights of HfO<sub>2</sub>-based integrated 1-transistor-1-resistor resistive random-access memory devices processed by batch atomic layer deposition," *Sci Rep*, vol. 6, no. 1, pp. 1–11, 2016.
- [86] M. Si, Z. Lin, Z. Chen, and P. D. Ye, "High-Performance Atomic-Layer-Deposited Indium Oxide 3-D Transistors and Integrated Circuits for Monolithic 3-D

- 
- Integration," *IEEE Trans Electron Devices*, vol. 68, no. 12, pp. 6605–6609, 2021, doi: 10.1109/TED.2021.3106282.
- [87] E. Pérez, C. Zambelli, M. K. Mahadevaiah, P. Olivo, and C. Wenger, "Toward Reliable Multi-Level Operation in RRAM Arrays: Improving Post-Algorithm Stability and Assessing Endurance/Data Retention," *IEEE Journal of the Electron Devices Society*, vol. 7, pp. 740–747, 2019, doi: 10.1109/JEDS.2019.2931769.
- [88] R. Han *et al.*, "Demonstration of logic operations in high-performance RRAM crossbar array fabricated by atomic layer deposition technique," *Nanoscale Res Lett*, vol. 12, no. 1, pp. 1–6, 2017.
- [89] P. Lin and Q. Xia, "Tutorial: Fabrication and three-dimensional integration of nanoscale memristive devices and arrays," *J Appl Phys*, vol. 124, no. 15, p. 152001, 2018.
- [90] M. Morita, T. Ohmi, E. Hasegawa, M. Kawakami, and M. Ohwada, "Growth of native oxide on a silicon surface," *J Appl Phys*, vol. 68, no. 3, pp. 1272–1281, 1990.
- [91] S. E. Potts, H. B. Profijt, R. Roelofs, and W. M. M. Kessels, "Room-Temperature ALD of Metal Oxide Thin Films by Energy-Enhanced ALD," *Chemical Vapor Deposition*, vol. 19, no. 4–6, pp. 125–133, 2013.
- [92] R. L. Puurunen, "Surface chemistry of atomic layer deposition: A case study for the trimethylaluminum/water process," *J Appl Phys*, vol. 97, no. 12, p. 9, 2005.
- [93] A. Fantini *et al.*, "Engineering of  $\text{Hf}_{1-x}\text{Al}_x\text{O}_y$  amorphous dielectrics for high-performance RRAM applications," in *2014 IEEE 6th International Memory Workshop, IMW 2014*, 2014. doi: 10.1109/IMW.2014.6849354.
- [94] K. Instruments, "Model 4200-SCS Semiconductor Characterization System Reference Manual."
- [95] A. Grossi *et al.*, "Electrical characterization and modeling of pulse-based forming techniques in RRAM arrays," *Solid State Electron*, vol. 115, pp. 17–25, 2016.
- [96] A. Grossi *et al.*, "Impact of the precursor chemistry and process conditions on the cell-to-cell variability in 1T-1R based  $\text{HfO}_2$  RRAM devices," *Sci Rep*, vol. 8, no. 1, pp. 1–11, 2018.
- [97] R. Degraeve *et al.*, "Causes and consequences of the stochastic aspect of filamentary RRAM," *Microelectron Eng*, vol. 147, pp. 171–175, 2015.

- 
- [98] M. K. Mahadevaiah *et al.*, "(Invited) Optimized HfO<sub>2</sub>-Based MIM Module Fabrication for Emerging Memory Applications," *ECS Trans*, vol. 92, no. 4, pp. 211–221, Jul. 2019, doi: 10.1149/09204.0211ecst.
- [99] Z. Yong *et al.*, "Tuning oxygen vacancies and resistive switching properties in ultra-thin HfO<sub>2</sub> RRAM via TiN bottom electrode and interface engineering," *Appl Surf Sci*, vol. 551, p. 149386, 2021.
- [100] Y. Cui, J. Y. Jeong, Y. Gao, and S. G. Pyo, "Process optimization of via plug multilevel interconnections in CMOS logic devices," *Micromachines (Basel)*, vol. 11, no. 1, p. 32, 2019.
- [101] K. R. Williams, K. Gupta, and M. Wasilik, "Etch rates for micromachining processing-Part II," *Journal of microelectromechanical systems*, vol. 12, no. 6, pp. 761–778, 2003.
- [102] D. Shamiryan, M. R. Baklanov, S. Vanhaelemeersch, and K. Maex, "Comparative study of SiOCH low-k films with varied porosity interacting with etching and cleaning plasma," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 20, no. 5, pp. 1923–1928, 2002.
- [103] V. M. Donnelly and A. Kornblit, "Plasma etching: Yesterday, today, and tomorrow," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 31, no. 5, p. 50825, 2013.
- [104] N. Selamoglu, C. N. Bredbenner, T. A. Gliniecki, and H. J. Stocker, "Tapered etching of aluminum with CHF<sub>3</sub>/Cl<sub>2</sub>/BCl<sub>3</sub> and its impact on step coverage of plasma-deposited silicon oxide from tetraethoxysilane," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 9, no. 5, pp. 2530–2535, 1991.
- [105] S. A. Vitale, J. Kedzierski, and C. L. Keast, "High density plasma etching of titanium nitride metal gate electrodes for fully depleted silicon-on-insulator subthreshold transistor integration," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 27, no. 6, pp. 2472–2479, 2009.
- [106] M. E. McNie *et al.*, "Performance enhancement and evaluation of deep dry etching on a production cluster platform," in *Micromachining and Microfabrication Process Technology VIII*, 2003, vol. 4979, pp. 34–42.

- 
- [107] R. A. Gottscho, C. W. Jurgensen, and D. J. Vitkavage, "Microscopic uniformity in plasma etching," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, vol. 10, no. 5, pp. 2133–2147, 1992.
- [108] P. Huang *et al.*, "A physics-based compact model of metal-oxide-based RRAM DC and AC operations," *IEEE Trans Electron Devices*, vol. 60, no. 12, pp. 4090–4097, 2013.
- [109] H. Li, P. Huang, B. Gao, B. Chen, X. Liu, and J. Kang, "A SPICE model of resistive random-access memory for large-scale memory array simulation," *IEEE Electron Device Letters*, vol. 35, no. 2, pp. 211–213, 2013.
- [110] N. Raghavan *et al.*, "Statistical insight into controlled forming and forming free stacks for HfO<sub>x</sub> RRAM," *Microelectron Eng*, vol. 109, pp. 177–181, 2013, doi: <https://doi.org/10.1016/j.mee.2013.03.065>.
- [111] K. G. Young-Fisher *et al.*, "Leakage Current-Forming Voltage Relation and Oxygen Gettering in HfO<sub>x</sub> RRAM Devices," *IEEE Electron Device Letters*, vol. 34, no. 6, pp. 750–752, 2013, doi: 10.1109/LED.2013.2256101.
- [112] E. Perez, M. K. Mahadevaiah, C. Wenger, C. Zambelli, and P. Olivo, "The role of the bottom and top interfaces in the 1<sup>st</sup> reset operation in HfO<sub>2</sub> based RRAM devices," in *2018 Joint International EUROSOI Workshop and International Conference on Ultimate Integration on Silicon (EUROSOI-ULIS)*, 2018, pp. 1–4.
- [113] H. Y. Lee *et al.*, "Scalability issue in Ti/HfO bipolar resistive memory with 1T-1R configuration by resistance pinning effect during 1<sup>st</sup> RESET and its solution," in *2013 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA)*, 2013, pp. 1–2.
- [114] P. Pouyan, E. Amat, S. Hamdioui, and A. Rubio, "RRAM variability and its mitigation schemes," in *2016 26th international workshop on power and timing modeling, optimization and simulation (PATMOS)*, 2016, pp. 141–146.
- [115] S. Yu, X. Guan, and H.-S. P. Wong, "On the switching parameter variation of metal oxide RRAM—Part II: Model corroboration and device design strategy," *IEEE Trans Electron Devices*, vol. 59, no. 4, pp. 1183–1188, 2012.
- [116] H.-X. Zheng *et al.*, "Overcoming Limited Resistance in 1T-1R RRAM Caused by Pinch-Off Voltage During Reset Process," *IEEE Trans Electron Devices*, vol. 66, no. 11, pp. 4706–4709, 2019, doi: 10.1109/TED.2019.2937278.



- 
- [117] Y. Hao *et al.*, "Uniform, fast, and reliable CMOS compatible resistive switching memory," *Journal of Semiconductors*, vol. 43, no. 5, p. 54102, May 2022, doi: 10.1088/1674-4926/43/5/054102.
- [118] H. A. du Nguyen, J. Yu, L. Xie, M. Taouil, S. Hamdioui, and D. Fey, "Memristive devices for computing: Beyond CMOS and beyond von Neumann," in *2017 IFIP/IEEE International Conference on Very Large-Scale Integration (VLSI-SoC)*, 2017, pp. 1–10. doi: 10.1109/VLSI-SoC.2017.8203479.
- [119] G. Pedretti and D. Ielmini, "In-Memory Computing with Resistive Memory Circuits: Status and Outlook," *Electronics (Basel)*, vol. 10, no. 9, 2021, doi: 10.3390/electronics10091063.
- [120] U. Sobe, K.-H. Rooch, and D. Mörtl, "Simulation and analysis of analog circuit and PCM (process control monitor) test structures in circuit design," 2007.
- [121] A. Chowdhury, J. Kim, and R. M. Wallace, "Alumina as a Hydrogen Barrier Layer for FeRAM Devices," in *2007 Non-Volatile Memory Technology Symposium*, 2007, pp. 49–52. doi: 10.1109/NVMT.2007.4389944.
- [122] M. Kalishettyhalli Mahadevaiah *et al.*, "Integration of memristive devices into a 130 nm CMOS baseline technology" in *Bio-inspired Information Pathways - From Neuroscience to Neurotronics*, Springer Series on Bio- and Neurosystems, 2023.
- [123] S.-T. Wei, B. Gao, D. Wu, J.-S. Tang, H. Qian, and H.-Q. Wu, "Trends and challenges in the circuit and macro of RRAM-based computing-in-memory systems," *Chip*, vol. 1, no. 1, p. 100004, 2022, doi: <https://doi.org/10.1016/j.chip.2022.100004>.
- [124] C. Walczyk *et al.*, "On the role of Ti adlayers for resistive switching in HfO<sub>2</sub>-based metal-insulator-metal structures: Top versus bottom electrode integration," *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 29, no. 1, p. 01AD02, 2011.
- [125] E. Pérez *et al.*, "Advanced temperature dependent statistical analysis of forming voltage distributions for three different HfO<sub>2</sub>-based RRAM technologies," *Solid State Electron*, vol. 176, p. 107961, 2021.
- [126] Y. Y. Chen *et al.*, "Tailoring switching and endurance/retention reliability characteristics of HfO<sub>2</sub>/Hf RRAM with Ti, Al, Si dopants," in *2014 Symposium on VLSI Technology (VLSI-Technology): Digest of Technical Papers*, 2014, pp. 1–2.

- 
- [127] B. Chakrabarti and E. M. Vogel, "Effect of Ti doping and annealing on multi-level forming-free resistive random-access memories with atomic layer deposited HfTiO<sub>x</sub> nanolaminate," *Microelectron Eng*, vol. 109, pp. 193–196, 2013, doi: <https://doi.org/10.1016/j.mee.2013.03.092>.
- [128] B. Butcher *et al.*, "Hot forming to improve memory window and uniformity of low-power HfO<sub>x</sub>-based RRAMs," in *2012 4th IEEE International Memory Workshop*, 2012, pp. 1–4.
- [129] J. Ma *et al.*, "Identify the critical regions and switching/failure mechanisms in non-filamentary RRAM (a-VMCO) by RTN and CVS techniques for memory window improvement," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 21–24.
- [130] M. K. Mahadevaiah, E. Perez, and C. Wenger, "Influence of specific forming algorithms on the device-to-device variability of memristive Al-doped HfO<sub>2</sub> arrays," *Journal of Vacuum Science & Technology B*, vol. 38, no. 1, p. 013201, Jan. 2020, doi: 10.1116/1.5126936.
- [131] M. Kalishettyhalli Mahadevaiah *et al.*, "Modulating the Filamentary-Based Resistive Switching Properties of HfO<sub>2</sub> Memristive Devices by Adding Al<sub>2</sub>O<sub>3</sub> Layers," *Electronics (Basel)*, vol. 11, no. 10, p. 1540, 2022.
- [132] A. Grossi *et al.*, "An automated test equipment for characterization of emerging MRAM and RRAM arrays," *IEEE Trans Emerg Top Comput*, vol. 6, no. 2, pp. 269–277, 2016.
- [133] G. Molas *et al.*, "Resistive memories (RRAM) variability: Challenges and solutions," *ECS Trans*, vol. 86, no. 3, p. 35, 2018.
- [134] P. Lorenzi, R. Rao, and F. Irrera, "Forming Kinetics in HfO<sub>2</sub>-Based RRAM Cells," *IEEE Trans Electron Devices*, vol. 60, no. 1, pp. 438–443, 2012.
- [135] C. Zambelli *et al.*, "Statistical analysis of resistive switching characteristics in ReRAM test arrays," in *2014 International Conference on Microelectronic Test Structures (ICMTS)*, 2014, pp. 27–31.
- [136] E. Pérez, C. Wenger, A. Grossi, C. Zambelli, P. Olivo, and R. Roelofs, "Impact of temperature on conduction mechanisms and switching parameters in HfO<sub>2</sub>-based 1T-1R resistive random-access memories devices," *Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena*, vol. 35, no. 1, p. 01A103, 2017.

- 
- [137] E. Pérez, A. Grossi, C. Zambelli, M. K. Mahadevaiah, P. Olivo, and C. Wenger, "Temperature impact and programming algorithm for RRAM based memories," in *2018 IEEE MTT-S International Microwave Workshop Series on Advanced Materials and Processes for RF and THz Applications (IMWS-AMP)*, 2018, pp. 1–3.
- [138] M. Zhang *et al.*, "A physical model for the statistics of the set switching time of resistive RAM measured with the width-adjusting pulse operation method," *IEEE Electron Device Letters*, vol. 36, no. 12, pp. 1303–1306, 2015.
- [139] N. J. Western, I. Perez-Wurfl, S. R. Wenham, and S. P. Bremner, "Point-contacting by localized dielectric breakdown with breakdown fields described by the Weibull distribution," *IEEE Trans Electron Devices*, vol. 62, no. 6, pp. 1826–1830, 2015.
- [140] S. Long *et al.*, "A model for the set statistics of RRAM inspired in the percolation model of oxide breakdown," *IEEE electron device letters*, vol. 34, no. 8, pp. 999–1001, 2013.
- [141] E. Y. Wu, J. Suñé, and W. Lai, "On the Weibull shape factor of intrinsic breakdown of dielectric films and its accurate experimental determination. Part II: experimental results and the effects of stress conditions," *IEEE Trans Electron Devices*, vol. 49, no. 12, pp. 2141–2150, 2002.
- [142] A. Grossi *et al.*, "Fundamental variability limits of filament-based RRAM," in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, pp. 4–7.
- [143] S. Long, C. Cagli, D. Ielmini, M. Liu, and J. Sune, "Analysis and modeling of resistive switching statistics," *J Appl Phys*, vol. 111, no. 7, p. 74508, 2012.
- [144] H. Li *et al.*, "Variation-aware, reliability-emphasized design and optimization of RRAM using SPICE model," in *2015 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2015, pp. 1425–1430.
- [145] J. Li, T. Zhang, Q. Duan, L. Li, Y. Yang, and R. Huang, "Engineering resistive switching behavior in TaO<sub>x</sub> based memristive devices for non-von Neuman computing applications," in *2018 China Semiconductor Technology International Conference (CSTIC)*, 2018, pp. 1–3.
- [146] X. Huang *et al.*, "HfO<sub>2</sub>/Al<sub>2</sub>O<sub>3</sub> multilayer for RRAM arrays: a technique to improve tail-bit retention," *Nanotechnology*, vol. 27, no. 39, p. 395201, 2016.

- 
- [147] P. Bousoulas, I. Michelakaki, E. Skotadis, M. Tsigkourakos, and D. Tsoukalas, "Low-power forming free  $\text{TiO}_{2-x}/\text{HfO}_{2-y}/\text{TiO}_{2-x}$ -trilayer RRAM devices exhibiting synaptic property characteristics," *IEEE Trans Electron Devices*, vol. 64, no. 8, pp. 3151–3158, 2017.
- [148] S. Maikap and W. Banerjee, "In Quest of Nonfilamentary Switching: A Synergistic Approach of Dual Nanostructure Engineering to Improve the Variability and Reliability of Resistive Random-Access-Memory Devices," *Adv Electron Mater*, vol. 6, no. 6, p. 2000209, 2020.
- [149] J. Woo *et al.*, "Improved synaptic behavior under identical pulses using  $\text{AlO}_x/\text{HfO}_2$  bilayer RRAM array for neuromorphic systems," *IEEE Electron Device Letters*, vol. 37, no. 8, pp. 994–997, Aug. 2016, doi: 10.1109/LED.2016.2582859.
- [150] C. Y. Huang, J. H. Jieng, W. Y. Jang, C. H. Lin, and T. Y. Tseng, "Improved resistive switching characteristics by  $\text{Al}_2\text{O}_3$  layers inclusion in  $\text{HfO}_2$ -based RRAM devices," *ECS Solid State Letters*, vol. 2, no. 8, 2013, doi: 10.1149/2.006308ssl.
- [151] U. Chand, K. C. Huang, C. Y. Huang, and T. Y. Tseng, "Mechanism of Nonlinear Switching in  $\text{HfO}_2$ -Based Crossbar RRAM With Inserting Large Bandgap Tunneling Barrier Layer," *IEEE Trans Electron Devices*, vol. 62, no. 11, pp. 3665–3670, Sep. 2015, doi: 10.1109/TED.2015.2471835.
- [152] W. Banerjee, X. Xu, H. Lv, Q. Liu, S. Long, and M. Liu, "Variability improvement of  $\text{TiO}_x/\text{Al}_2\text{O}_3$  bilayer nonvolatile resistive switching devices by interfacial band engineering with an ultrathin  $\text{Al}_2\text{O}_3$  dielectric material," *ACS Omega*, vol. 2, no. 10, pp. 6888–6895, Oct. 2017, doi: 10.1021/acsomega.7b01211.
- [153] C. S. Peng, W. Y. Chang, Y. H. Lee, M. H. Lin, F. Chen, and M. J. Tsai, "Improvement of resistive switching stability of  $\text{HfO}_2$  films with Al doping by atomic layer deposition," *Electrochemical and Solid-State Letters*, vol. 15, no. 4, 2012, doi: 10.1149/2.011204esl.
- [154] S. Yu *et al.*, "Improved uniformity of resistive switching behaviors in  $\text{HfO}_2$  thin films with embedded Al layers," *Electrochemical and Solid-State Letters*, vol. 13, no. 2, 2010, doi: 10.1149/1.3267050.
- [155] S. Yu, Y. Wu, Y. Chai, J. Provine, and H. S. P. Wong, "Characterization of switching parameters and multilevel capability in  $\text{HfO}_x/\text{AlO}_x$  bi-layer RRAM devices," in *International Symposium on VLSI Technology, Systems, and Applications, Proceedings*, 2011, pp. 106–107. doi: 10.1109/VTSA.2011.5872251.

- 
- [156] J. W. McPherson, J. Kim, A. Shanware, H. Mogul, and J. Rodriguez, "Trends in the ultimate breakdown strength of high dielectric-constant materials," *IEEE Trans Electron Devices*, vol. 50, no. 8, pp. 1771–1778, Aug. 2003, doi: 10.1109/TED.2003.815141.
- [157] A. S. Sokolov *et al.*, "Comparative study of Al<sub>2</sub>O<sub>3</sub>, HfO<sub>2</sub>, and HfAlO<sub>x</sub> for improved self-compliance bipolar resistive switching," *Journal of the American Ceramic Society*, vol. 100, no. 12, pp. 5638–5648, Dec. 2017, doi: 10.1111/jace.15100.
- [158] J. Reuben, M. Biglari, and D. Fey, "Incorporating variability of resistive RAM in circuit simulations using the Stanford–PKU model," *IEEE Trans Nanotechnol*, vol. 19, pp. 508–518, 2020.
- [159] J. Reuben, D. Fey, and C. Wenger, "A modeling methodology for resistive ram based on stanford-pku model with extended multilevel capability," *IEEE Trans Nanotechnol*, vol. 18, pp. 647–656, 2019.
- [160] D. Ielmini and V. Milo, "Physics-based modeling approaches of resistive switching devices for memory and in-memory computing applications," *J Comput Electron*, vol. 16, no. 4, pp. 1121–1143, 2017.
- [161] J. Borghetti, D. B. Strukov, M. D. Pickett, J. J. Yang, D. R. Stewart, and R. S. Williams, "Electrical transport and thermometry of electroformed titanium dioxide memristive switches," *J Appl Phys*, vol. 106, no. 12, p. 124504, 2009.
- [162] F.-Y. Yuan *et al.*, "Conduction mechanism and improved endurance in HfO<sub>2</sub>-based RRAM with nitridation treatment," *Nanoscale Res Lett*, vol. 12, no. 1, pp. 1–6, 2017.
- [163] D. Ielmini, F. Nardi, and C. Cagli, "Physical models of size-dependent nanofilament formation and rupture in NiO resistive switching memories," *Nanotechnology*, vol. 22, no. 25, p. 254022, 2011.
- [164] E. Miranda and J. Suñé, "Analytic modeling of leakage current through multiple breakdown paths in SiO<sub>2</sub>/sub 2/films," in *2001 IEEE International Reliability Physics Symposium Proceedings. 39th Annual (Cat. No. 00CH37167)*, 2001, pp. 367–379.
- [165] E. A. Miranda, C. Walczyk, C. Wenger, and T. Schroeder, "Model for the resistive switching effect in HfO<sub>2</sub> MIM structures based on the transmission properties of narrow constrictions," *IEEE Electron Device Letters*, vol. 31, no. 6, pp. 609–611, Jun. 2010, doi: 10.1109/LED.2010.2046310.

- 
- [166] L. M. Prócel *et al.*, “Experimental evidence of the quantum point contact theory in the conduction mechanism of bipolar HfO<sub>2</sub>-based resistive random-access memories,” *J Appl Phys*, vol. 114, no. 7, Aug. 2013, doi: 10.1063/1.4818499.
- [167] X. Lian *et al.*, “Quantum point contact model of filamentary conduction in resistive switching memories,” in *2012 13th International Conference on Ultimate Integration on Silicon, ULIS 2012*, 2012, pp. 101–104. doi: 10.1109/ULIS.2012.6193367.
- [168] A. Grossi *et al.*, “Electrical characterization and modeling of 1T-1R RRAM arrays with amorphous and poly-crystalline HfO<sub>2</sub>,” *Solid State Electron*, vol. 128, pp. 187–193, 2017.
- [169] J. Wang and F. Zhuge, “Memristive Synapses for Brain-Inspired Computing,” *Adv Mater Technol*, vol. 4, no. 3, p. 1800544, 2019.
- [170] S. Yu, Y. Wu, R. Jeyasingh, D. Kuzum, and H.-S. P. Wong, “An electronic synapse device based on metal oxide resistive switching memory for neuromorphic computation,” *IEEE Trans Electron Devices*, vol. 58, no. 8, pp. 2729–2737, 2011.
- [171] J. D. Kendall and S. Kumar, “The building blocks of a brain-inspired computer,” *Appl Phys Rev*, vol. 7, no. 1, p. 11305, 2020.
- [172] W. Wan *et al.*, “A compute-in-memory chip based on resistive random-access memory,” *Nature*, vol. 608, no. 7923, pp. 504–512, 2022.
- [173] E. Aljalbout, F. Walter, F. Röhrbein, and A. Knoll, “Task-independent spiking central pattern generator: A learning-based approach,” *Neural Process Lett*, vol. 51, no. 3, pp. 2751–2764, 2020.
- [174] A. Aldweesh, A. Derhab, and A. Z. Emam, “Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues,” *Knowl Based Syst*, vol. 189, p. 105124, 2020.
- [175] F. Zahari, E. Pérez, M. K. Mahadevaiah, H. Kohlstedt, C. Wenger, and M. Ziegler, “Analogue pattern recognition with stochastic switching binary CMOS-integrated memristive devices,” *Sci Rep*, vol. 10, no. 1, pp. 1–15, 2020.
- [176] E. A. Cartier *et al.*, “Reliability Challenges with Materials for Analog Computing,” in *2019 IEEE International Reliability Physics Symposium (IRPS)*, 2019, pp. 1–10. doi: 10.1109/IRPS.2019.8720599.

- 
- [177] S. Yu, "Neuro-inspired computing with emerging nonvolatile memories," *Proceedings of the IEEE*, vol. 106, no. 2, pp. 260–285, 2018.
- [178] C. Wenger *et al.*, "Inherent stochastic learning in CMOS-integrated HfO<sub>2</sub> arrays for neuromorphic computing," *IEEE Electron Device Letters*, vol. 40, no. 4, pp. 639–642, 2019.
- [179] C. Turchetti, *Stochastic models of neural networks*, vol. 102. IOS Press, 2004.
- [180] Q. Xia and J. J. Yang, "Memristive crossbar arrays for brain-inspired computing," *Nat Mater*, vol. 18, no. 4, pp. 309–323, 2019.
- [181] M. K. Mahadevaiah *et al.*, "Reliability of CMOS Integrated Memristive HfO<sub>2</sub> Arrays with Respect to Neuromorphic Computing," in *2019 IEEE International Reliability Physics Symposium (IRPS)*, 2019, pp. 1–4. doi: 10.1109/IRPS.2019.8720552.
- [182] S. G. Kim, J. S. Han, H. Kim, S. Y. Kim, and H. W. Jang, "Recent advances in memristive materials for artificial synapses," *Adv Mater Technol*, vol. 3, no. 12, p. 1800457, 2018.
- [183] T. Zhang, K. Yang, X. Xu, Y. Cai, Y. Yang, and R. Huang, "Memristive Devices and Networks for Brain-Inspired Computing," *physica status solidi (RRL)–Rapid Research Letters*, vol. 13, no. 8, p. 1900029, 2019.
- [184] M. Ziegler, C. Wenger, E. Chicca, and H. Kohlstedt, "Tutorial: Concepts for closely mimicking biological learning with memristive devices: Principles to emulate cellular forms of learning," *J Appl Phys*, vol. 124, no. 15, p. 152003, 2018.
- [185] S. Petzold *et al.*, "Forming-Free Grain Boundary Engineered Hafnium Oxide Resistive Random-Access Memory Devices," *Adv Electron Mater*, vol. 5, no. 10, p. 1900484, 2019.
- [186] M. Ignatov, M. Ziegler, M. Hansen, and H. Kohlstedt, "Memristive stochastic plasticity enables mimicking of neural synchrony: Memristive circuit emulates an optical illusion," *Sci Adv*, vol. 3, no. 10, p. e1700849, 2017.
- [187] J. C. López, "Quantifying synaptic efficacy," *Nat Rev Neurosci*, vol. 3, no. 5, p. 332, 2002.
- [188] M. Sowinska *et al.*, "In-operando hard X-ray photoelectron spectroscopy study on the impact of current compliance and switching cycles on oxygen and carbon

defects in resistive switching Ti/HfO<sub>2</sub>/TiN cells," *J Appl Phys*, vol. 115, no. 20, p. 204509, 2014.

- [189] M. Zhao *et al.*, "Investigation of statistical retention of filamentary analog RRAM for neuromorphic computing," in *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 34–39.



---

## Appendix A

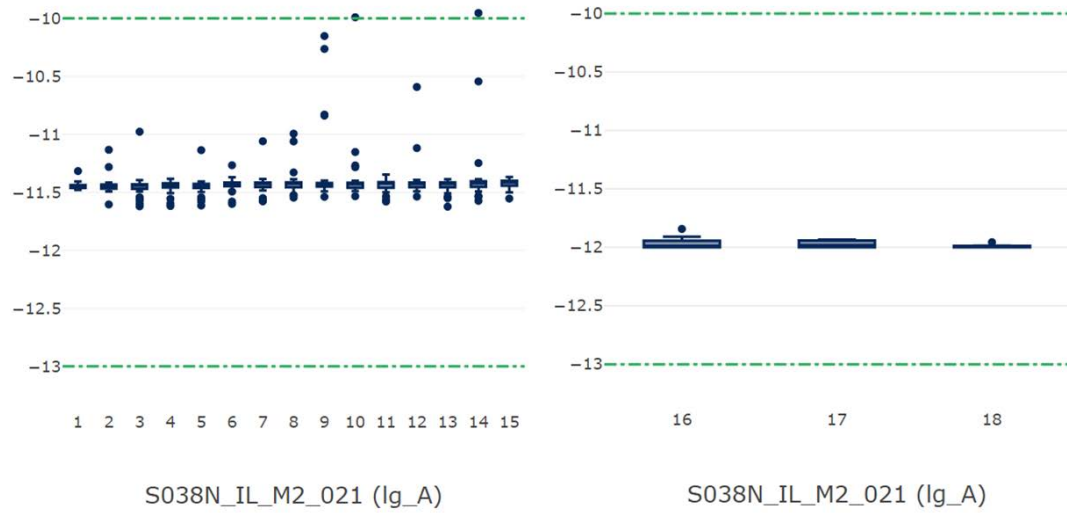
---

### **BEOL In-line Measurement Results**

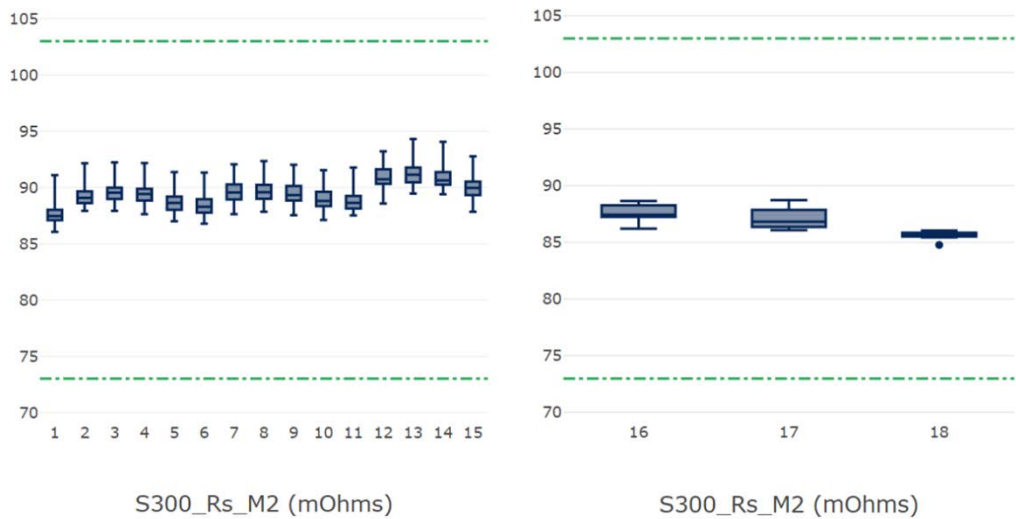
**Standard Met 2 in-line Measurements**

**Wafers 1-15 standard SG13S**

**Wafers 16-18 MEMRES + SG13S**



**Leakage current measurement**

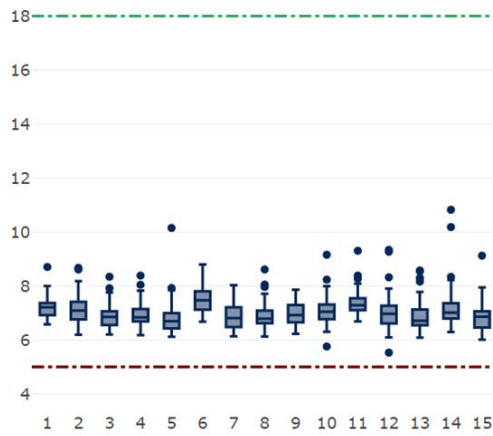


**Sheet resistance measurement**

---

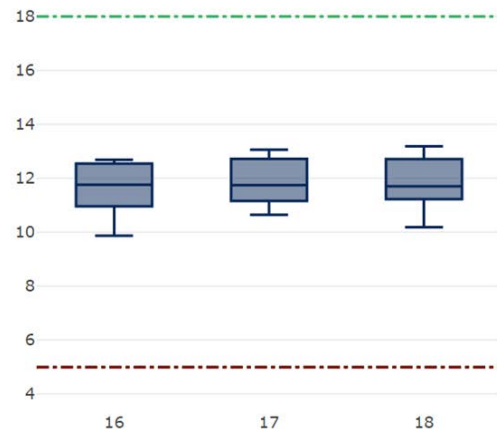
## Standard Via 2 in-line Measurements

### Wafers 1-15 standard SG13S



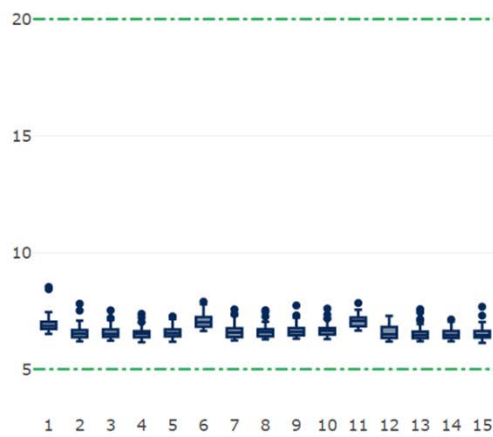
S194D\_RkVia2 (Ohm)

### Wafers 16-18 MEMRES + SG13S

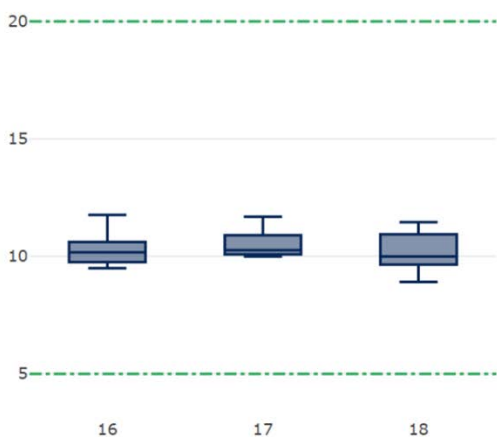


S194D\_RkVia2 (Ohm)

### Kelvin contact measurement



S179N\_RccVia2 (Ohm/via)



S179N\_RccVia2 (Ohm/via)

### Standard resistance measurement

---

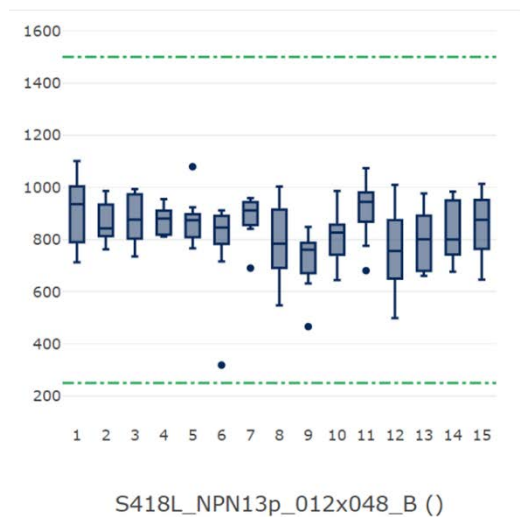
## Appendix B

---

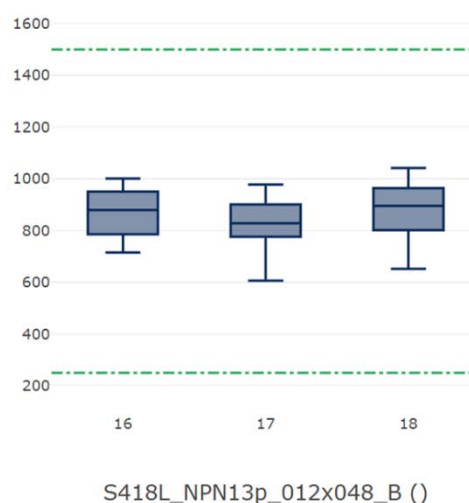
### **FEOL In-line Measurement Results**

## Standard HBT Measurements

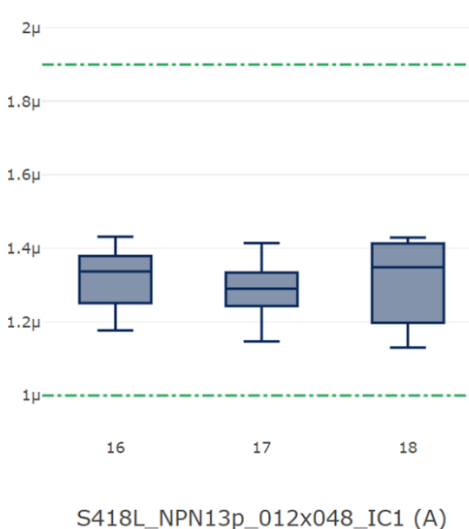
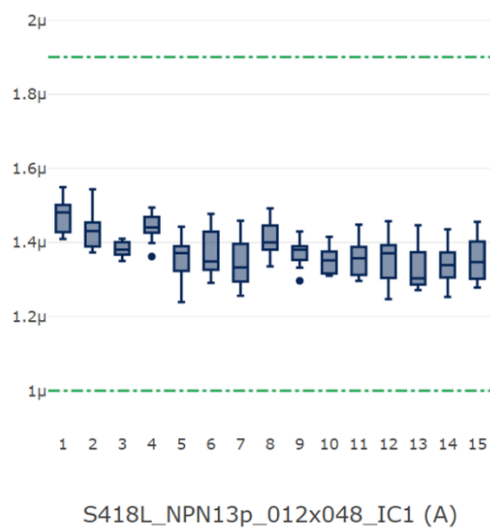
### Wafers 1-15 standard SG13S



### Wafers 16-18 MEMRES + SG13S



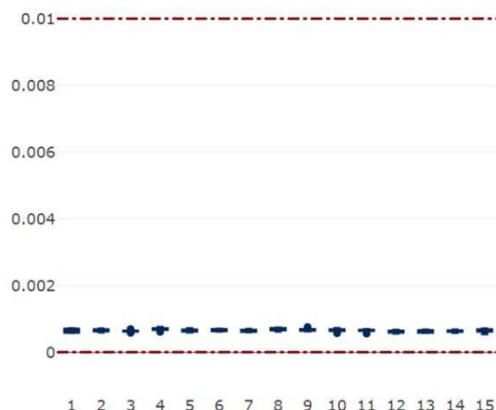
### Current gain of an HBT with an emitter size of $0.12 \mu\text{m} \times 0.48 \mu\text{m}$



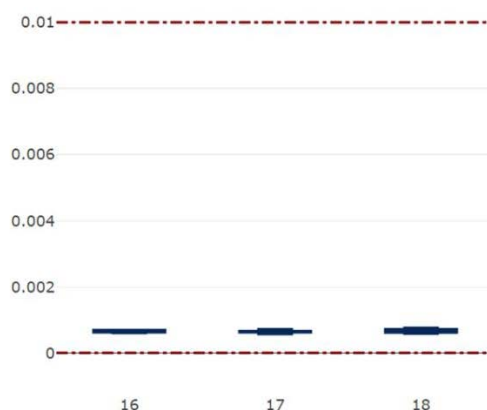
### Collector current of an HBT with a base-emitter voltage of 0.7 V

**Wafers 1-15 standard SG13S**

**Wafers 16-18 MEMRES + SG13S**

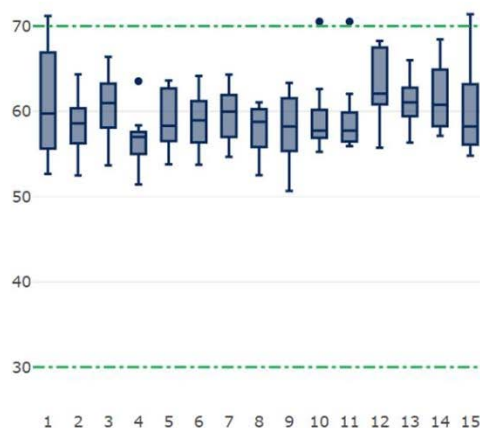


S418L\_NPN13p\_012x048\_IC3 (A)

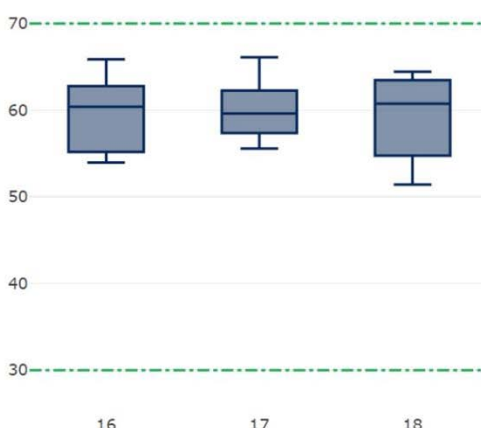


S418L\_NPN13p\_012x048\_IC3 (A)

**Collector current of an HBT with a base-emitter voltage of 0.9 V**



S418L\_NPN13p\_012x048\_RE (Ohm)

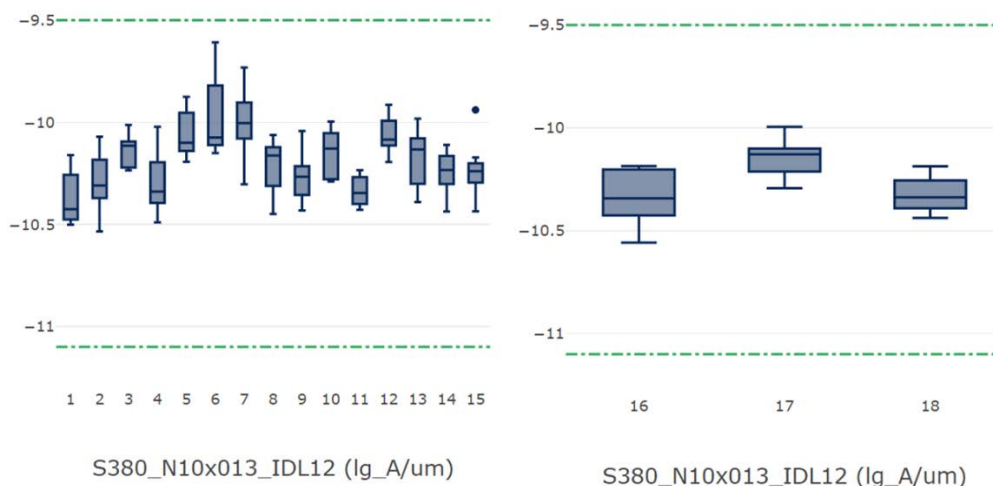
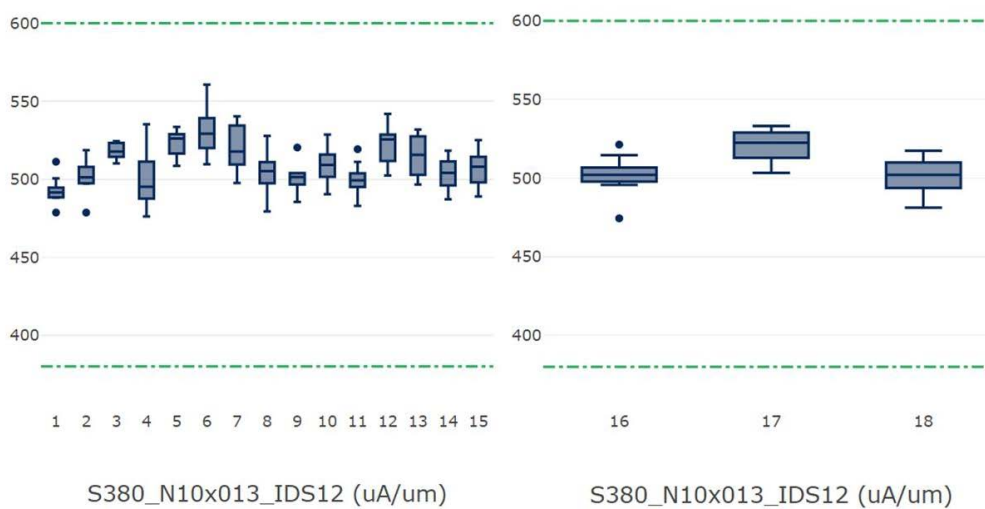


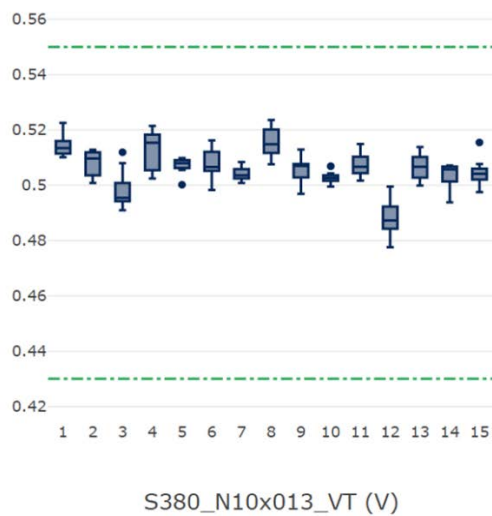
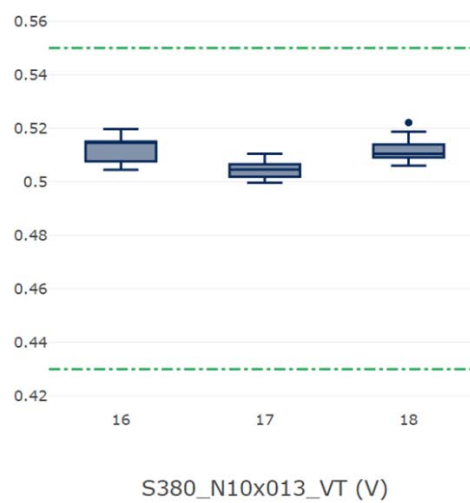
S418L\_NPN13p\_012x048\_RE (Ohm)

**Resistance of the emitter**

---

## Standard nMOS Transistor Measurements

**Wafers 1-15 standard SG13S****Wafers 16-18 MEMRES + SG13S****Drain leakage current of a nMOS with L = 130 nm and W = 10  $\mu\text{m}$** **Drain saturation current of a nMOS with L = 130 nm and W = 10  $\mu\text{m}$**

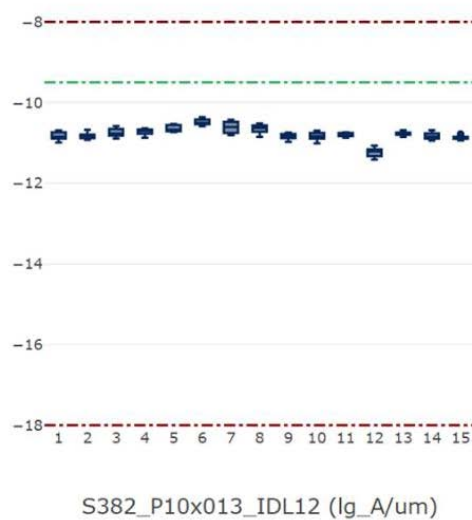
**Wafers 1-15 standard SG13S****Wafers 16-18 MEMRES + SG13S**

**Threshold voltage of a nMOS with  $L = 130$  nm and  $W = 10$   $\mu$ m**

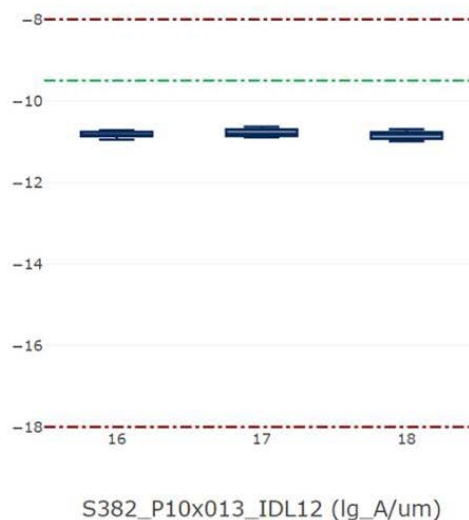


## Standard pMOS Transistor Measurements

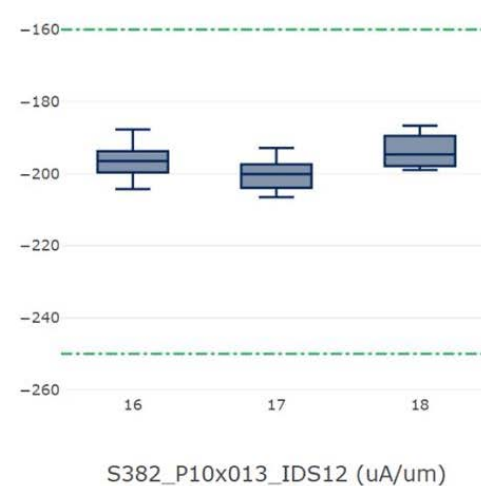
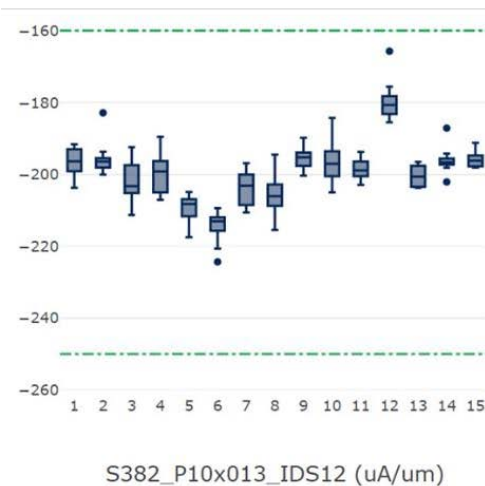
### Wafers 1-15 standard SG13S



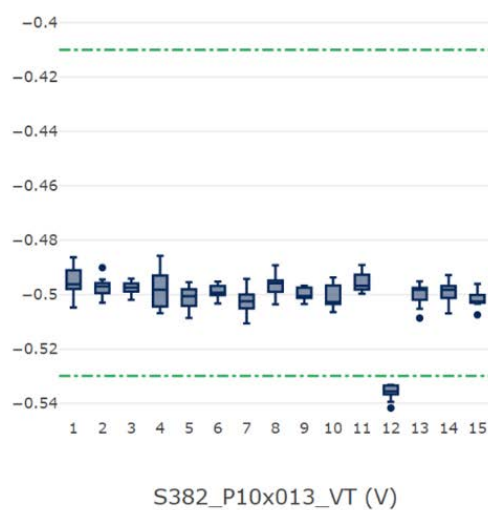
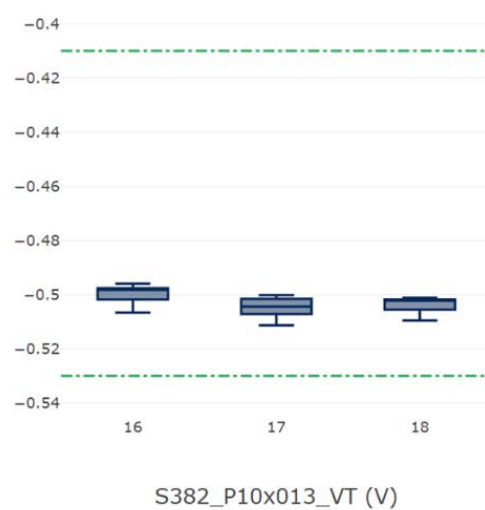
### Wafers 16-18 MEMRES + SG13S



### Drain leakage current of a pMOS with $L = 130$ nm and $W = 10$ $\mu\text{m}$



### Drain saturation current of a pMOS with $L = 130$ nm and $W = 10$ $\mu\text{m}$

**Wafers 1-15 standard SG13S****Wafers 16-18 MEMRES + SG13S**

**Threshold voltage of a pMOS with  $L = 130$  nm and  $W = 10$   $\mu$ m**