UNIVERSITAT DE BARCELONA

Optimization of Memristors for Information Storage and Neuromorphic Computing

Kaichen Zhu

PhD thesis dissertation

Optimization of Memristors for Information Storage and Neuromorphic Computing

Author: Mr. Kaichen Zhu

Director: Dr. Mario Lanza

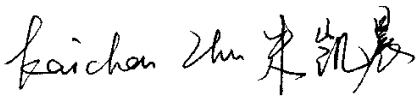
UNIVERSITAT DE BARCELONA

Optimization of Memristors for Information Storage and Neuromorphic Computing

Memoria presentada para optar al grado de doctor por la

Universidad de Barcelona

Programa de doctorado en Nanociencias



Autor: Mr. Kaichen Zhu

Director: Dr. Mario Lanza

Tutor: Dr. Juan Marcos Fernandez Pradas

UNIVERSITAT DE BARCELONA

To my family,

Acknowledgements

I would like to take this opportunity to express my deep gratitude to my PhD supervisor Professor Mario Lanza. We conceived many interesting ideas and enjoyed doing the experiment together in the laboratory and discussing the papers. His hardworking attitude and selfless sharing let me know what a real professor should be like. With the great help of Mario and a huge amount of fabrication and characterization, I am now a “little” expert on two-dimensional materials and memristors, and I do have a lot of fun during my PhD study.

I also would like to appreciate all the group members in Lanza’s group for their kind help in research and life, especially Chengbin Pan, Xianhu Liang, Fei Hui, Yuanyuan Shi, Tingting Han, Na Xiao, Xu Jing, Bin Yuan, Shaochuan Chen, Tao Wang, Mahata, Marco, Jonas, Xuehua Li, Chao Wen, Yiping Xiao, Yaqing Shen and so on. The senior students trained me well with solid foundations, and I also learned a lot from teaching younger students. When I write the names down, all kinds of good memories flash back to me.

I also would like to appreciate Professor Blas Garrido for hosting me and my tutor Professor Juan Marcos Fernandez for coordinating my PhD program at the University of Barcelona. And thanks a lot for the help from the friends there, especially Dr Giovanni Vescio, Juan L. Frieiro, and Julia L. Vidrier. Special thanks to Haicheng Cao, Zhen Yang and Yinchang Ma at King Abdullah University of Science and Technology (KAUST) who helped me adapt to Saudi life. Special thanks to Ren Li for helping me with the pulse measurement of the memristor. These days' and nights’ measurements are exciting and fruitful. Special thanks to Fernando Aguirre and Sebastian Pazos regarding the neuron networking training and programming.

All the technicians from the Institute of Functional Nano and Soft Materials (FUNSOM) of Soochow University, the University of Barcelona and KAUST deserve my acknowledgement for their technical advice and help in using the equipment in the institute, as well as the administration staff for all their help.

Finally, I would like to thank my family’s strong support and my girlfriend Chunchun Zhang’s accompany. 最后我想感谢我的家人，谢谢他们这么多年来对我的关心，牵挂，鼓励和等待。

Index

Contents

Abstract

The total amount of global information to be stored and computed is increasing exponentially since the beginning of the 21st century. The semiconductor industry made great efforts to push the silicon-based integrated circuits (ICs) to follow Moore's Law to achieve better performance and meet the requirements. However, the basic element of the modern ICs, the transistor, is approaching its physical limitation of further scaling down. There are two main solutions are proposed to solve these problems. One is "More than Moore" by replacing the transistor with a new electronic device, for example, a memristor. The other one is "More Moore" by introducing two-dimensional (2D) materials into the transistor structure. Compared to transistors, memristors are competitive with simpler device structure, scalability, and three-dimensional stackability. 2D materials have been included in the technology progression plan in the International Roadmap for Device and Systems (IRDS).

In this PhD thesis, I combined the advantages of both memristor and 2D materials and carried out a deep study on 2D materials-based memristors, especially using a 2D insulator hexagonal boron nitride (h-BN) as the resistive switching medium. I fabricated h-BN based memristors with van der Waals structure and demonstrated that the device has the capability of stable and highly controllable tristate operation by controlling the current compliance (CC) and reset voltage (V_{RESET}). I integrated the h-BN based memristor with a commercial complementary metal-oxide-semiconductor (CMOS) circuit, and h-BN memristors with a size of 0.053 μm^2 have been achieved. The formed one transistor one memristor (1T1M) cell can show high endurance (millions of cycles), multistate switching, and synaptic behaviour like spike-time-dependent-plasticity (STDP), which are useful for implementing a spiking

neural network (SNN). The circuit based on two 1T1M cells can also show matrix operations like OR and IMP logic. Apart from the 2D materials-based memristor, I also conducted several experiments based on the metal-oxide memristor for comparison.

Our study is an important step toward the 2D materials-based memristor, and its successful integration into silicon-based ICs could inspire other scientists to study 2D materials-based devices integrated into a real microchip with functional circuitry.

Abstract in official language

La cantidad total de información que se almacena y calcula a nivel global está aumentando exponencialmente desde principios del siglo XXI. La industria de los semiconductores esta haciendo grandes esfuerzos para impulsar los circuitos integrados (CI) basados en silicio para cumplir la Ley de Moore logrando un mejor rendimiento y a la vez que cumple con los requisitos. Sin embargo, el elemento básico de los circuitos integrados modernos, el transistor, se está acercando a su limitación física de reducción de tamaño. Se han propuesto dos soluciones principalmente para resolver estos problemas. La primera es la llamada "More than Moore" que se basa en reemplazar el transistor con un nuevo dispositivo electrónico, por ejemplo, un memristor. El segunda es la conocida como "More Moore" mediante la introducción de materiales bidimensionales (2D) en la estructura del transistor. En comparación con los transistores, los memristores son competitivos con una estructura de dispositivo más simple, escalabilidad y capacidad de apilamiento tridimensional. Los materiales 2D se han incluido en el plan de progreso tecnológico en la Hoja de ruta internacional para dispositivos y sistemas (IRDS).

En esta tesis doctoral, combiné las ventajas de los memristores con los materiales 2D y realicé un estudio profundo sobre los memristores basados en materiales 2D. Concretamente, he usado el nitruro de boro hexagonal (h-BN) fabricado mediante *Chemical Vapour Deposition* (CVD) como medio de conmutación resistiva. Fabriqué memristores basados en h-BN con estructura de van der Waals y demostré que el dispositivo tiene la capacidad de operar en tres estados estables y altamente controlables mediante el control de la corriente límite (CC) y el voltaje de reset (V_{RESET}). Integré este memristor basado en h-BN con un circuito comercial CMOS logrando memristores h-BN con un área de 0,053 μm^2. La celda formada por un transistor y un memristor (1T1M)

puede mostrar alta durabilidad (millones de ciclos), conmutación multi-estado, coexistencia de conmutación volátil y no volátil y comportamiento sináptico como *spike-time-dependent-plasticity* (STDP), que son útiles para implementar la red neuronal de impulsos (SNN). El transistor en serie con el memristor puede suprimir efectivamente las corriente parásitas. Además, los circuitos memristivos basados en dos celdas 1T1M también puede mostrar operaciones matriciales como las operaciones lógicas OR e IMP. Además del memristor basado en materiales 2D, también realicé varios experimentos usando memristors basados en óxidos metálicos comparar.

Nuestro estudio es un paso importante hacia el memristor basado en materiales 2D, y su integración exitosa en circuitos integrados de silicio. Esto podría inspirar a otros científicos a estudiar dispositivos basados en materiales 2D integrados en un microchip real con circuitos funcionales.

List of Figures

List of Tables

List of acronyms

IC: integrated circuit

2D: two dimensional

IRDS: International Roadmap for Device and Systems

h-BN: hexagonal boron nitride

CMOS: complementary metal oxide semiconductor

1T1M: one-transistor-one-memristor

IoT: Internet of Things

IDC: International Data Corporation

Ge: germanium

Si: silicon

SiO_2: silicon dioxide

Al: aluminium

Au: gold

high-k: high dielectric constant

MIM: Metal/insulator/metal

TiO_2: Titanium oxide

HfO_2: Hafnium oxide

TaO_x: tantalum oxide

Pt: platinum

Au: gold

Ti: titanium

Cu: copper

Ag: silver

Ni: nickel

RRAM: Resistive random access memory

AAO: anode aluminium oxide

BEOL: back end of line

CMP: chemical mechanical polishing

MPW: Multi-Project-Wafer

FoMs: Figure of merits

SMU: source-measure unit

RVS: ramped voltage stress

CVS: constant voltage stress

VLS: voltage list stress

I-V: current-voltage

I-t: current-time

CC: Current compliance

HRS: high resistive state

LRS: low resistive state

V_{SET}: set voltage

V_{RESET}: reset voltage

WGFMU: waveform-generator-fast-measurement-unit

V_{READ}: read voltage

t_w: pulse width

t_{SET}: set time

RTN: random telegraph noise

TDDB: time-dependent-dielectric-breakdown

CBRAM: Conductive bridge random access memory

OxRAM: Oxide resistive random access memory

Ta: Tantalum

Cr: chromium

W: tungsten

TiN: titanium nitride

ITO: indium tin oxide

Hf: hafnium

Zr: Zirconium

ALD: atomic layer deposition

MoS_2: molybdenum disulfide

WSe_2: tungsten diselenide

FET: field-effect transistor

CVD: chemical vapor deposition

LPE: liquid phase exfoliation

S-LRS: Soft-LRS

XTEM: Cross-section transmission electron microscope

CAFM: Conductive atomic force microscope

SEM: Scanning electron microscopy

PMMA: polymethyl-methacrylate

$FeCl_3$: ferric chloride

HCl: hydrochloric acid

FIB: focused ion beam

CNF: conductive nanofilaments

QPC: quantum point contact

G_0: quantum conductance

1M: single memristor

CV: Coefficient of variance

HAADF: high-angle annular dark-field

STEM: scanning transmission electron microscope

V_G: gate voltage

Chapter 1:
Dissertation summary

1.1 Introduction

With the rise of technologies such as big data, cloud computing and the Internet of Things (IoT), the total amount of global information has increased exponentially since the beginning of the 21st century, and most of the information is stored in the form of digital data (the records show that 94% of the stored information is digital information) (Figure 1.1a) [1]. The International Data Corporation (IDC) predicts that the global data will grow from 33 Zettabytes in 2018 to 175 Zettabytes by 2025 (Figure 1.1b) [2]. A huge amount is predicted in 2018 before the COVID-19 pandemic, while the number should be much bigger with the extensive application of online courses and cloud offices. For example, Zoom, a multi-end cloud video conference software, experienced exponential growth (Figure 1.1c) [3]. Nowadays most of the digital data are in the form of pictures, videos, movies, and data streaming, where special attention is needed.

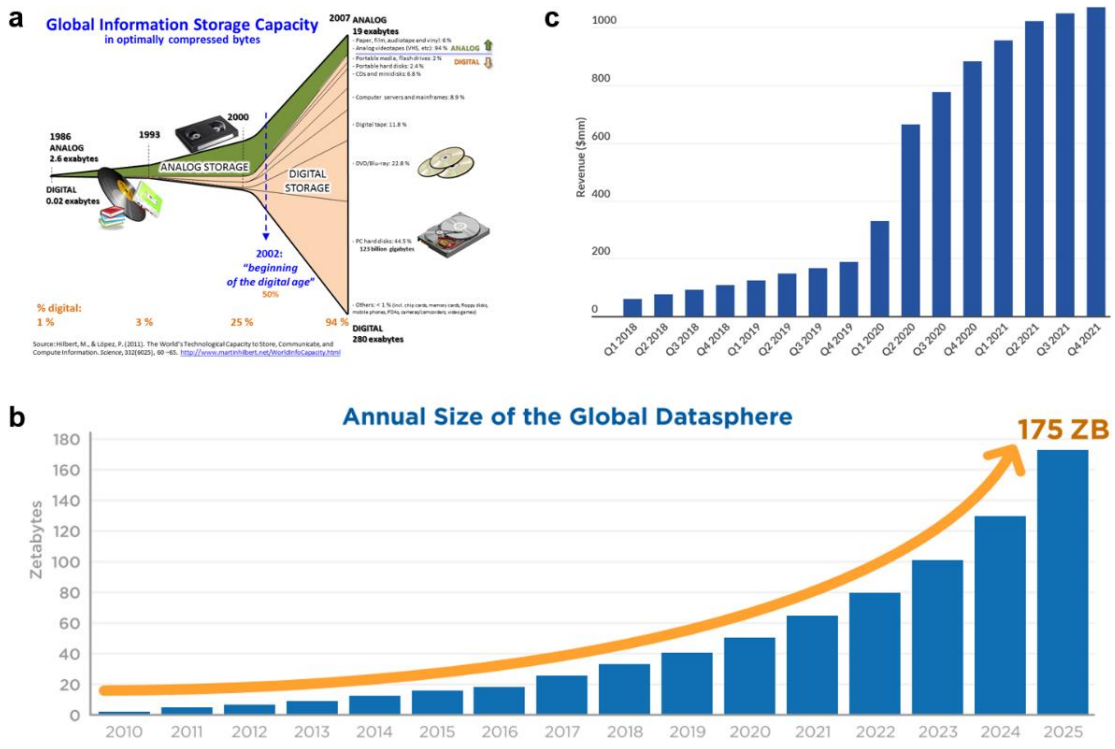


Figure 1.1 Extensive demand for digital information storage in the world market

Integrated circuits (ICs) have enabled the development of electronic systems capable of performing complex operations quickly (~0.1 ns per operation) and efficiently (~1 pJ per operation) [4]. The term IC refers to the monolithic electrical connection of two or more discrete electronic devices to create functionalities superior to those of individual devices. The first ICs were developed in the late 1950s and included transistors, resistors and capacitors made of different materials: germanium (Ge) or silicon (Si) for the body of the transistors, silicon dioxide (SiO_2) for the dielectric of the capacitors and transistor gates, aluminium (Al) for the electrodes and gold (Au) for the interconnections [5–7]. Today, ICs are often referred to as silicon technology, but they contain hundreds of materials, all of which have controlled morphologies and specific properties.

With the establishment of ICs, both industry and academia are working hard to "cramming more components onto integrated circuits" to get a more powerful computer. On the large scale, the development of ICs follows Moore's Law predicted by Gordon E. Moore in 1965 [8-9]. The number of transistors on microchips doubles every two years, as shown in Figure 1.2. With chip size staying the same, it means the size of the transistor should be scaled down. State-of-the-art ICs for data computation and storage are made from ~11.8 billion FinFETs from the 5-nm node, reaching an integration density of 1.34×10^8 transistors per mm^2 [10].

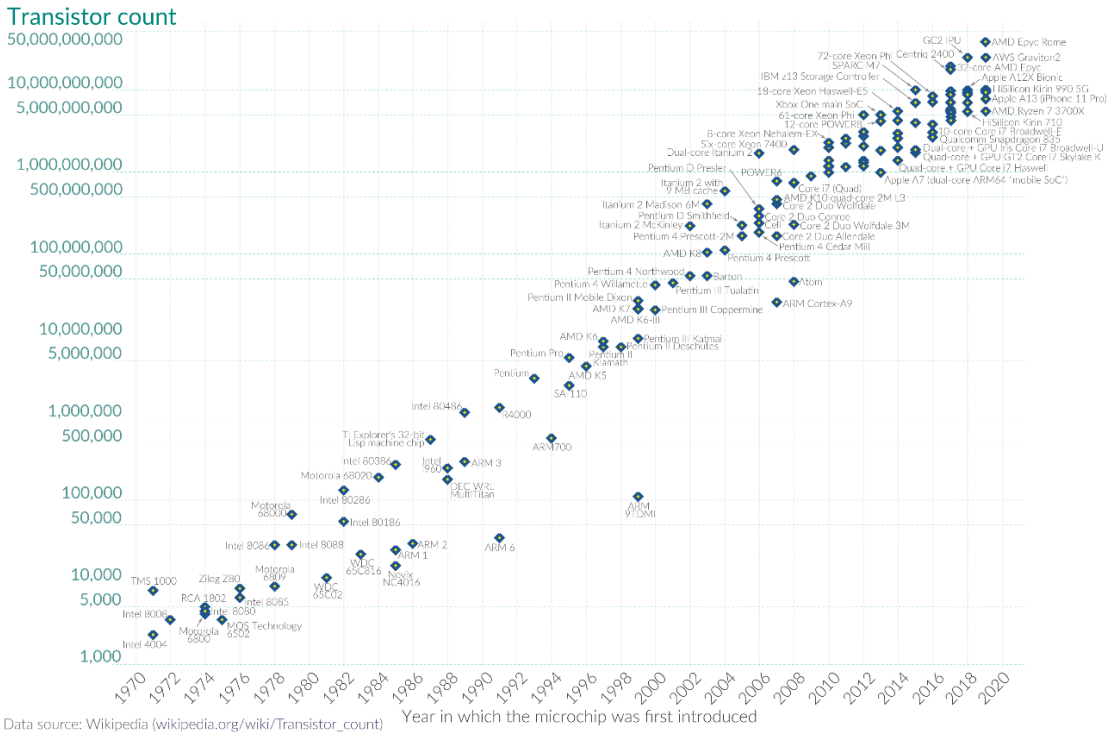


Figure 1.2 Moore's Law: number of transistors on microchips doubles every two years

In the microscopic view, the scaling down of the basic component of IC, the transistor, should follow Dennard's Law [11]. A transistor is a three-terminal device with gate voltage controlling the current flow between the drain and the source. Dennard's Law states the rule of how to design the scaled-down transistor structure with parameters like the length and width of channel and voltage applied, and evaluate the speed, power consumption and cost. For example, the electric field will stay still, while the operation voltage is scaled simultaneously with the scaled dimension.

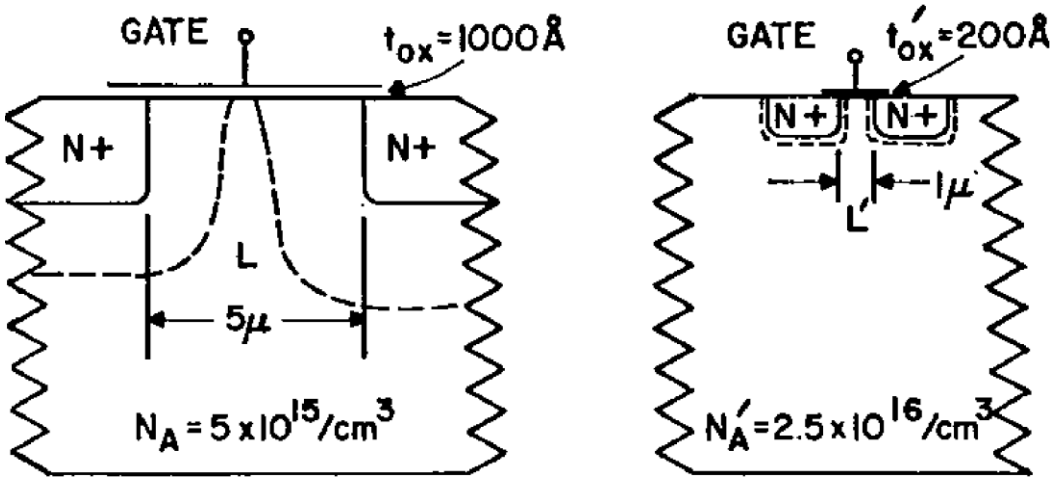


Figure 1.3 Dennard's Law illustrating transistor scaling with a value of 5

Since the 1990s, the semiconductor industry has issued roadmaps for the development of semiconductor technology every two years to coordinate hundreds of research groups, manufacturers and suppliers to meet Moore's law. For example, high dielectric constant (high-k) gate dielectrics, strained silicon, FinFET and the metal gate have been introduced to help extend the technology node to 5 nm [4]. However, the feature size miniaturization of transistors based on electron mechanism is gradually moving towards the limit of physical size. Current leakage in the vertical direction across the dielectrics and the cross-talk in the horizontal direction and heat between adjacent devices are unavoidable for smaller devices with thinner dielectrics and higher device density. Also, the short-channel effect and thermal death caused by miniaturization are more difficult. The final result is that under the influence of the triple factors of performance, power consumption and cost, researchers have changed their research focus to develop new electronic devices based on different storage mechanisms to replace the traditional three-terminal transistors and introduce novel materials in the transistor.

1.2 Main contribution of the thesis

1.2.1 Objectives of the Thesis

This thesis mainly focuses on the design, fabrication, characterization and optimization of h-BN based memristors, with the main goal of achieving a circuit-level integration with the mature Si-based CMOS technology. The h-BN based memristive circuit should have improved performance compared to a single device and should show circuit-level operation between two or more operation cells.

To achieve the main goal, the task can be divided into three main objectives:

1) Design and fabricate the competitive h-BN based memristors. It is worth noting that, to achieve a circuit-level integration, all the fabrication methods involved should be scalable. Optimization methods including different metallic electrode combinations and deposition methods, graphene insertion, wet transfer of 2D materials, and 1T1M cell structure should be implemented.
2) Characterize h-BN based memristors in both structure and electrical performance. The correct layer structure of 2D materials and device structure should be carefully examined by cross-sectional TEM. Raman spectrum is an optional method to statistically analyse 2D materials. Memristors with device area smaller than 0.1 μm^2, endurance more than 1×10^6, operation speed faster than 10 MHz (100 ns), and switching energy less than 0.1 pJ are recommended. Device-to-device variability should be analysed. Synaptic behaviour like spike-time-dependent-plasticity (STDP), multilevel, and both volatile and non-volatile switching capability should be explored.
3) Understand the switching mechanism of the h-BN based memristor for further optimization. Conductive AFM should be used to study the h-BN locally with a nanometer scale. Different scale simulations should be developed from material level, device level to circuit level.
4) Integrate the h-BN based memristor with the CMOS circuit. Matrix operations are expected to see, including the vector-matrix multiplication and logic OR and IMP. The more advanced function with more operation cells should be simulated including spiking neural network (SNN).

1.2.2 Key findings

I would like to draw attention to the main contributions of this work and the importance can be stated because of the following reasons:

We present the first 2D material-based memristors that exhibit three stable and well-distinguishable resistive states. By using a multilayer h-BN stack sandwiched by multilayer graphene electrodes, we fabricate 5 μm × 5 μm cross-point Au/Ti/Graphene/h-BN/Graphene/Au memristors that can switch between each two or three resistive states, depending on the current limitation and reset voltage used

We present an industry-compatible approach to fabricate high-performance CMOS/2D materials hybrid microchips. We observe that the CMOS transistors provide outstanding control over the currents flowing across the h-BN memristors and remarkably increase their performance. We present the longest endurance and best conductance controllability ever reported for any 2D materials based memristor. The hybrid CMOS/2D materials one-transistor-one-memristor cells also exhibit spike-timing-dependent plasticity, which is suitable for the implementation of spiking neural networks with high accuracy and low power consumption during image classification.

We present the first fabrication of memristors that are initially in a low resistive state, which exhibit homogenous initial resistance and switching voltages. When used as electronic synapses in a neuromorphic system to classify images from the CIFAR-10 dataset, the memristors offer 1.83 better throughputs per area and consume 0.85 less energy than standard memristors (i.e., with the necessity of forming), which stems from 63% better density and 17% faster operation. It is demonstrated in the results that tuning the local properties of materials embedded in memristive electronic synapses is an attractive strategy that can lead to improved neuromorphic performance at the system level.

Metal oxide-based memristors are fabricated as comparison samples. Failure experiences are shared and suggestions are provided.

1.2.3 Thesis Outline

The thesis is divided into five chapters: Chapter 1 provides the dissertation overview and introduces the most relevant aspects of the thesis. Chapter 2 provides background information on 2D materials based memristor. Chapter 3 focuses on the fabrication and characterization of Au/Ti/Graphene/h-BN/Graphene/Au memristors with the tristate operation. Chapter 4 focuses on the integration of two-dimensional layered materials at the back-end-of-line of silicon microchips for high-performance memristive circuits. Chapter 5 focuses on the fabrication and characterization of memristors with an initial low-resistive state for efficient neuromorphic systems. Chapter 6 summarizes the main results of this thesis, as well as the conclusions that are drawn from the research.

1.3 List of publications

Only the publications included below shall be considered for the evaluation of this PhD Dissertation. A reproduction of each publication can be found on the indicated pages below. A complete list of the author's publications (updated on September 29th 2022) is included in the scientific curriculum vitae (Appendix A).

Article 1: **Kaichen Zhu**, Bin Yuan, Xianhu Liang, Xu Jing, Chao Wen, Marco A. Villena, Mario Lanza*, Graphene-Boron Nitride-Graphene Cross-Point Memristors with Three Stable Resistive States, ***ACS Applied Materials & Interfaces***. 11, 37999, 2019.

** The author's contribution: plan and perform the experiments, evaluate the results and write the main parts of the manuscript.*

Article 2: **Kaichen Zhu**, Xianhu Liang, Bin Yuan, Marco A. Villena, Chao Wen, Tao Wang, Shaochuan Chen, Mario Lanza*, Fei Hui, Yuanyuan Shi, Tristate resistive switching in heterogenous van der Waals dielectric structures, ***2019 IEEE International Reliability Physics Symposium (IRPS)***, 11th March – 4th April 2019, Monterey, California (USA). DOI: 10.1109/IRPS.2019.8720485.

** The author's contribution: plan and perform the experiments and write the main parts of the manuscript.*

Article 3: **Kaichen Zhu**, Mohammad Reza Mahmoodi, Zahra Fahimi, Yiping Xiao, Tao Wang, Kristýna Bukvišová, Miroslav Kolíbal, Juan B. Roldan, David Perez, Fernando Aguirre, Mario Lanza*, Memristors with Initial Low-Resistive State for Efficient Neuromorphic Systems, ***Advanced Intelligent Systems*** 2200001, 2022.

** The author's contribution: plan and perform the experiments, evaluate the results, and write the main parts of the manuscript.*

Article 4: The article has been SUBMITTED; **Kaichen Zhu**, Fernando Aguirre, Sebastian Pazos, Xinyi Li, Yaqing Shen, Juan B. Roldan, Ren Li, Yue Yuan, Bin Fang, Tao Wang, Hossein Fariborzi, Guenther Benstetter, Xixiang Zhang, Tibor Grasser, Daniele Ielmini, Huaqiang Wu, Mario Lanza*, Hybrid 2D/CMOS microchips for memristive technologies.

** The author's contribution: plan and perform the experiments, evaluate the results and write the main parts of the manuscript.*

Article 4: **Kaichen Zhu**, Chao Wen, Areej A. Aljarb, Fei Xue, Xiangming Xu, Vincent Tung, Xixiang Zhang, Husam N. Alshareef, Mario Lanza*, "The development of integrated circuits based on two-dimensional materials", ***Nature Electronics***. 4, 775-785, 2021.

** The author's contribution: perform the literature research and write the manuscript.*

Article S1: Ying Zuo, Huizi Lin, Jingchun Guo, Yue Yuan, Hanglin He, Yutong Li, Yiping Xiao, Xuehua Li, **Kaichen Zhu**, Tao Wang, Xu Jing, Chao Wen, Mario Lanza*, "Effect of the Pressure Exerted by Probe Station Tips in the Electrical Characteristics of Memristors", ***Advanced Electronic Materials***, 1901226, 2020.

** The author's contribution: plan and perform parts of the experiments.*

Article S2: Xuehua Li, Tommaso Zanotti, Tao Wang, **Kaichen Zhu**, Francesco Maria Puglisi, Mario Lanza*, "Random Telegraph Noise in Metal-Oxide Memristors for True Random Number Generators: A Materials Study", ***Advanced Functional Materials***. 2102172, 2021.

** The author's contribution: perform parts of the experiments.*

Article S3: Chao Wen, Xuehua Li, Tommaso Zanotti, Francesco Maria Puglisi, Yuanyuan Shi, Fernan Saiz, Aleandro Antidormi, Stephan Roche, Wenwen Zheng, Xianhu Liang, Jiaxin Hu, Steffen Duhm, Juan B. Roldan, Tianru Wu, Victoria Chen, Eric Pop, Blas Garrido, **Kaichen Zhu**, Fei Hui, Mario Lanza*, "Advanced Data Encryption using 2D Materials", ***Advanced Materials***. 2100185, 2021.

** The author's contribution: plan and perform parts of the experiments and write parts of the manuscript.*

Article S4: Tao Wang, Yuanyuan Shi, Francesco Maria Puglisi, Shaochuan Chen, **Kaichen Zhu**, Ying Zuo, Xuehua Li, Xu Jing, Tingting Han, Biyu Guo, Kristýna Bukvišová, Lukáš Kachtík, Miroslav Kolíbal, Chao Wen, Mario Lanza*, "Electroforming in metal-oxide memristive synapses", ***ACS Applied Materials & Interfaces***. 12, 11806, 2020.

** The author's contribution: plan and perform parts of the experiments.*

Article S5: Yingwen Liu, **Kaichen Zhu**, Fei Hui, Bin Yuan, Chenhui Zhang, Yinchang Ma, Xixiang Zhang, and Mario Lanza*, "Inkjet Printing: A Cheap and Easy-to-Use Alternative to Wire Bonding for Academics", ***Crystal Research & Technology*** 57, 2100210, 2022.

** The author's contribution: plan and perform the experiments, evaluate the results and write parts of the manuscript.*

Article S6: Wenwen Zheng, Fernan Saiz, Yaqing Shen, **Kaichen Zhu**, Yingwen Liu, Clifford McAleese, Ben Conran, Xiaochen Wang, Mario Lanza*, "Defect-free Metal Deposition on Two-dimensional Materials via Inkjet Printing Technology", ***Advanced Materials***. 2104138, 2021.

** The author's contribution: perform the experiments and evaluate of the results.*

Article S7: Yaqing Shen, Wenwen Zheng, **Kaichen Zhu**, Yiping Xiao, Chao Wen, Yingwen Liu, Xu Jing, Mario Lanza*, "Variability and Yield in h-BN-Based Memristive Circuits: The Role of Each Type of Defect", ***Advanced Materials***. 2103656, 2021.

** The author's contribution: perform the experiments and evaluate of the results.*

Chapter 2: Two-dimensional materials based memristors for data storage and computing

2.1 More than Moore using a new electronic device - memristor

Memristor is a two-terminal electronic device with a simple metal/insulator/metal (MIM) sandwich structure. Its resistance can be controlled between high resistive state (HRS) and low resistive state (LRS) by electrical stresses applied to the electrodes [12]. This resistance change is also named resistive switch (RS) behaviour. Most memristors have two stable resistive states, which can be used to simulate binary codes "0"s and "1"s to store digital information. The device usually uses metal oxides such as titanium oxide (TiO_2), silicon oxide (SiO_2), hafnium oxide (HfO_2), tantalum oxide (TaOx) and other insulators as the resistive switching medium, and conventional metals such as platinum (Pt), gold (Au), titanium (Ti), copper (Cu), silver (Ag) and nickel (Ni) as the electrodes. Memristors show many excellent properties compared to traditional memory, such as small size (feature size 2 nm) [13], low energy consumption (each switching < 1 fJ) [14], fast switching speed (85 ps) [15], high endurance (> 10^{12} cycles) [16], long retention time (> 10 years) [17], easy fabrication and three-dimensional stackability. Figure 2.1 shows the typical sandwich structure with 2D materials as resistive switching medium.

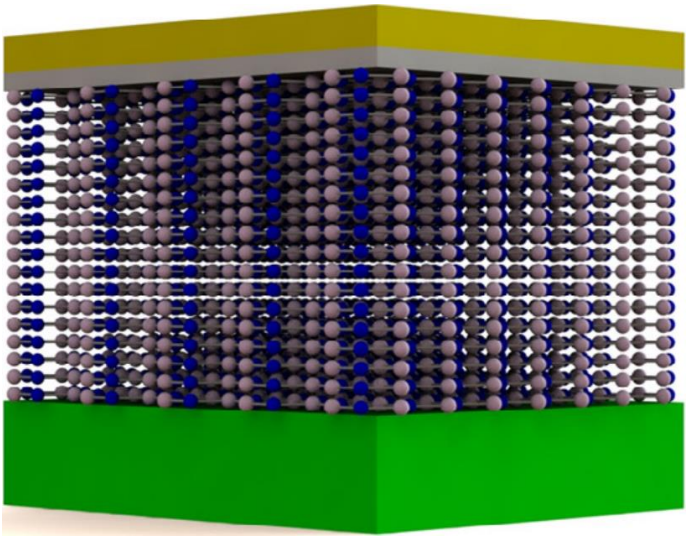


Figure 2.1 Schematic of a simple MIM sandwich structure

The observation of the memristive phenomenon can be traced back to two centuries ago, but it was not confirmed theoretically by Professor Leon Chua of the University of California, Berkeley until 1971 [18], and the combination of experiment and theory can be traced back to the memristor developed by HP laboratory in 2008 [19]. The most direct application of memristor is as resistive random access memory, but its dynamic nonlinear resistive process also implies the possibility that it can be used in neuromorphic computing. In 2000, IBM company used ternary transition metal oxide $SrZrO_3$ (perovskite material) to fabricate the memristor and provided a standard test process for its memory application direction: from current-voltage sweep to pulse, from endurance to retention and multilevel resistive state measurement, as well as the exploration of temperature dependence and scalability [20]. By 2004, Samsung officially determined the name of resistive random access memory (RRAM) using binary transition metal oxide nickel oxide and integrated it into a 0.18 μm complementary metal oxide semiconductor (CMOS) process with a one-transistor-one-memristor (1T1M) architecture [21-23]. Big institutes and companies like Imec [24-26], Intel [27], Nokia [28], TSMC [29-30], Fujitsu [31-32], SanDisk [33], Spansion [34], GlobalFoundries [35] and Institute of Microelectronics of the Chinese Academy of Sciences [36-37] all have works related. Companies have a great advantage that they can integrate the memristor with the commercial transistor to achieve better controllability. In the year 2019, Fujitsu Semiconductor released an RRAM chip with 8 Mbits and an endurance of 10^6 [32]. SanDisk released a 32 Gb RRAM test chip with 24 nm technology [33]. Institute of Microelectronics of the Chinese Academy of Sciences successfully integrated RRAM with 14 nm FinFET technology [37].

2.1.1 Basic device structures and 3D stackability

I have tried three different device structures depending on the research purposes and the materials used.

The first device structure is the common bottom electrode structure. This structure has a bottom electrode and RS medium all over the substrate. The top electrodes are identical dots which can be fabricated by a shadow mask, inkjet printing or through an anode aluminium oxide (AAO) template. The common bottom electrode has the advantages of easy fabrication and high cleanness. This structure is also useful to study

materials that cannot go through the lithography and lift-off process, for example, organics, perovskite and liquid phase exfoliated 2D materials. However, the device prepared by this method needs a large area to allow the engagement of the probe station tip, and the area is usually greater than 50 μm × 50 μm. Probe station tips have a minimum radius of about 5 μm [38], but mechanical wear will inevitably increase the contact area during use. The mechanical stress applied to the devices' active area could also affect the electrical properties [39]. (Detailed information about the effect of mechanical stress on memristors is shown in **Article S1**).

The second device structure is the crosspoint structure. The crosspoint structure can solve the problems of device miniaturization and unwanted mechanical stress. The device's active area is defined by the small crosspoint area while the probe station tip can be applied to the large metallic pads. With photolithography or direct laser writer, a device size of a few μm can be achieved. With electron beam lithography, hundreds or tens of nm can be achieved. However expensive lithography instruments are needed to pattern the electrodes. Also, the electrode fabricated by metal lift-off could have the problem of metal or polymer residues and tall electrode edges, which can cause device early failure, especially at the crosspoint edge region.

The last structure is the via-hole structure. Memristors are typically involved at the back end of line (BEOL) to CMOS technology, while the transistor part can be embedded under the surface [40]. This BEOL method is also friendly for researchers from academia since the device size is defined by the size of the via hole. Researchers can easily get devices with diameters in tens or hundreds of nm, even using normal photolithography with a few μm resolutions. Since high accuracy alignment, local etching, and chemical mechanical polishing (CMP) are necessary, the via-hole structure is normally provided by the foundries. The typical price for TSMC Multi-Project-Wafer (MPW) shared tapeouts of 180 nm technology node is \$1,250/mm^2 [41].

Last but not least, the crossbar structure is the ultimate structure for high-density storage with a few tens of Gb density or neuromorphic computing. The topology of the crossbar is particularly well suited to artificial neural networks. The three-dimensional stackability can further improve the storage density and reduce the storage cost per bit.

Some experiments will also adopt a planar structure, but this plane is mostly a supplementary experiment for convenient characterization of the resistive switching mechanism, rather than a characterization structure of device performance or a structure used in industry [42].

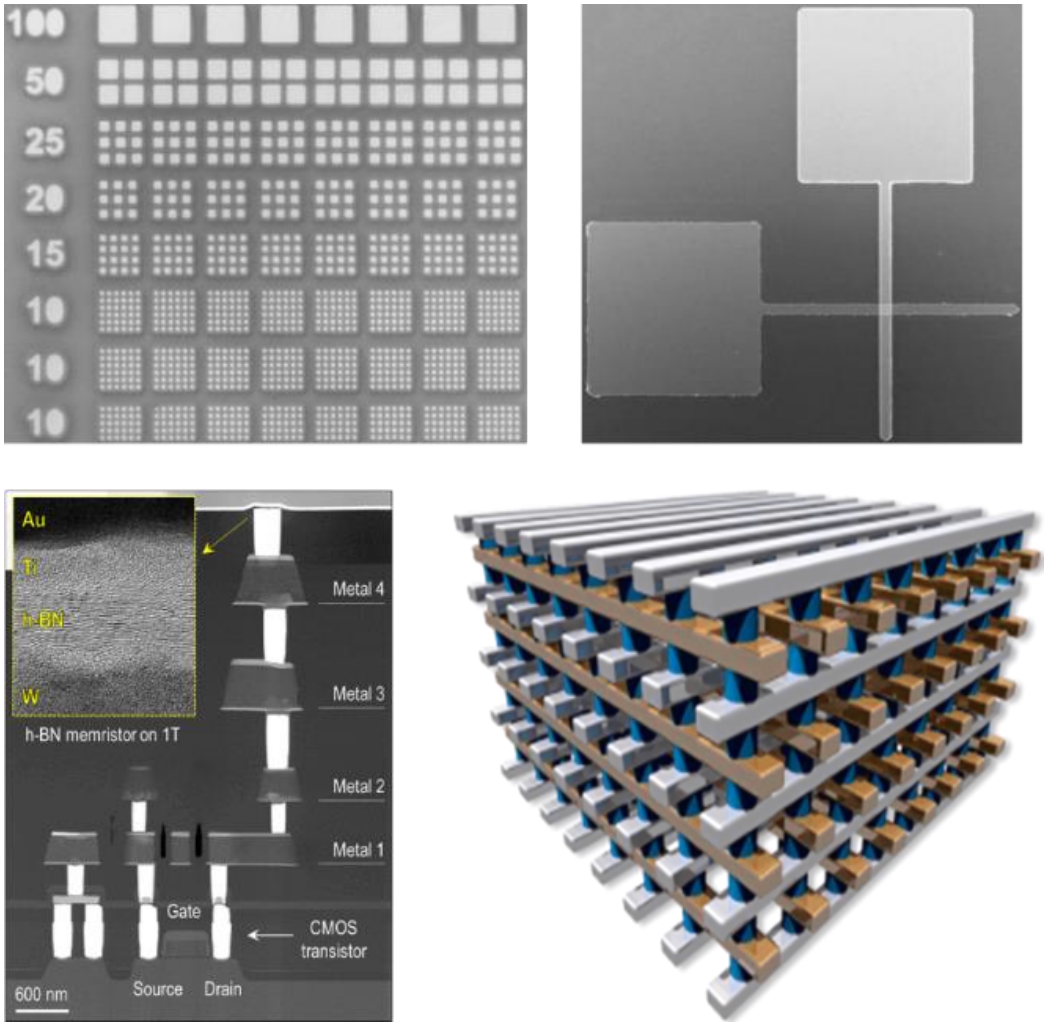


Figure 2.2 Basic cell structures including common bottom electrode, crosspoint, and via hole structure. 3D stackability of crossbar structure.

Table 2.1 Comparison of different device structures

	Common bottom electrode	**Crosspoint structure**	**Via hole structure**
Device area	Large, > 2,500 μm^2 Switching performance could be different when scaling down	Small, μm / nm Easy for probe station tip engagement High-density 3D	Small, nm scale
Equipment requirement	Easy fabrication Material friendly	Lithography needed	CMOS technology
Device quality	Clean Mechanical stress of tip	Polymer, metal residue high electrode edge	Good

2.1.1.1 Fabrication of a typical crosspoint structure

Crosspoint structure is the most commonly used structure in academia. Figure 2.3 shows the fabrication procedure of a typical crosspoint structured device.

Step 1: Pattern and deposit the bottom electrode. First, spin coat and bake the photoresist, which is used as a sacrifice layer for thin film patterning. Second, the light exposure through the pattern of the photomask. A post-exposure bake (PEB) is optional depending on the photoresist, although it may help to reduce the impact of the standing wave effect and make the chemical reaction more sufficient. Thirdly, the developer will remove the unwanted part of the pattern. For positive photoresist, the part exposed by the light will be dissolved in the developer. Appropriate concentration and developing time are needed to get the targeted resolution and sharp edge. In academic laboratories, contact photolithography technology commonly uses a Hg lamp (g, h, line) as a light source to fabricate a few μm patterns because of its low cost, wafer-scale productivity, and versatility for different photomasks.

Step 2: Deposit the dielectric film. Here the deposition methods could be thermal/electron beam evaporation, sputtering, atomic layer deposition, or transfer for 2D materials.

Step 3: Pattern and deposit the top electrode.

Here the pattern methods could be photolithography, direct laser writing or electron beam lithography.

Another photolithography and etching are optional to open a window on the bottom electrode if the dielectric film is covered all over the surface.

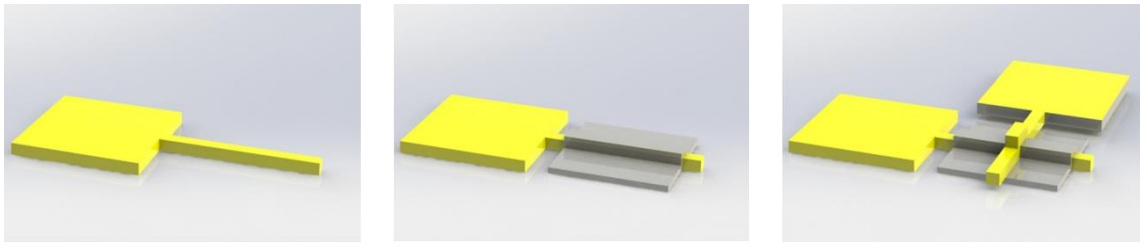


Figure 2.3 Schematics showing the fabrication of crosspoint structured device

2.1.1.2 Failure analysis of different device structures

For the common bottom electrode structure, since the physical gap between the target sample and shadow mask is inevitable. If the contact is not processed well, the pattern will be blurred and the device area will increase. For the mechanical stress problem, another protective metal is suggested to be deposited on top, like Pt.

For the crosspoint structure, the unwanted metal tall wall near the edge of the metal wire could cause earlier device failure. The irregular protuberance could cause the discontinuity of the dielectric grown on top. As the electrical breakdown tends to happen at the edge region, especially the bottom electrode edge. One thin layer of the polymer was spin-coated on top for better observation. For photolithography, negative photoresists and positive photomasks could help eliminate this problem.

For the via hole structure, it is important to have the correct etching recipe to expose the conductive via. Over etching will break the dielectric on top, especially layered structured 2D materials. Conformal deposition methods are needed.

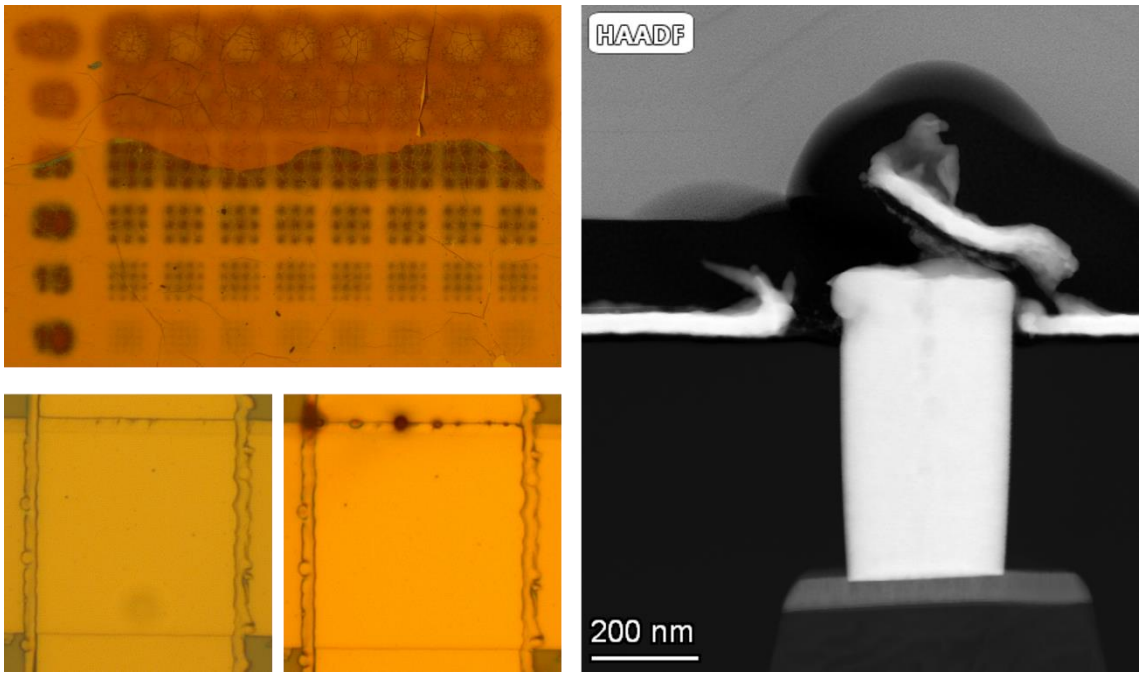


Figure 2.4 Failure analysis of different device structures.

For the device fabrication part, always check the detailed recipe parameter stored in the equipment in case some other people changed your recipe.

2.1.2 Electrical properties

Electronic measurements are the final test of electronic devices since the electronic properties are a determining factor in evaluating whether the electronic device works properly. Correctly and reasonably calibrate the instrument, set the test conditions, minimize interference and error, and find out the tips that can help to improve the accuracy and speed of the test. When investigating a new technology for electronic applications, it is of primary importance to define the key figures of merits (FoMs) and compare them against the requirements of the IEEE International Roadmap for Devices and Systems (IRDS).

Generally, the electrical properties of the memristor are measured by the probe station and the semiconductor parameter analyser. Figure 2.5 shows the image of a typical probe station and two commercial semiconductor parameter analysers. Middle: 4200 from Keithley. Right: B1500A from Agilent. The electrical stress is normally applied to the top electrode and the bottom electrode is grounded. Three main forcing functions could be applied to measure memristors electrically. With a single source-measure-unit (SMU), voltage stress like ramped voltage stress (RVS), constant voltage stress (CVS), and voltage list sweep stress (VLS) could be applied to get a current-voltage (I-V) plot, current-time (I-t) plot and current-pulse number plot, respectively. SMU can apply voltage and measure current simultaneously. Corresponding current stress could also be applied based on some specific applications. By adding the pulse slot (B1530A WGFMU for B1500A, 4225-PMU for 4200), we can generate and measure fast pulses with a minimum pulse width of 10 ns.

Figure 2.5 Images of probe station and semiconductor parameter analysers

Note: RVS and CVS are also known as voltage sweep and voltage bias mode. RVS is also referred to as DC mode, to compare with pulse mode. VLS works similar to pulse mode but cannot control the pulse width. 1 ms for 1 data point. VLS mode has the advantage of applying current compliance, while pulse mode cannot have current compliance. One resistor can be connected to reduce the current.

With RVS mode, we can collect the data to plot the most commonly used I-V curves for the memristor. I-V curves are helpful to obtain an overview of the characteristics (like unipolar or bipolar switching) and for determining the threshold voltages for the pulse operation. RVS mode only needs the most fundamental setup – 1 SMU and ground for the simplest configurations. For the RVS mode using voltage linear sweep, there are 3 main parameters to decide for the programming:

1) Start and stop voltage,

2) Voltage step,

3) Current compliance (CC).

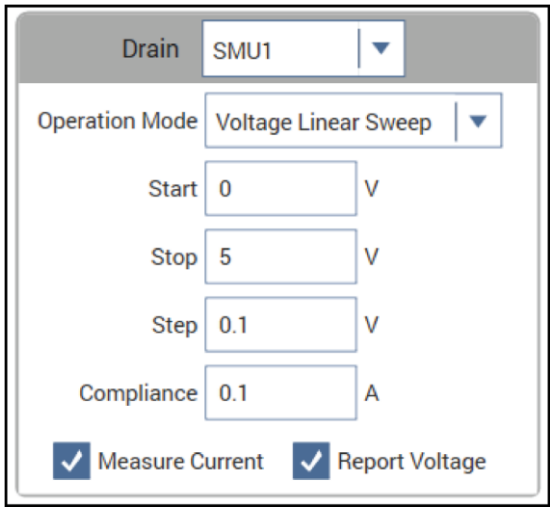


Figure 2.6 Parameter setting of the voltage linear sweep mode.

Since the typical memristor has a MIM structure, with an increased voltage linear sweep from start voltage to stop voltage, the devices always start from a high resistive state (HRS) with low current. Once the voltage reaches a threshold voltage, the current will show a sudden jump to low resistive state (LRS). This resistance switching from HRS to LRS is called the SET process, and the threshold voltage is called set voltage (V_{SET}), indicated by arrow 1 in Figure 2.7. In the process of SET, a limiting current, named current compliance (CC), is used to constrain the maximum current flow in the device to prevent the device from receiving too much energy and being permanently broken down. Depending on the voltage polarity to reset the device back to HRS, the switching mode can be categorised as bipolar and unipolar switching. Unipolar memristor tends to have a higher operation current especially in the reset region, with a current level of more than 1 mA.

And the threshold voltage at which point HRS turns to LRS is called reset voltage (V_{RESET}). If the LRS can go back to HRS automatically without applying any stress, this behaviour is named threshold switching.

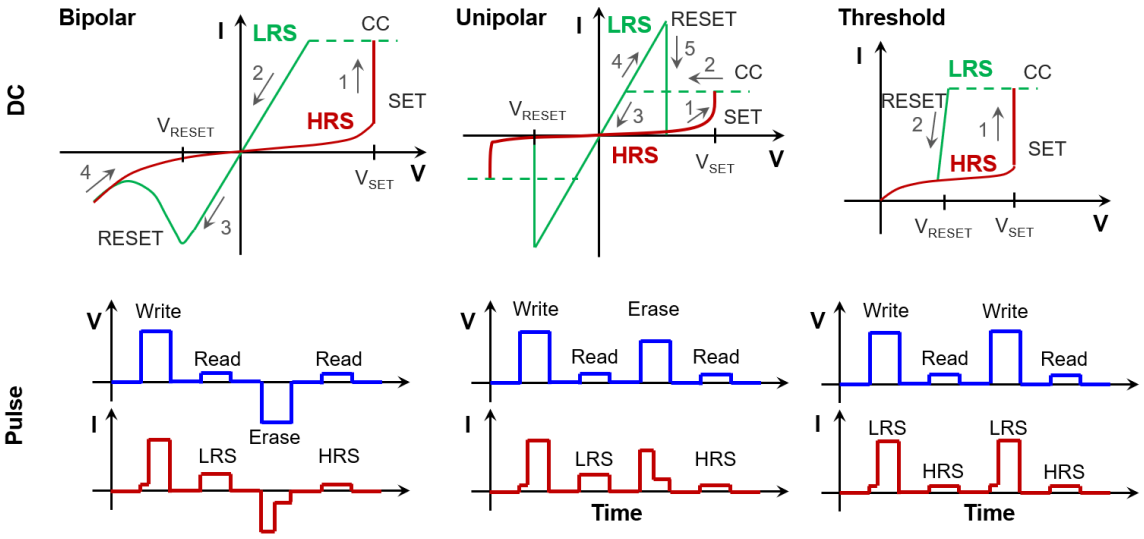
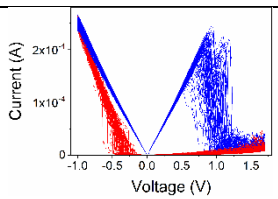
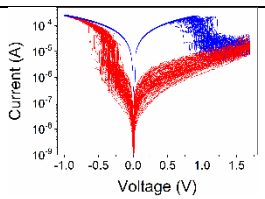
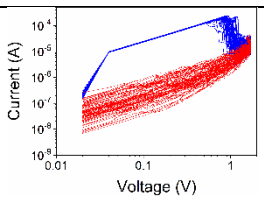
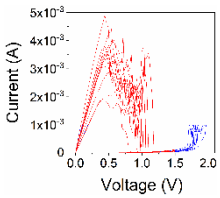
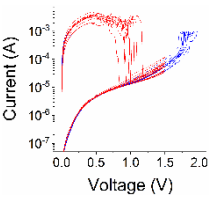
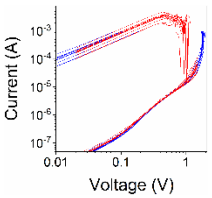
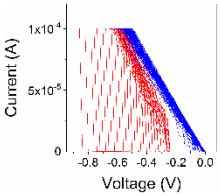
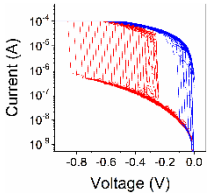


Figure 2.7 Typical unipolar, bipolar and threshold switching under dc and pulse stresses

Linear plots are used to check the Ohmic conduction in the LRS, while semi-log plots are used to check on-off ratio. Log-log plots are used to check the conduction mode in HRS as shown in Figure 2.8.

Table 2.2 Different plotting methods for I-V curves

	Linear plot	Semi-log plot	Log-log plot
Bipolar			
Unipolar			
Threshold			

Pulse mode works similar to the clock signal in the real computer with a pulse shorter than 100 ns (Fujitsu RRAM works with a frequency of 10 MHz [32]), while I-V sweeps do not match the operating conditions of realistic devices. Pulse mode works with minimized energy, which limits the catastrophic damage to the device.

For pulse measurement, 2 waveform-generator-fast-measurement-unit (WGFMU) are needed. One is used to apply pulse stress, and the other one is used to measure current. There are 3 main parameters to decide to program the pulses of WGFMU1:

1) Pulse width (t_W) and height,

2) Time between pulses.

3) Rise/fall time.

Pulse width and height need to be carefully adjusted based on the electrical performance of the device. Time between pulses are normally used to characterize synaptic behaviour like potentiation and depression.

WGFMU2 share the same total pulse time with voltage always 0V. There are 2 main parameters to decide for the measurement event of currents. 1) Measurement points, 2) Current range. More measurement points are useful to obtain the information of t_{SET}/t_{RESET}, while minimized measurement points are useful to obtain endurance.

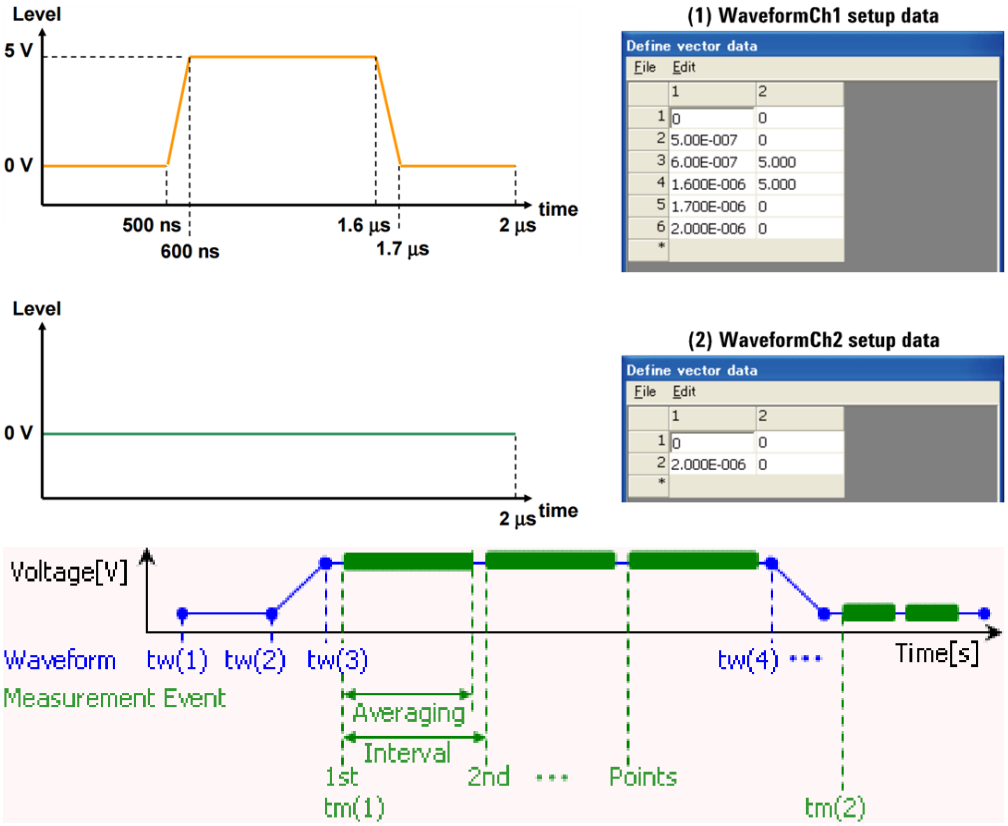


Figure 2.8 Parameter setting of pulse programming and measuring

There are 3 main parameters to decide for the I-t measurement.

1) the value of the constant voltage applied,
2) the measurement speed. The speed is defined by the delay factor, filter factor and A/D integration time.
3) the number of measurement points.

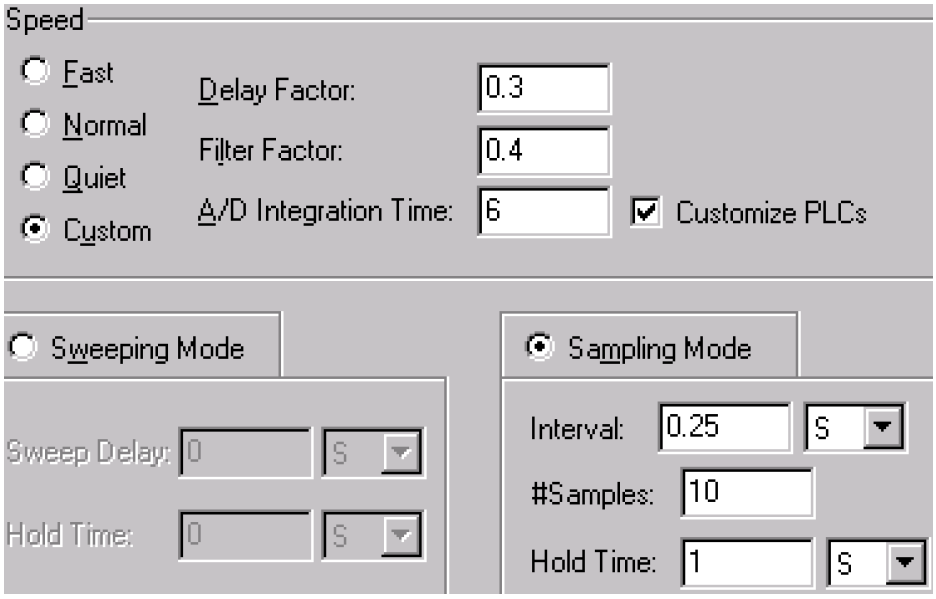


Figure 2.9 Parameter settings for I-t measurement

A small value of constant voltage with a slow measurement speed is used to measure the retention time of the memristor's different resistive states. Here the constant voltage is also named the read voltage.

A small value of constant voltage with a fast measurement speed is used to measure the random telegraph noise (RTN). (Detailed information about the RTN measurement on both h-BN and metal-oxide based memristors are shown in **Article S2-3**).

A constant voltage with a value larger than the read voltage but smaller than the breakdown voltage is used to measure the time-dependent-dielectric-breakdown (TDDB). (Detailed information about the TDDB measurement are shown in **Article S4**).

Table 2.3 FoMs from different stress mode to measure memristor

	I-V	Pulse	I-t
Stress applied	Ramp voltage stress	Pulse voltage	Constant voltage stress
Measured value	Current	Current, time	Current
FoMs extracted	Distribution of V_{SET} and V_{RESET} Endurance plot with on-off ratio window	Switching time: t_{SET} and t_{RESET} Endurance plot with on-off ratio window	Small V_{READ} with long time interval → Measuring retention Small V_{READ} with short time interval → Measuring RTN Voltage larger than V_{READ} but small than V_{BD} → Measuring TDDB
Advantages		1- reduce the total energy consumption of devices; By reducing the Joule heating effect, damage to small nano devices can be avoided. 2- transients can be studied 3- generate clock signals for digital circuits, analog repetitive control lines, such as memory read cycles.	
Problem		Need to be connect with a resistor avoid permanent short circuit	

Table 2.4 Traditional and emerging memory technologies [74-75]

	Traditional Technologies				Emerging Technologies			
	DRAM	SRAM	Improved Flash		FeRAM	MRAM	PCRAM	Memristor
			NOR	NAND				
Knowledge level	Mature		Advanced		Product		Advanced	Early stage
Cell elements	1T1C	6T	1T		1T1C	1T1R	1T1R	1M
Half pitch (F) (nm)	50	65	90	90	180	130	65	3-10
Read time (ns)	< 1	< 0.3	< 10	< 50	< 45	< 20	< 60	< 50
Write/erase time (ns)	< 0.5	< 0.3	10^5	10^6	10	20	60	< 250
Retention time (years)	seconds	N/A	> 10	> 10	> 10	> 10	> 10	> 10
Write op. voltage (V)	2.5	1	12	15	0.9~3.3	1.5	3	< 3
Read op. voltage (V)	1.8	1	2	2	0.9~3.3	1.5	3	< 3
Write endurance	10^{16}	10^{16}	10^5	10^5	10^{14}	10^{16}	10^9	10^{15}
Write energy (fJ/bit)	5	0.7	10	10	30	1.5×10^5	6×10^3	< 50
Density (Gbit/cm^2)	6.67	0.17	1.23	2.47	0.14	0.13	1.48	250
Voltage scaling	Fairly scalable					No	Poor	Promising
Scalability	Major technological barriers				Poor		Promising	Promising

2.1.3 Material combination and switching mechanism

The commercialization of any application requires a solid and predictable understanding of its underlying mechanisms. The switching mechanism of memristors is based on the material combination of the electrodes and the switching medium. Figure 2.10 shows two typical material combinations with different switching mechanisms. Metal 1 is the adhesion layer. Metal 2 and 4 are the inert bottom electrode and protective electrodes. The insulator is the resistive switching medium. The switching mechanism depends on the section of Metal 3 and the optional insulator 2.

Conductive bridge random access memory (CBRAM) tends to higher off-ratio. CBRAM is based on the creation and formation of conductive nano-filaments (CNF) made of active metal penetration (Metal 3). Metal 3 is an active electrode which will participate in the resistive switching.

Oxide resistive random access memory (OxRAM) tends to have multilevel switching capability for analogue applications. Insulator 2 is optional and can work as an oxygen reservoir. Bilayer oxides are used for better performance. [43]

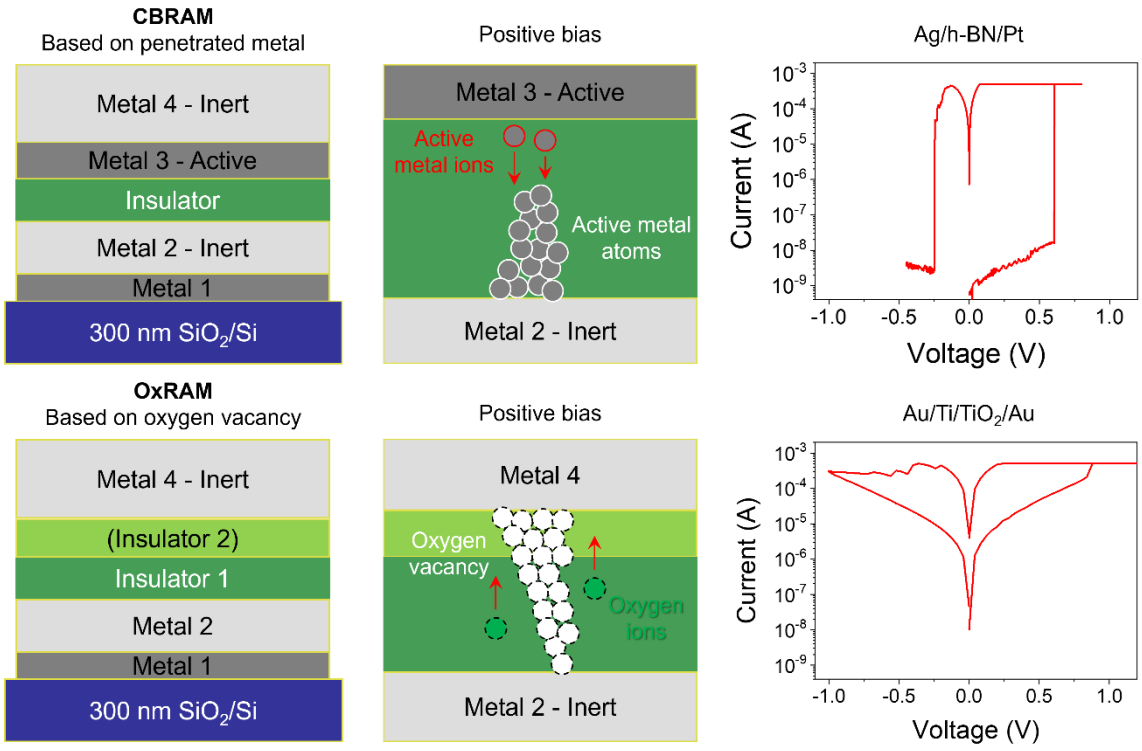


Figure 2.10 Material combinations and corresponding switching mechanism, I-V curves

2.1.3.1 Selection of substrates and adhesion layers

For memristors, we want the current always flows in the MIM structure. So the substrate should be a highly resistive substance to avoid unnecessary leakage current. 300 nm SiO_2/Si substrates are highly recommended, while glass and ITO substrates are used for transparent devices, and plastic or paper are used for flexible devices. Inert metals like Pt and Au have poor adhesion with the substrate, especially when they are deposited by thermal or electron beam evaporation. A thin adhesion layer like Ti, Tantalum (Ta) [44] or chromium (Cr) with a thickness of 5 nm is enough to provide necessary adhesion.

2.1.3.2 Selection of bottom and top electrodes

In traditional devices, the electrode mainly plays the role of transmitting electrons, while in memristor devices, the electrode can affect the transition mechanism of the device. For memristors, bottom electrodes are normally made of inert electrodes, including tungsten (W), Pt, Au, graphene, silicon-based electrodes such as highly doped polysilicon, nitrogen-based electrodes such as titanium nitride (TIN), and oxygen-based electrodes such as indium tin oxide (ITO). Highly conductive Si or active metals can also be used as bottom electrodes; however, inert electrodes provide a more reliable switching performance with less variability.

Top electrodes can be inert or active materials. Active electrode materials include metal electrodes such as Cu, Ag, Ti and Ni, under such a combination, the main resistance mechanism lies in the dissolution of the active metal electrode under positive pressure and the deposition of nano conductive filaments in the dielectric layer, as well as the fracture of conductive filaments under directional voltage or high current.

The insertion of Ti, hafnium (Hf), Zirconium (Zr), and Ta buffer layers over HfOx dielectrics show improved resistive switching performance. The thicker Ti layer, the thicker interface thickness, which will decrease the forming voltage of ZrO_2 based memristors [45]. Ti contributes to the low power memristors. Because the insertion acts as an oxygen-gettering layer, which acquiesces to control the dielectric strength of the HfOx dielectrics, the forming voltage, leakage current and on-off ratio [46]. Cu and Ag tend to be CBRAM with a large on-off ratio. Ag tends to be a threshold resistive switching. For memristors using active metals, a protective inert electrode is necessary to prevent oxidation or scratch.

Sputtered Au electrode (with the same cross-sectional area and line length) has a lower resistivity and better high current durability than e-beam evaporated electrode, indicating it is a better method for much more scaled device with a line width of few or tens of nanometers considering from the aspect of the electrical property [47]. (Detailed information about the effect of metal deposition methods on memristors is shown in **Article S6**).

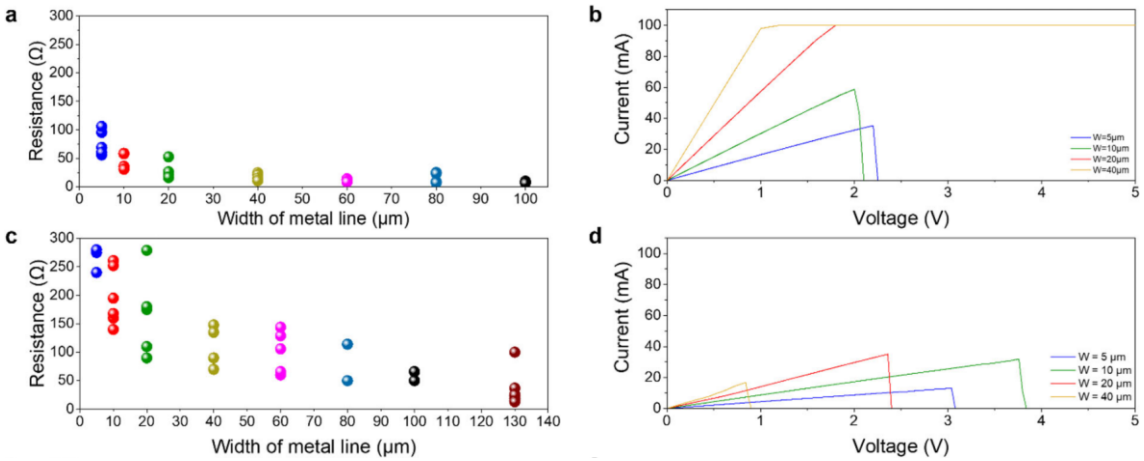


Figure 2.11 Comparison of different interconnection technologies

Electron beam deposited Au could induce distortion and defects at the interfaces between the h-BN and Au. (Detailed information about the effect of metal deposition methods on 2D materials is shown in **Article S7**). This may affect the transistor quality when 2D materials are used as the channel material, while these defects could improve the memristor behaviour when h-BN is used as the resistive switching medium.

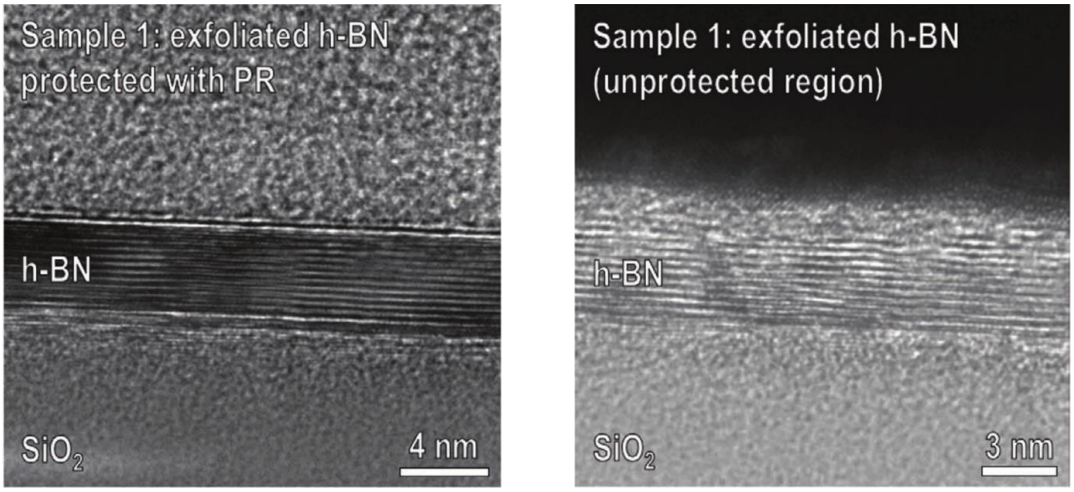


Figure 2.12 XTEM of h-BN with and without protection under Au deposition

Table 2.4 lists the advantages and disadvantages of several mainstream electrode materials.

Table 2.5 Selection of metallic electrodes

	Advantages	Disadvantages
Ti	As adhesion layer Strong oxygen absorption capacity as an oxygen reservoir [48] Lower V_{SET} by forming a strong Ti-O bond [49-50] Lower operation current	A high vacuum is needed. High melting point (not for thermal evaporation)
Ta	As an oxygen ion getting layer Promising in analogue switching [51]	
Ag	High ionic mobility, surface activation energy, for threshold switching [38,52] High switching speed	Not for long-term non-volatile memory
Cu	High ionic mobility [53] For Threshold switching [52]	
Ni	High resistivity, even no need CC [54] RTN signal [55-56]	Not for high-density crossbar array
Au	Chemically inert but has immigration [57]	Not suggested for CMOS [58] High operation current
Pt	Inert metal, no ion immigration [59] For analogue switching	Hard to etch
TiN	Conductive nitrides Inert from an aspect of CBRAM [62-63] CMOS compatible	
Gr	As ion blocking layer Lower operation current Lower power consumption	high-temperature direct growth or transfer needed

2.1.3.3 Selection of resistive switching medium

Thin films have several crystal structures: amorphous, single crystalline and polycrystalline, while polycrystalline dielectrics show the best resistive switching behaviour. 2D materials based memristors provide the best demonstration. Mechanical exfoliation provides the perfect single crystalline materials. If we check the papers about memristors based on mechanical exfoliated 2D materials.

Table 2.6 Selection of dielectrics

	Characteristics
HfO_2	Initially as high k dielectric for high-performance CMOS
	Defect rich good for resistive switching [24]
	Both unipolar and bipolar switching
TiO_x	Several stable phases, like oxygen deficient Magneli phase, Ti_4O_7. [64]
	Good for multilevel and analog switching
	Low endurance and short retention time
	Promising for high endurance 10^{12} [16]
TaOx	XPS and EELS reveal its switching mechanism is due to the shift between the non-conductive Ta_2O_5 phase and conductive TaO_2 [17]
SiOx	Based on highly mature Si technology [65-66]
NiO	Reliable unipolar switching with 1D1M architecture [21]
ZrOx	Doped ZrOx has high oxygen diffusivity and high formation energy [23]
Al_2O_3	Low reset current of 1 μA [67]
VOx	Threshold switching [68]
	Bipolar [69]
NbOx	Threshold switching 1081K [70]
	Analog switching for neuromorphic computing [71]
Perovskite	Light operation
	Cannot be patterned by photolithography.
	Flexible
Organics	Based on the materials used [72]
	Most organics cannot be patterned by photolithography

Atomic layer deposition (ALD) has become a popular choice owing to its excellent controllability and reproducibility down to the atomic layer level. In the processes, the deposition cycle is a key factor for determining resultant film thickness and other properties. Sputtering is an efficient method to synthesize uniform, stable, highly crystalline and stoichiometric TiO_2 film yielding reproducible and endurable switching behaviour. In these processes, an ion beam is bombarded on a target material in a vacuum maintaining the RF power [65].

With the same Au/TiO_2/Au structure but different deposition methods for the dielectrics, devices fabricated by sputtered TiO_2 show the typical bipolar switching behaviour, while devices fabricated by ALD only show a permanent breakdown and stay in LRS. It can be explained that ALD-grown TiO_2 films seem to be more compact and freer of defects than the sputtered ones, as the TEM contrast of the ALD thin film is much more homogeneous than the sputtered one. The irreversible LRS in ALD thin film is induced at the high defect density region. (Detailed information about the effect of the dielectric deposition methods on memristors is shown in **Article S5**)

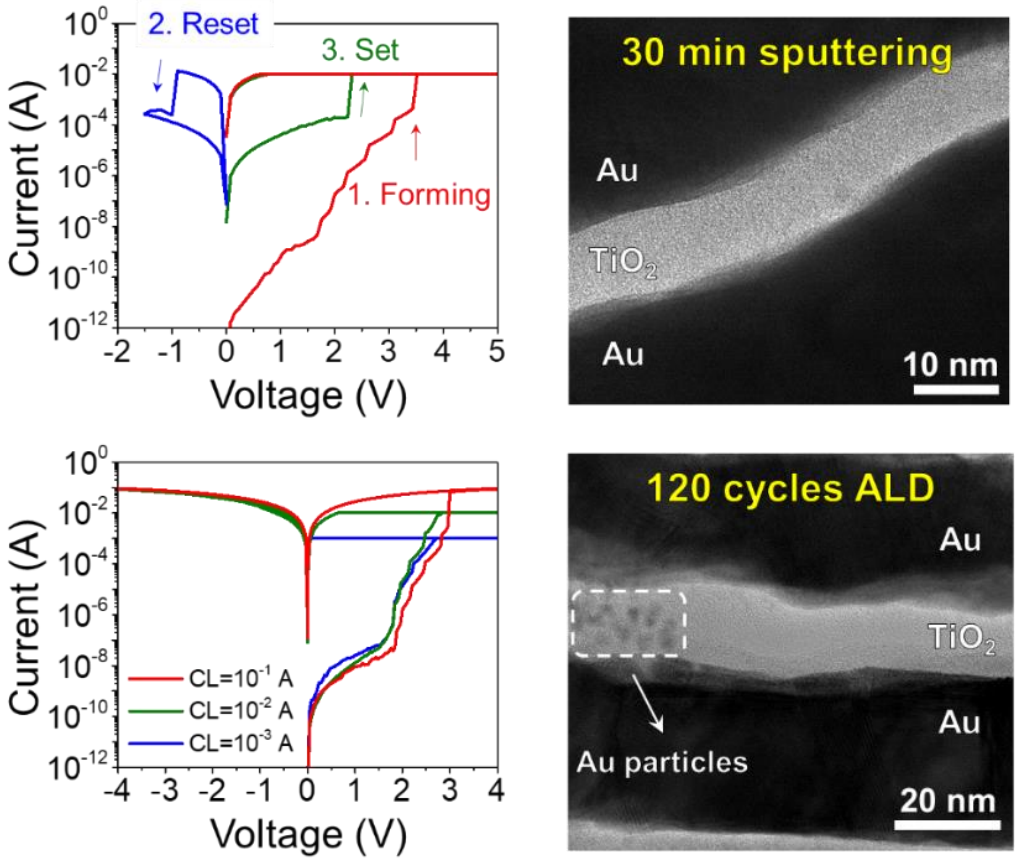


Figure 2.13 RS in Au/TiO_2/Au device with TiO_2 deposited by sputtering or ALD

2.1.4 Large crossbar circuit implementation: active or passive array

Large memristor circuits are normally structured as a crossbar matrix. The row of the matrix is normally named as a word line while the column is named as bit line. Each operation cell is defined by the crosspoint structured word line (top electrode) and bit line (bottom electrode). The peripheral circuit controls the word line and bit line to conduct write, erase or read operations on arbitrary selected cells.

In the passive array, each cell is one memristor, which ensures this structure has the minimum cell size of $4F^2$ (F is the feature size). Considering the memristor's 3D stackability, the effective cell area can be further decreased to $4F^2/N$ (N is the number of stacked layers). However, the passive array normally has an Ohmic conductive LRS, which could cause the misread of adjacent devices due to sneak path currently. To avoid sneak path current, the half-select method is used to alleviate the problem. Different kinds of selectors and nonlinear memristors are developed. Starting from 2012, nonlinear memristors became a hot topic for the implementation of large crossbar circuits. In 2012, J. J. Yang et al. demonstrate a nonlinear device with TiO_{2-x}/TaO_x oxide heterostructure. One immediate advantage of a nonlinear device is that it will decrease the current level, especially the LRS current.

Among different large crossbar circuit implementation methods, the active 1T1M array is more attractive due to this cell structure can: 1) accept both bipolar and unipolar switching; 2) avoid current overshoot. Current overshoots will generate excessive defects that the conductive filaments tend to grow laterally and increase in diameter or even multiple conductive filaments can be generated. 3) Lower the operation current and power with transistor constrain. 4) have better controllability to achieve multilevel switching by controlling the gate voltage. 1T1M cell size depends on the size of the transistor. With CMOS technology, the transistor and the memristor can be vertically stacked to relieve the problem. If the transistor is based on bulk silicon, it will have the problem of 3D integration.

Compared to 1T1M, 1D1M has better stackability. However, 1D1M cells only accept memristors with unipolar switching behaviour.

Among the 26 papers from the big companies and institutes [20-37] or papers published on Nature or Science [76], 16 papers used 1T1M cell, 2 papers used 1D1M cell, and 4 papers used 1R cell with the non-linear switching device.

2.2 More Moore using new materials - 2D materials

2D materials refer to the materials with crystalline layered structures consisting of monolayer atoms, which are connected by strong covalent bonds in the plane and weak van der Waals forces outside the plane. Materials with polycrystalline structures and a few layers can also be categorized as 2D materials. Monolayer 2D materials can be obtained by breaking the relatively weak van der Waals force. For example, the most representative 2D material, graphene, can be obtained by exfoliating the tape from the bulk graphite; At the same time, the graphene can be transferred to any targeted substrate.

Graphene was first exfoliated in 2004 [81] and the "groundbreaking experiments regarding the two-dimensional material graphene" makes the discoverers Andre Geim and Konstantin Vovoselov win the Nobel Prize in 2010 [82]. 2D materials represented by graphene show its great potential to extend Moore's law. This will not just be a replacement between materials, but an iteration of the whole material system.

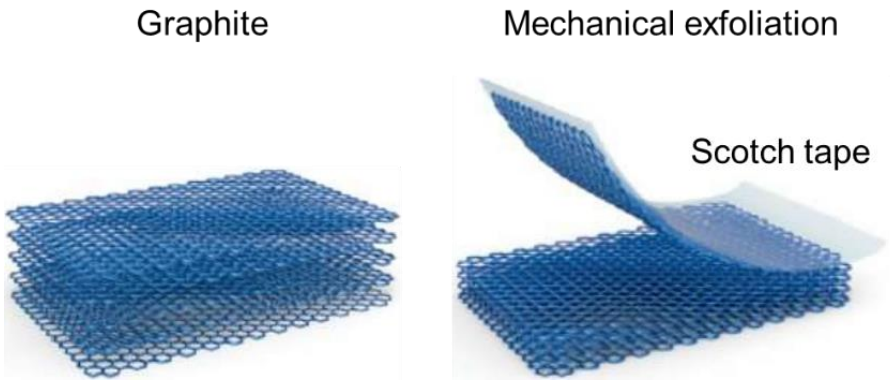


Figure 2.14 Exfoliation of monolayer graphene from the bulk graphite

Reducing the dimension of a system usually can change or even obtain extraordinary mechanical, electrical, optical or magnetic properties. Graphene, as the representative of crystal materials with only one atomic layer or a few atomic layers, shows rich, novel and interesting physical phenomena. Its excellent physical and chemical properties open a broad market for its application prospects. In the past decade or two, 2D materials have been proved to show unparalleled potential in many emerging applications, including logic devices [83-84], nonvolatile memory [85-86], energy storage [87], sensors [88] and flexible [89-90] electronic devices.

The use of semiconducting 2D materials such as molybdenum disulfide (MoS_2) and tungsten diselenide (WSe_2) as channel materials in field-effect transistors (FETs) could provide improved electrostatic control because the channel depth would be very low (few angstroms) and accurately known, leading to less vertical leakage current towards the substrate, less horizontal phonon scattering and reduced short channel effects. In the newly updated 2021 International Roadmap for Devices and Systems (IRDS) [4], 2D materials can be used as channel materials from 2028, and beyond-CMOS in the form of 2D devices.

YEAR OF PRODUCTION	2021	2022	2025	2028	2031	2034
	G51M30	G48M24	G45M20	G42M16	G40M16/T2	G38M16/T4
Logic industry "Node Range" Labeling (nm)	"5"	"3"	"2.1"	"1.5"	"1.0 eq"	"0.7 eq"
IDM-Foundry node labeling	i7-f5	i5-f3	i3-f2.1	i2.1-f1.5	i1.5e-f1.0e	i1.0e-f0.7e
Logic device structure options	FinFET	finFET LGAA	LGAA	LGAA	LGAA-3D	LGAA-3D
Platform device for logic	finFET	finFET	LGAA	LGAA	LGAA-3D	LGAA-3D
Frequency scaling - node-to-node	-	0.02	0.16	0.09	-0.08	-0.01
CPU frequency at constant power density (GHz)	3.13	2.83	3.53	2.50	1.48	0.86
Power at iso frequency - node-to-node	-	-0.16	-0.27	-0.05	-0.06	-0.08
Power density - relative	1.00	1.12	1.04	1.59	2.51	4.27
LOGIC TECHNOLOGY ANCHORS						
Patterning technology inflection for Mx interconnect	193i, EUV DP	193i, EUV DP	193i, EUV DP	193i, High-NA EUV	193i, High-NA EUV	193i, High-NA EUV
Beyond-CMOS as complimentary to platform CMOS	-	-	-	2D Device, FeFET	2D Device, FeFET	2D Device, FeFET
Channel material technology inflection	SiGe25%	SiGe50%	SiGe50%	Ge, 2D Mat	Ge, 2D Mat	Ge, 2D Mat
Process technology inflection	Conformal Doping, Contact	Channel, RMG	Lateral/AtomicEtch	Non-Cu Mx	3DVLSI	3DVLSI
Stacking generation inflection	2D	3D-stacking: W2W, D2W Mem-on-Logic	3D-stacking: W2W, D2W Mem-on-Logic	3D-stacking, Fine-pitch stacking, P-over-N, Mem-on-Logic	3D-stacking, 3DVLSI: Mem-on-Logic with Interconnect	3D-stacking, 3DVLSI: Logic-on-Logic

Figure 2.15 2D materials in the 2021 IRDS

2.2.1 Evaluation of typical 2D materials

The most important characteristic of 2D materials is the layered structure, which is normally characterized by the cross-sectional TEM images. The second important property is based on its crystalline structure. Monolayer single crystalline MoS_2 is highly recommended due to its bandgap as channel material. Polycrystalline h-BN is highly recommended as a resistive medium due to its current confinement. Here we show 3 typical materials based on their conductivity or bandgap.

The first category is conductors including graphene and MXene, which can be used as electrodes. 30-inch graphene thin films have been grown and can be transferred by the roll-to-roll method as transparent electrodes. It shows better performance compared with commercial indium tin oxides based transparent electrodes [91]. Graphene-based interconnect has been integrated into 7 nm FinFET technology and

achieves better resistance to electromigration and high current density compared to the conventional Cu interconnect [92]. The graphene/Cu combination can improve the conductivity and high current density endurance [93]. Wafer scalable MXene (Ti_3C_2) could be used as a gate, source, and drain electrode of transistors [94], and electrode array for flexible electronics [95].

The second category is a semiconductor which includes the transition metal dichalcogenides, III-VI family and black phosphates family. The first MoS_2 transistor was fabricated back in 2011 and the material is mechanically exfoliated [96]. Wafer-scale growth is achieved by chemical vapour deposition (CVD) [97] and liquid phase exfoliation (LPE) [98]. CVD provides a statistical analysis of 200 devices to extract the variability and show the mobility in the range of 30~45 $cm^2/V/S$ [99]. Big companies like Imec [100-108], Intel [109-110], TSMC [111-112], and Samsung [113-115] all have works related. Imec focused on the wafer scale growth and transfer of MoS_2 and WS_2. Based on the wafer-scaled materials, a huge amount of transistors of 15,000 devices are statistically analysed. Intel focused on the replacement of channel materials of FEOL by engineering the NMOS and PMOS based on MoS_2 and WSe_2 to achieve ultimate CMOS scaling, respectively. TSMC focused on ultralow resistive, metallic contact engineering in cooperation with MIT.

The last category is insulators mainly h-BN. h-BN thin sheets were firstly exfoliated in 2008 [116], however, the monolayer h-BN was prepared by a combination with reactive ion etching [117] or a high-energy electron beam [118] due to the strong interplane bonding compared to graphene or MoS_2. h-BN is also widely studied as the candidate for future gate dielectrics or substrates of graphene-based electronics. h-BN supported graphene devices provide an order of magnitude better mobility and two orders of magnitude smaller charge fluctuations than the same devices on SiO_2 substrate [119-120]. Conductive atomic force microscopy (CAFM) is a perfect tool to study its electrical properties from the material level to the nanoscale. CAFM I-V sweeps show the direct tunnelling in few-layer h-BN (1~3 layer), and Fowler-Nordheim tunnelling for thicker h-BN with more than 4 layers. CAFM confirms the exfoliated h-BN is a defect-free dielectric [121] with a high barrier height of 3.07 eV and dielectric strength of 7.94 MV/cm [122]. CAFM mapping demonstrates the electrical uniformity of exfoliated h-BN in a more visual way. These values are comparable with the traditional dielectric SiO_2. J. Kong group at MIT used low-pressure CVD with ammonia borane as the precursor to grow single layer h-BN, and used atmospheric pressure CVD with

borazine precursor to grow multilayer h-BN on Cu foil [123-124]. These two methods are adopted as academic references by a commercial company “Graphene Supermarket” to grow single-layer and multi-layer h-BN, respectively [125-126]. From 2018 to 2022, Nature or Science published 4 papers regarding the wafer-scale methods to grow h-BN using different substrates and mechanisms [127-130], demonstrating its potential.

2.2.2 Scalable synthesis of 2D materials

It is widely used in cutting-edge research in physics, physical properties and devices. Most of the intrinsic properties of graphene, such as quantum Hall effect, massless Dirac fermion gas and superconductivity, are obtained on stripped high-quality samples. When using CVD, vacancies, impurities, atomic misalignments, strained bonding, wrinkles and thickness fluctuations in the 2D sheet can easily appear on substrate imperfections (i.e., grain boundaries, steps, impurities); moreover, the CVD process often results in polycrystalline 2D sheets, which intrinsically contain many vacancies and strained bonds. However, such pilot demonstrations to produce single-crystal 2D monolayers via CVD still did not reach the commercial stage, and state-of-the-art samples still contain grain boundaries, as well as other imperfections like multilayer islands and wrinkles [78]. Moreover, these approaches to producing single-crystal 2D materials via CVD have been only demonstrated for monolayers, which are not suitable for many applications. Another important challenge to address when using the CVD approach is the thermal budget. CVD growth of 2D materials requires the use of high temperatures >900 ºC to achieve good crystallinity, but the maximum temperatures allowed during and BEOL processes are ~450 ºC [19].

When using LPE-based 2D materials to construct electronic devices and circuits, the main problem is that the junctions between different flakes contain multiple bonding defects that can severely degrade the performances, such as room temperature mobility in graphene (i.e., ~140,000 cm^2/Vs in mechanically exfoliated [87] vs. 95 cm^2/Vs in LPE [88]) and leakage current in h-BN (i.e., in this case, comparisons are not so straightforward because the thinnest LPE h-BN films reported are ~100 nm thick, which show ~10-6 A/cm^2 at 1V [89], and for such thickness exfoliated h-BN shows no leakage current at 1 V). Furthermore, the minimum thickness and surface roughness of 2D materials films made by LPE (i.e., ~10 nm and ~3 nm, respectively) [56,90] are much higher than that of CVD sheets (i.e., monolayer and <150 pm, respectively) [91]. This

impedes the use of LPE 2D materials in some applications, such as gate dielectric in ultra-scaled FETs [92]. Furthermore, so far the device-to-device variability of devices made of 2D materials produced by LPE is higher than CVD-grown counterparts, as one device that is formed exactly on a flake junction is expected to behave very different than others that are junctions-free. In fact, the smallest LPE-based FET that we could find in the literature has L = 50 μm [56], while using the CVD method FETs with L = 8.2 nm have been readily achieved [93].

It is important to understand that, while state-of-the-art 2D materials synthesized by CVD and LPE (i.e., which contain native defects) still cannot be employed reliably for some high integration density applications (i.e., gate dielectric in FETs), the inherent presence of defects may make them suitable for others, such as resistive switching technologies [64,94].

Table 2.7 compares the advantages and disadvantages of the mainstream growth methods for 2D materials. CVD is highly recommended for large-scale IC integration.

Table 2.7 Comparison of different growth methods of 2D materials

	Working principle	**Crystal status**	**Size**	**Yield**	**Application**
Mechanical exfoliation	Mechanical stress from weak interlayer van der Waals force	Perfect layered structure	< 50 μm	Low efficiency, high randomness, low yield	For fundamental physics
LPE	Gentler shear stress applied by ultrasonication	Perfect layered structure		Efficient and low-cost	For large area and flexible electronics
CVD	Surface reaction	Polycrystalline (single crystal hundred of μm)	4 inch	High efficiency	For ICs
PVD	Sputtering			High efficiency	
Epitaxial growth	Thermal decomposition of SiC	High quality	μm	Low temperature 400 ºC. High speed 1 Å/s	For Physics, quantum metrology

2.2.3 Performance evaluations

The quality evaluation for 2D materials is more important than growing one batch of 2D materials. We will judge the quality of 2D materials from two main aspects 1) from the material level, and 2) from the device or circuit level.

From a materials point of view, claims about high-quality, single-crystallinity and low density of defects in 2D materials rely on one/few transmission electron microscopy (TEM) image/s and/or macroscale measurements like Raman spectroscopy and X-ray photoelectron spectroscopy (XPS). However, TEM covers a very small part of the samples, Raman and XPS have a very low lateral resolution, and none of them provides information about the electrical properties of the materials. The electrical performance of 2D materials has been often evaluated at the device level using electrical parameters like in-plane carriers mobility in FETs [56] (not suitable for dielectrics) or out-of-plane leakage current in large ($>10^4$ μm^2) metal/insulator/metal (MIM) structures [22]; however, these collect the total current driven by the area of the device under test, and they are blind to inhomogeneities produced by local defects. In this context, nanoelectronic characterization using atomic force microscopy (AFM) based techniques emerges as a powerful and simple method to evaluate the quality of 2D materials [120] statistically. For example, most reports measure the thickness of a 2D material showing a single-line cross-section of an AFM map at the edge, while they should show the spectra of the height of the whole scan to get statistical validation of the thickness and surface roughness [121] — this is even more important in ultra-thin (<4 nm) thick 2D materials grown by CVD, as the surface roughness can be comparable to the thickness of the material. Such kind of good practices should be adopted in all studies to provide statistical validation of their claims.

From a device and circuit perspective, performance evaluation requires abundant statistical information about yield, variability, reliability and stability [46]. Clear definitions of the yield criteria and acceptable variability windows should be provided, and they should match as much as possible the industrial requirements. For example, multiple articles have reported proof-of-concept demonstrations of RS in 2D materials, but most of them did not provide any competitive advantage compared to state-of-the-art RS technologies. Similarly, future studies on 2D materials-based devices and ICs should include deeper reliability analyses and lifetime predictions based on industrial characterization methods. Similarly, most claims about the endurance of 2D

materials-based RS devices have been based on current vs. cycle plots in only one device and showing just one/few cycles per decade [61], which overestimates the real endurance and ignores cycle-to-cycle and device-to-device variability. The correct method to evaluate endurance was used in Ref. [33], which plotted the current vs. cycle plot for 48 devices for 1,500 cycles, demonstrating for the first time a good on/off ratio for all cycles of all devices. Perhaps establishing a universal certification standard (similar to that of photovoltaics by the National Renewable Energy Lab) would be useful [122].

Table 2.8 Characterization methods for 2D materials from material to the device level

Methods	Material level	Device level
Optical microscope	Number of layers by colour contrast	Crack, wrinkles, and large area uniformity
SEM	Topographic information, including roughness, edge, wrinkles, ruptures and folds	Correct device structure
(in-situ) TEM	Layered structure	Switching mechanism
STM	Inspect defects like interstitial, vacancy	Switching mechanism
Raman	Number of layers, atomic structure of edges, disorder and defects Stacking order of layers Effect of strain	Switching mechanism
AFM & CAFM	Topography and local defects	Switching mechanism
XPS	Crystal structure	-
Probe station	-	Device performance Device-to-device variability
Probe card	-	Yield of the building block Functionality of whole circuit

2.3 2D materials based memristors

One common strategy to enhance the performance of memristors is to insert new materials in their structure, including perovskites, polymers, carbon nanotubes, nanowires, and two-dimensional (2D) materials. Among them, the use of 2D materials seems to be the one that leads to the best results and has already led to new prototypes with very attractive properties that metal/TMO/metal counterparts do not show., and

Different methods are introduced to engineer relatively controllable defects. From the single device level study of mechanical exfoliate 2D materials, gentle oxygen plasma is used to modify the insulating h-BN and high-temperature annealing (induce oxidation) is used to modify the semiconducting MoS_2.

2.3.1 Graphene

In 2012, graphene is used as the electrodes of SiOx to achieve memristor with high transparency and flexibility [141]. Graphene insertion of HfOx memristor will decrease the reset current, thus the power consumption more than one order of magnitude [142]. Monolayer graphene is enough to block or constrain the penetration of metallic ions or vacancies due to its permeability. By engineering the pore size of inserted graphene, we can control the resistance level of both LRS and HRS [143]. Graphene is widely used in different materials systems, including TaOx [144], AlOx [145], ZnO [146], perovskite [147] and parylene [148] etc, and the common advantage is the lower operation current and stable switching performance. Graphene insertion could also be used to enhance the threshold switching behaviour as selectors.

2.3.2 MoS_2

Oxidized MoS_2 based memristor shows great thermal stability and can operate at high temperatures up to 340 ℃ [149]. The defects introduced by the oxidisation are necessary for the resistive switching behaviour.

MoS_2-Pd nanoparticles hybrid insertion can decrease the operation voltage of HfO_x based memristor [150].

Memristor with single layer MoS2 has been proposed to show the coexistence of bipolar and unipolar switching behaviour [151-152]. However, limited cycles are achieved with around 150 cycles and 20 cycles are shown in each paper.

The reset current can be reduced by the insertion of MoS_2 layer [153]. Vertical structure using MoS_2 can also show synaptic behaviour although the on/off ratio is low less than 5 [154]. Planar structure using MoS_2/graphene can show synaptic behaviour with near-linear weight update [155-158]. LPE MoS_2 has also been used for memristive devices [159] and can be integrated with pressure sensors and quantum light emit diodes [160].

2.3.3 h-BN

2D materials with a perfect single crystalline structure (i.e mechanical exfoliated 2D materials) cannot produce resistive switching behaviour. Increased voltage stress can only in catastrophic material damage [161].

Defects engineering should be introduced for mechanical exfoliated h-BN. Ultra low power oxygen plasma [162-163]

Although nanoscale statistical analysis of I-V test on CVD grown h-BN shows unprecedented reliability even after 100 sweeps and lower variability compared to HfO_2 films, CVD grown h-BN always contains many defective regions like grain boundaries, which will inevitably decrease its reliability as a gate dielectric. However, this early breakdown is normally reversible, which is the fundamental behaviour of a resistive switching device [164].

Memristors with metallic electrodes and multilayer hexagonal boron nitride (h-BN) as RS medium (i.e. Au/Ti/h-BN/Au) have exhibited the coexistence of bipolar and threshold RS [165], which has been used to build electronic synapses for neuromorphic systems [166]. The switching mechanism is attributed to the grain boundaries-assisted active metal penetration and the generation of boron vacancies.

h-BN based memristor connected with the MoS_2 based transistor can show linear conductance change under the control of gate voltage [167].

Mechanically exfoliated BP also needs the introduction of defects by using oxidation.

2.4 Conclusions

In summary, in this chapter, I first introduced the main device structures, the electrodes, resistive switching mediums, and the cell structure of memristors, and explained how to select them based on the materials used and application. A general idea of the switching mechanism is described while it still depends on a case-by-case analysis. Then I explained the detailed fabrication and characterization methods to get the FoMs of memristors. Failure analysis and corresponding suggestions are given. After that, I introduce the functionality of 2D materials, with a special focus on scalable growth methods. Finally, 2D materials-based memristors are introduced by using different 2D conductors, 2D semiconductors or 2D insulators.

Chapter 3: Graphene/h-BN/graphene cross-point memristors with three resistive states

Memristor is a novel two-terminal electronic memory device with a simple metal/insulator/metal (MIM) sandwiched structure. Two-dimensional (2D) material-based memristors have shown several properties that are not shown by traditional ones, such as high transparency, robust mechanical strength and flexibility, superb chemical stability, enhanced thermal heat dissipation, ultralow power consumption, the coexistence of bipolar and threshold resistive switching, and ultrastable relaxation when used as electronic synapse (among others). However, several electrical performances often required in memristive applications, such as the generation of multiple stable resistive states for high-density information storage, still have never been demonstrated.

Here, we present the first 2D material-based memristors that exhibit three stable and well-distinguishable resistive states. By using a multilayer hexagonal boron nitride (h-BN) stack sandwiched by multilayer graphene electrodes, we fabricate 5 μm × 5 μm cross-point Au/Ti/graphene/h-BN/graphene/Au memristors that can switch between every two or three resistive states, depending on the current compliance (CC) and reset voltage used. Basically, we generate an intermediate state between the high resistive state and the low resistive state (LRS), named soft-LRS (S-LRS), which may be related to the formation of a narrower conductive nanofilament across the h-BN because of the ability of graphene to limit metal penetration (at low CCs). Three resistive states have been fitted well with Ohm law, hybrid quantum point contact mode and Poole-Frenkel mode, respectively. All the 2D materials have been fabricated using the scalable chemical vapour deposition (CVD) approach, which is an immediate advantage compared to other works using mechanical exfoliated 2D materials by enabling large-scale device fabrication. Benefiting from this, we can further confirm the device reliability and low device-to-device variability. The successful implementation of this experiment makes the multi-level 2D material-based memristor possible.

3.1 Transfer and characterization of CVD-grown 2D materials

The ability to produce large-area, high-quality, continuous and uniform single crystal 2D materials is a common pursuit of the scientific community. Compared with mechanical exfoliation, liquid phase exfoliation (LPE) and epitaxial growth methods, CVD has the advantages of scalability, low variability between devices and the capability to be transferred onto an arbitrary substrate.

This chapter will first briefly describe the synthesis and transfer of CVD-grown h-BN. Then characterize the quality of the transferred 2D materials. The layered structure and defective state of 2D materials will be an important standard to evaluate the quality of 2D materials after transfer. In this characterization process, I will use the optical microscope, scanning electron microscope (SEM), cross-sectional transmission electron microscope (XTEM), Raman spectrum and conductive atomic force microscope (CAFM).

3.1.1 Growth of 2D materials by CVD

In our group's previous work, multilayer h-BN with large grain size and low defect density were grown by CVD using Ni-doped copper foil as the catalytic growth substrate and liquid borazane as the precursor. The experimental CVD setup is shown in Figure 3.1. The air inlet, precursor chamber, growth chamber and exhaust port of the CVD furnace are displayed from the lower left corner to the upper right corner respectively.

We choose CuNi alloy foil as the growth substrate, because the grain size of h-BN grown on this substrate is large, which can decrease the defect density of the h-BN. Before growth, the 25 μm thick CuNi foil was annealed in a hydrogen atmosphere at 1050 ℃ for 2 hours, which could enhance the flatness of the film and increase the grain size of the alloy. At the same time, annealing in a hydrogen atmosphere will also reduce the native oxides on the metal foil to obtain a better catalytic effect.

In this experiment, borazane was selected as the precursor and placed in an independent chamber 60 cm away from the catalytic CuNi substrate and the temperature was controlled to around 80 ℃. In the h-BN growth process, the hydrogen carrier gas passes through the precursor chamber and carries the precursor molecules to be

deposited on the CuNi foil in the growth furnace. The mono- or multilayer h-BN starts to grow under a high temperature of 1070 ℃ and low pressure of 50 Pa. By controlling the growth time, we can control the thickness of the multilayer h-BN. In addition, by controlling the content of Ni in CuNi foil, we can control the grain size and growth rate of h-BN. At the same time, we found that when the atomic weight of Ni is between 10% and 20%, h-BN has the largest grain size and the fastest growth rate. Finally, the roughness of the obtained h-BN is less than 0.2 nm, which also shows that the samples obtained in this study have excellent morphology [171].

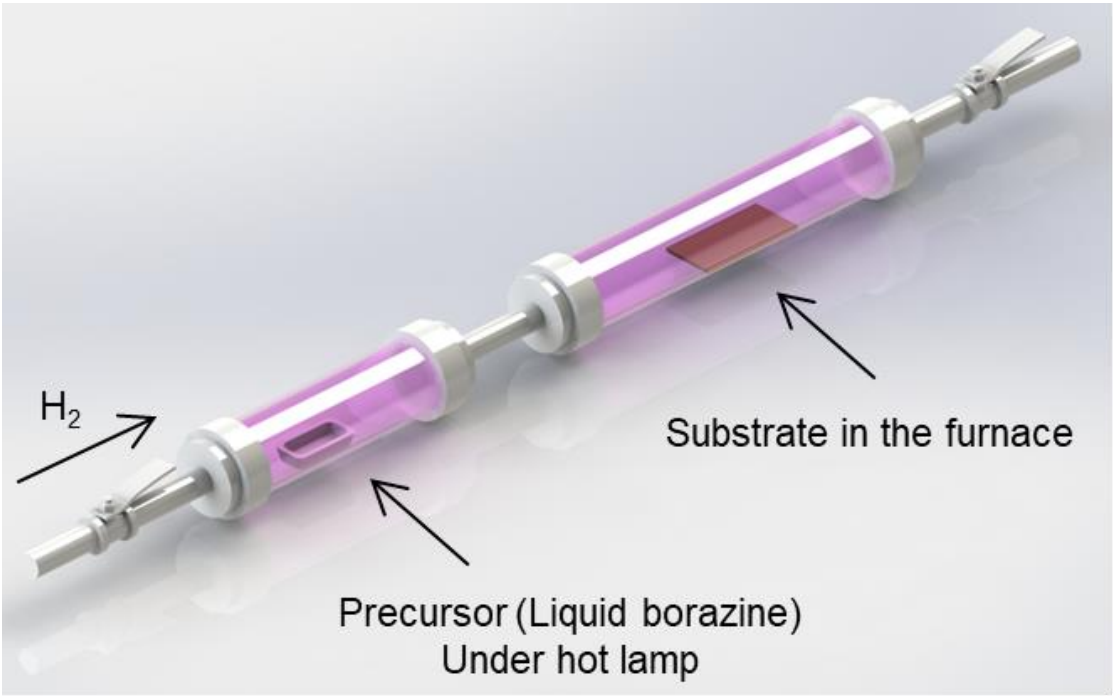


Figure 3.1 Schematic of CVD setup for h-BN growth

3.1.2 Wet transfer of CVD-grown 2D materials

The 2D materials used in this chapter are grown on Cu foil substrate by CVD, for both graphene and h-BN. Different applications need to transfer the CVD-grown 2D materials on a specific conductive or insulating substrate. The basic requirement for the transfer is that the original continuity and uniformity of 2D materials shall not be damaged during the whole transfer process. There should be no cracks, metal or polymer residual after transfer. To achieve this goal, I used the polymer support wet

transfer method. The wet etching method is the most commonly used due to its simplicity [172].

The first step is to spin coat a polymer support layer on the surface of the 2D material. The supporting layer should provide support and protection for the whole transfer process and prevent crushing. The supporting layer can be removed without any residue after the transfer. Here I choose the anisole solution of polymethyl-methacrylate (PMMA) with a concentration of 40 mg/ml. I spin coat PMMA solution on its upper surface as a support layer (spin rate 3500 rpm 45 s, heat at 100 ^{o}C 5 min to remove the solvent). PMMA has many advantages including better mechanical strength than other polymers, low viscosity, good wetting capability, flexibility and good dissolubility in acetone. High mechanical strength can help flatten the 2D materials during the transfer. The PMMA with a shorter chain length and a molecular weight of less than 1,000,000 is recommended. After the growth of 2D material film on the metal foil, both sides of the metal foil have 2D materials. The upper surface normally has better quality (This is also suggested by companies. The company will mark the upside with a star and with notice "this size up"). Without the supporting layer, the 2D materials will dissolve and disappear in the etcher. Figure 3.2a cross-sectional SEM shows that the thickness of PMMA prepared under this process is about 350 nm, and the mechanical strength of PMMA film under this thickness will be enough to protect the 2D materials from crushing in the subsequent etching process.

The second step is to etch and dissolve metal substrates such as Cu or Ni with metal corrosive ferric chloride ($FeCl_3$) solution (concentration: 0.1 g/ml). The environment temperature will affect the etching speed of the metal substrate. The higher the temperature is, the faster the metal will dissolve. The etching time is normally around a few hours at 25 ℃. The etching time shall be extended one hour more to ensure the metal residue is etched completely after the Cu is invisible to the naked eye. The 2D materials on the lower surface without PMMA protection will be dissolved and broken in the etcher. If there is no PMMA protection, the whole 2D material will break and dissolve in the etcher solution.

In the third step, rinse/float 2D materials with dilute hydrochloric acid (HCl) with a concentration of 3 mol/L and DI water in sequence to remove the residual metal and etcher solution and neutralize the alkali solution. Fe^{3+} ions react with Cu to form Fe^{2+} ions, then the Fe^{2+} ion will be oxidized back to Fe^{3+} by oxygen later. HCl solution is used to neutralize the hydroxyl ions produced in these reactions.

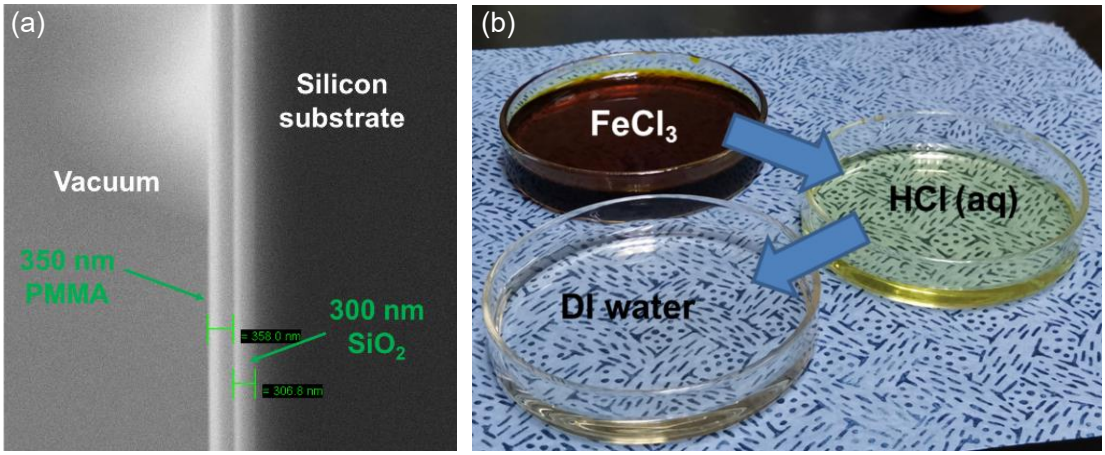


Figure 3.2 (a) Cross-sectional SEM of the spin-coated PMMA on 300 nm SiO_2/Si substrate. (b) Photo of solution for wet transfer.

Step 4: Pick up and transfer the PMMA/2D material thin film to the target substrate. Dry overnight at room temperature or heat on a 50 ℃ hot plate for 5 minutes to remove moisture and make the 2D material adhere to the substrate surface with a good van der Waals contact. So it will not be washed away in the subsequent lithography process. It should be noted that high-temperature heating cannot be used when there is still much water left, because the residual water will react with the metal electrode at high temperature, affecting the electrical performance of the electrode. A high temperature of 150 ℃ or can be used to flatten the wrinkles and provide better adhesion between 2D materials and the substrate after there is no visible water can be seen.

Finally, soak the sample in acetone for more than 12 hours to dissolve and remove the PMMA protective layer. Finally, we can get the 2D material attached to the target substrate. Before removing PMMA by acetone immersion, another drop of PMMA solution can be dropped on the sample surface to alleviate the stress on the 2D material surface to obtain a smoother surface [172].

The wet transfer method supported by the polymer can obtain clean 2D materials with fewer defects, but the cracks and folds that may still appear in the transfer process and the possible residue of PMMA after transfer may affect the electrical characteristics of the 2D material-based memristor after transfer. High-temperature annealing (250 ℃ ~ 350 ℃) is an efficient way to remove the PMMA residues on the surface [173]. High temperature can also help the conformal contact with the patterned metallic electrode [174].

3.1.3 Material level characterization of 2D materials

3.1.3.1 SEM

SEM has been proved to be a very useful technical means to characterize 2D materials, especially the topography and the wrinkles. I transferred the h-BN onto Au/SiO_2/Si substrate and took SEM images to check the quality of the transferred h-BN. Figure 3.3 shows the SEM top view of the multilayer h-BN on the Au/300 nm SiO_2/Si substrate. The SEM image confirms the successful transfer of h-BN, and the h-BN thin film maintained good uniformity and continuity in tens of microns. However, the image also reflects some problems that may occur in the process of CVD growth and wet transfer, such as the wrinkles after the transfer, that is, the white tall lines in the image. Its conductivity distribution was later characterized by the CAFM. The native wrinkles on the as-grown Cu substrate originate from the different thermal expansion coefficients between the h-BN and the Cu foil substrate. When they are cooled from higher growth temperature to room temperature, the Cu foil shrinks and h-BN expands, resulting in wrinkles. While the wrinkles after transfer are mainly due to the evaporation of water underneath the transferred h-BN. The dome shape of the wafer has a larger surface area. The wrinkle can be alleviated by floating the 2D material in DI water for a longer time during the transfer process or high-temperature annealing after the transfer [174].

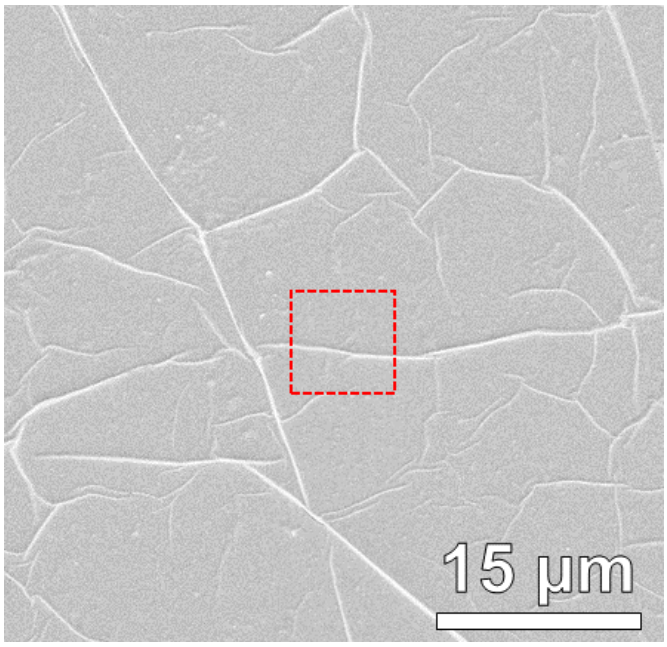


Figure 3.3 Top view SEM of transferred h-BN on Au substrate

3.1.3.2 TEM

Cross-sectional transmission electron microscopy (XTEM) is used to confirm the quality of the h-BN. The quality of the 2D materials is mainly judged by the correct layered structure, the correct number of layers, local defects and defect density.

Special sample preparation is needed to take XTEM. Firstly, the CVD-grown h-BN was transferred to an Au substrate, followed by a layer of Au deposited on top by electron beam evaporation. This Au/h-BN/Au sandwich structure was used to enhance the contrast for a better TEM observation of the layer structure. Secondly, focused ion beam (FIB) etching technology was used to obtain a thin lamella with a thickness less than 100 nm for the TEM characterization. From the XTEM image taken from the above-mentioned Au/h-BN/Au structure, it can be seen that the h-BN has the correct 2D layered structure. The number of layers is around 16 layers, and the thickness of each layer is about 0.33 nm. The green arrow marks the local defects of h-BN, including lattice mismatch and so on. Although the existence of these defect states has the problem of vertical leakage current in the application of transistor gate material, they can show their strength in the application direction of memristor: the defective states will assist in the formation of nano conductive filaments.

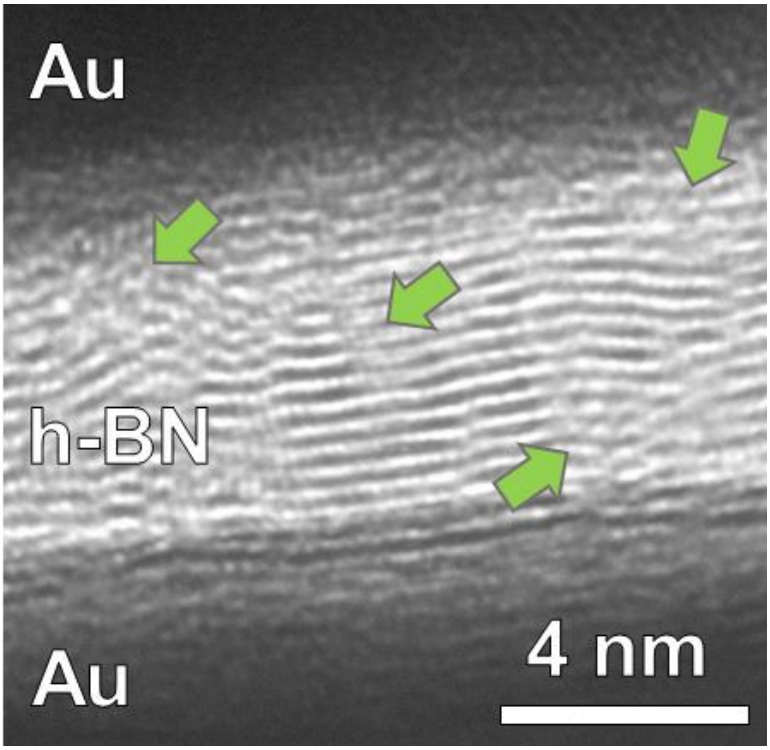


Figure 3.4 XTEM image of the Au/h-BN/Au stack with defects highlighted

3.1.3.3 CAFM

To study the electrical characteristics of local defects displayed by XTEM, I further characterized the transferred h-BN by CAFM. The advantage of CAFM characterization is that it can obtain the topographic and current information of the same region at the same time so that the electrical properties of the materials hidden under the morphology can be compared.

Figure 3.5 shows the collected CAFM topographic and current map of transferred h-BN on Au film with a scanning area of 8 μm × 8 μm. The CAFM image is obtained at a voltage of 5 V with the voltage applied to the AFM tip and the Au substrate kept grounded. The tall white line area corresponding to the wrinkle in the SEM shows no current signal, indicating that the wrinkle is morphologically raised and electrically insulating. Although the topographic map shows similar roughness apart from the wrinkle region, the current map in Figure 2.5b shows that the multilayer h-BN grown by CVD contains regions with higher conductivity (red column) than the surrounding regions (blue planes). These high current regions correspond to the local defects which support the out-of-plane charge transport, such as lattice mismatch. This observation was verified with the results shown in the XTEM image.

It is worth noting that these defects are necessary for the correct operation of the memristor. The previous work of our group shows that high-quality and defect-free h-BN cannot produce any resistance change phenomenon, and electrical breakdown will only cause unrecoverable physical damage to the 2D material obtained by mechanical exfoliation [166]. These defects provide a path for recoverable ion migration as a resistive switching medium.

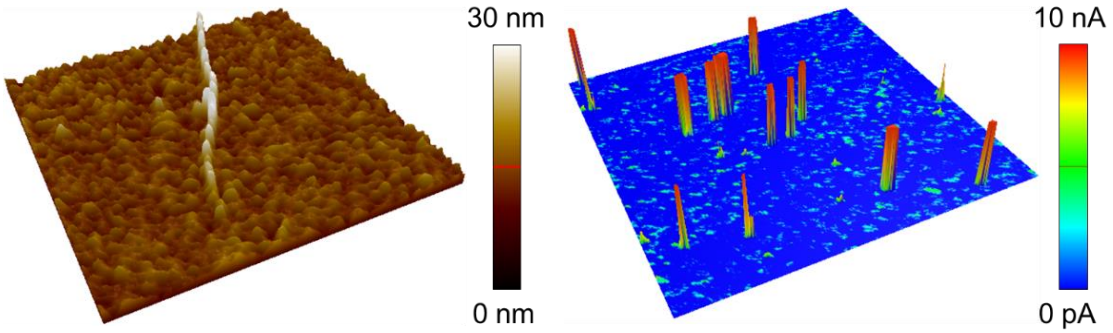


Figure 3.5 CAFM topography and current mapping of transferred h-BN on Au film

3.1.3.4 Raman

Raman spectroscopy is an ideal characterization to characterize 2D materials, which can be used to determine the number and the quality of layers, strain, doping, disorder and functional groups. Compared to XTEM, no special sample preparation is needed. The measurement is fast and non-destructive.

To confirm the two-dimensional nature of multilayer graphene and h-BN used in this experiment, I transferred one piece of multilayer graphene and one piece of multilayer h-BN on a 300 nm SiO_2/Si substrate and collected the corresponding Raman spectra. I collected multiple Raman spectra at 8 randomly distributed locations for graphene and 3 randomly distributed locations for h-BN, which makes the collected data more convincing. Randomly selected multi-characterization points can also show the uniformity of materials, to ensure the low variability between the later fabricated devices.

For graphene, the typical G peak and 2D peak prove that the number of layers of graphene is 3 to 5, and the lower D peak also proves the existence of local defects and lattice mismatch. D peak is related to elastic backscattering, and there is the peak with the highest intensity at the grain boundary, which can be attributed to translational symmetrical fracture or defect [175].

For the h-BN, the Raman spectra show the typical E_{2g} h-BN peak [176].

In this proof experiment, the main purpose is to prove the uniformity of the layered structure and layer number distribution of graphene and h-BN.

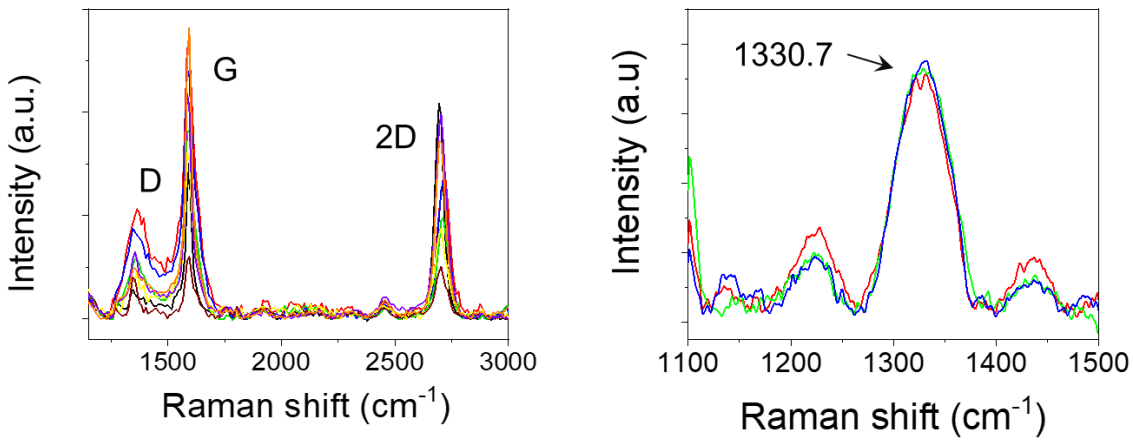


Figure 3.6 Raman signal of the transferred multilayer graphene and h-BN films

3.2 Fabrication of graphene/h-BN/graphene van der Waals structure memristor

This chapter mainly introduces the device structure of the 2D material-based memristor to be prepared, the reasons for the selection of each layer of materials and their deposition methods.

For 2D materials, I choose h-BN as the resistive switching medium because it is a reliable insulating material with a band gap of more than 5 eV. High resistivity also means a large switching ratio, which makes it possible to realize multilevel switching. I choose multilayer graphene as the insertion layer due to its high conductivity and its strong mechanical strength and small pore size that can limit the penetration of metal ions. Both the 3-5 layers of graphene and 15-18 layers of h-BN are grown on Cu foil by CVD and transferred by wet transfer.

For the metal electrode, I choose Au as the inert bottom electrode and Ti as the active top electrode. Ti has been proved to be a reliable top electrode for h-BN based memristors. For the deposition method, I choose the electron beam evaporation method since it can successively evaporate two metals in a high vacuum environment without breaking the vacuum. This method is of great significance for the preparation of top electrode Au/Ti bilayer, because Ti is very active and very easy to be oxidized. The oxidized Ti will not only affect the device's performance but also cause a misinterpretation of the switching mechanism. In addition, the thermal evaporation method can not be used to evaporate metal Ti.

Figure 3.7 shows the structural diagram of the van der Waals structure. The material and thickness from top to bottom are 40 nm Au/10 nm Ti/graphene/h-BN/graphene/40 nm Au/10 nm Ti/300 nm SiO_2/Si. The active metal Ti layer on the top layer will participate in the resistive switching and penetrate into the switching medium to form conductive filaments during the set process. The bottom metal Ti layer is an adhesive layer, which makes the inert metal Au of the bottom electrode adhere better to the SiO_2/Si substrate. The top inert metal Au will not participate in the resistive switching, which is mainly used to enhance the conductivity and prevent the oxidation of the active metal Ti.

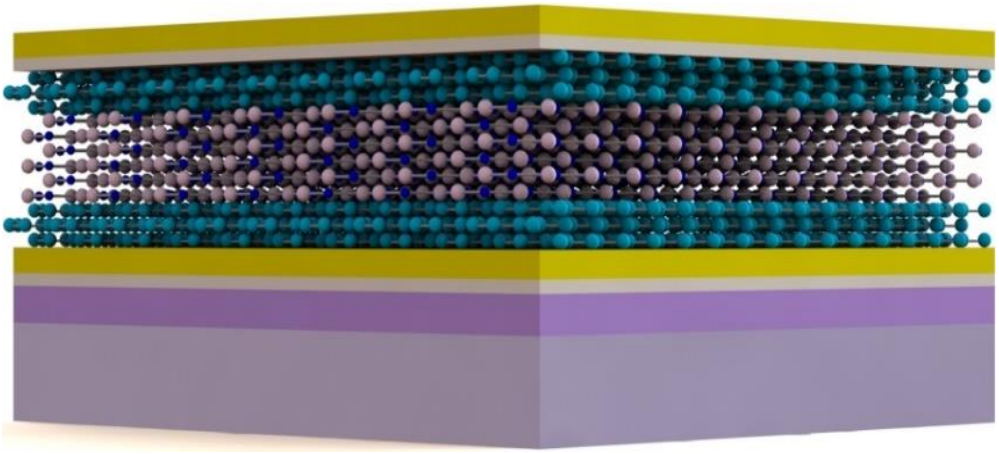


Figure 3.7 Schematic of the van der Waals structured memristor

For the cell structure of the memristor, I choose the crosspoint structure, because it is the right structure suitable for the industry to study resistive switching. Combined with photolithography, we can obtain devices of small size and convenient for the probe engagement of the probe station.

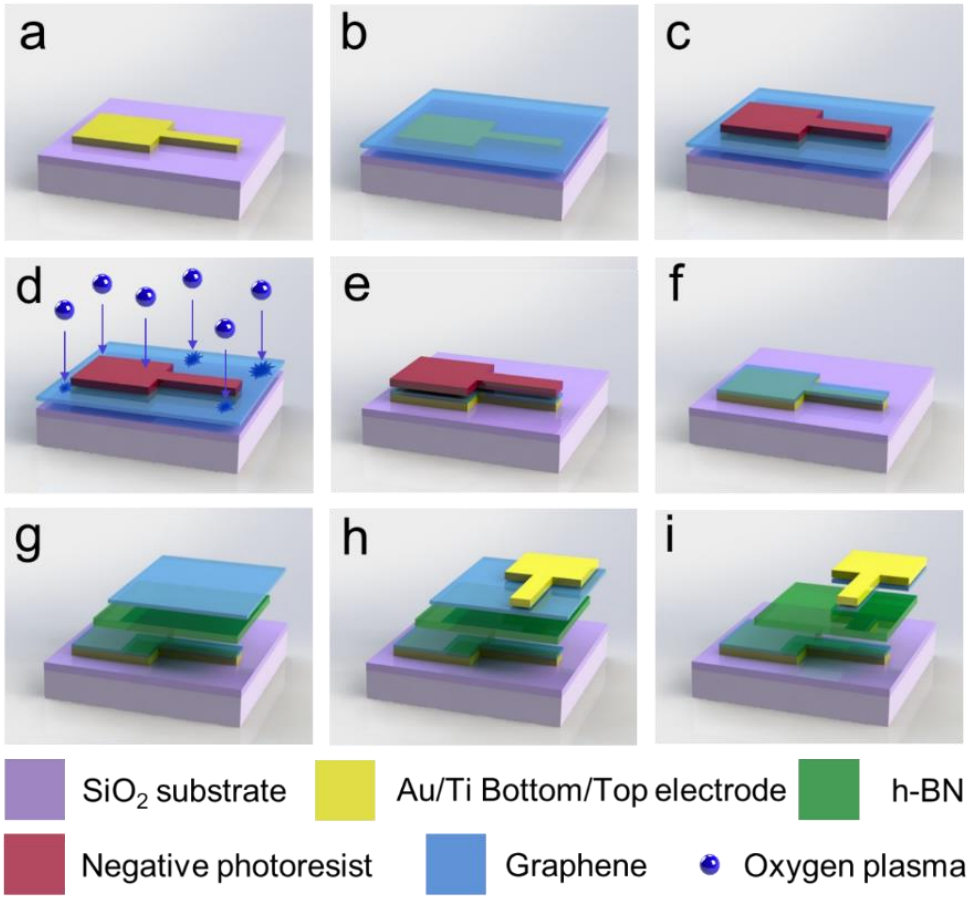


Figure 3.8 Schematic of fabrication of Au/Ti/graphene/h-BN/graphene/Au memristor

3.2.1 Preparation of the SiO_2/Si substrate

In this experiment, a silicon wafer with a 300 nm SiO_2 layer is selected as the substrate and the wafer is cut into a size of 2 cm × 2 cm. The SiO_2/Si substrate was cleaned by ultrasonication in acetone, ethanol and DI water respectively for 15 minutes. Among them, acetone and ethanol are used to dissolve and wash away the organic substances. DI water is used to remove potassium, sodium and other metal ions. Finally, rinse the substrate with running DI water and dry it with dry nitrogen gas.

3.2.2 Fabrication of bottom electrodes

The typical photolithography used for preparing metal electrodes is as followed: Firstly, spin coat and soft bake the positive photoresist on the cleaned SiO_2/Si substrate with a spin speed of 500 rpm for 5 s and 3500 rpm for 45 s. The photoresist will be solidified and the solvent in the photoresist will be volatilized after 100 ℃ for 3 minutes of heating. Secondly, expose and pattern the photoresist with a mask aligner through a photomask. Under the existing conditions in our laboratory, the minimum linewidth of 2 microns can be obtained. Thirdly, develop and remove the exposed area. The developer will remove the area exposed to the lights. Fourthly, deposit the bottom electrode by electron beam evaporation. The principle of electron beam evaporation is to evaporate the material sources by the high-energy electron beam. The chamber was always vacuumed to a high vacuum of 10^{-5} Torr to obtain a high-quality thin film. The instrument can continuously evaporate Ti and Au without breaking the vacuum to avoid the oxidation of the Ti layer. The physical parameters of metal evaporation are:

Table 3.1 Evaporation parameters of the metal electrode

Materials	Vacuum level	Power	Rate
Ti	10^{-5} Torr	5.1 %	0.2 Å/s
Au	10^{-5} Torr	16.1 %	0.4 Å/s

Finally, lift-off by acetone to show the bottom electrode. Each electrode consisted of one 100 μm × 100 μm metal pad for probe station tip engagement connected to a ~150 μm long wire, with a width of 5, 10, 15, 20, 25, or 50 μm.

3.2.3 Bottom graphene transfer and pattern with negative photoresist

To avoid the increase of the device area due to the conductivity of graphene, I need to limit the device area to the crosspoint area with negative photoresist and plasma etching. In the actual preparation process, I first use the same lithography mask as the bottom electrode to pattern a layer of negative glue with the same shape as the bottom electrode. During lithography, the same bottom electrode mask is used, but a pattern is offset accordingly, that is, 10 μm mask alignment of linewidth 5 μm Mask with m linewidth, 15 μm Mask alignment of linewidth 10 μm line width mask, and so on, as shown in Figure 3.9a. The schematic diagram and physical diagram of the patterned photoresist are shown. Then I dry etch the graphene without negative photoresist protection with 100 W oxygen plasma for 5 minutes (see Figure 3.9c). Compared with Figure 3.9b, we can see that the dark green multilayer graphene under the optical microscope has been etched clean. Although there are still some residual graphite islands, the high resistance under the electrical test indicates that the residual graphite are discontinuous and non-conductive. Finally, the negative photoresist is washed away by acetone to show a graphene layer with the same shape as the bottom electrode (see Fig. 3.9c and Fig. 3.9d). The graphene electrode with a slightly larger linewidth ensures the complete coverage of the bottom electrode without leakage.

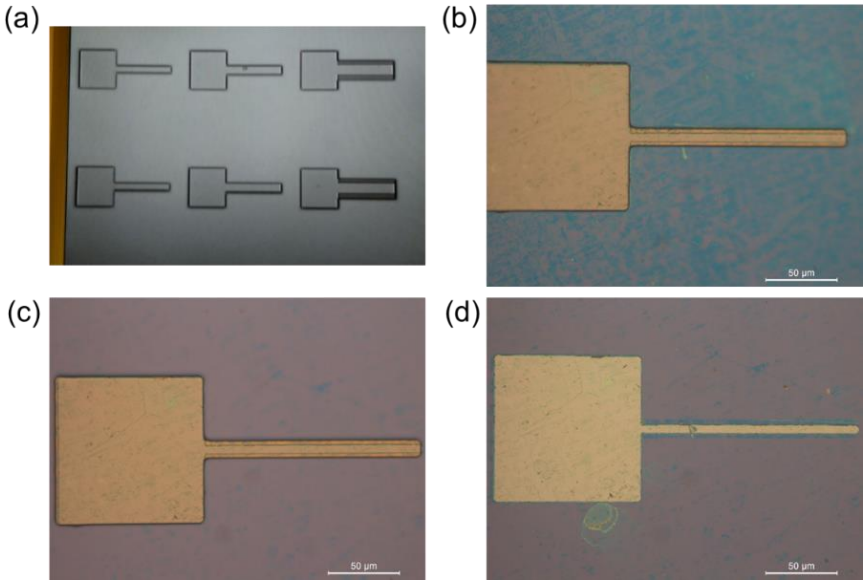


Figure 3.9 Lithography shift alignment (a), and patterned negative photoresist covering the graphene on the bottom electrode (b). The width of the bottom electrode and photoresist are 5 μm and 10 μm, respectively. Optical images of the device after oxygen ion etching (c) and after photoresist removal (d).

3.2.4 Transfer the h-BN and top graphene layer

Patterned h-BN is not required here because h-BN is an insulator and will not crosstalk with surrounding devices.

3.2.5 Pattern deposit the top electrode

The steps are the same as 3.2.2, but the mask shape rotates 90 ° with the top electrode to form a crosspoint structure to reduce the device area.

3.2.6 Plasma defines the top graphene electrode

The steps are the same as those in 3.2.3. At this time, the top electrode will be used as a mask for oxygen ion etching.

3.3 Characterization of the crosspoint structure

After the device is successfully fabricated, I characterized the morphology of the electrode structure respectively to ensure the correct transfer of 2D materials, obtain the correct intersection shape and obtain the accurate device area.

After the device was prepared, I first confirmed that the intersection structure was accurately prepared by an optical microscope. Figure 4.1a below is the physical top view of the optical microscope of the prepared electrode intersection structure and its array. The intersection area of each column from left to right is 5 μm × 5 μm, 10 μm × 10 μm, 15 μm × 15 μm, 20 μm × 20 μm, 25 μm × 25 μm and 50 μm × 50 μm. Each row has the same device area. The square metal plate used for the needle under the probe stage has an area of 100 μm × 100 μm. The top and bottom electrodes can be easily identified through the colour difference: the electrode with the metal plate on the left and the colour green is the bottom electrode under h-BN, while the electrode with the metal plate on the top and the colour gold is the top electrode without shelter. The lower left corner of the picture shows the crack of 2D material, and the line connected by black spots at the upper end of the image is the boundary of the transferred 2D material, indicating that the 2D material has been successfully transferred.

I also collected the SEM image of the crosspoint structure device to confirm the minimum area of the device. The SEM top view of the prepared device confirms that the area at the minimum intersection of the device is 5 μm × 5 μm. The boundary of the electrode is clear and free of burrs, indicating that the maturity and reliability of the lithography and lift-off process are relatively high. The wrinkles on the electrode also confirmed that the 2D material was successfully transferred. The dielectric and electrode surfaces are also relatively clean without excessive metal and polymer residues, indicating that the transfer step of 2D materials is well implemented.

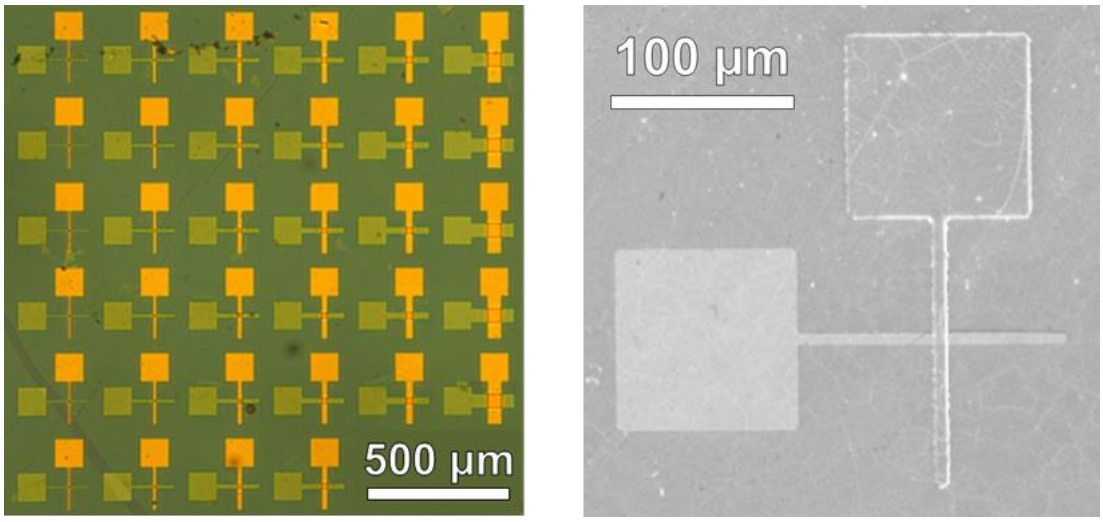


Figure 3.10 Optical and SEM images of the crosspoint array and single device

3.4 Electrical characterization of the van der Waals device

The electrical characterization of the experiment was obtained by a standard room temperature probe station and semiconductor parameter analyzer. Electrical tests are to apply ramped voltage stress (RVS) to the top electrode and keep the bottom electrode grounded. The application of this voltage direction is helpful for us to judge the active electrode involved in the positioning process. Since the h-BN on the bottom electrode is not etched, two probes on the same bottom electrode are needed to confirm the Ohmic connection between the ground tip and the bottom electrode.

3.4.1 Analysis of forming process

I first analysed the electrical forming process of the device, which is the first current-voltage (I-V) curve of a fresh device. Generally, the forming process requires a voltage higher than the subsequent set voltage. The electrical forming curve in Figure 3.11a shows that the initial resistance state of the fresh device is HRS, and the current is about 10^{-7} A at 0.1 V, A sudden current jump to the CC of 1 mA can be detected until the voltage rises to 5 V, indicating the occurrence of the first dielectric breakdown of the h-BN insulating layer. To confirm that the initial high resistance depends on the h-BN layer, I prepared the control sample of structure Au/Ti/graphene/graphene/Au: the initial high current of the sample further confirmed that the sharp increase of the current lies in the breakdown of the h-BN, which also shows that the out-of-plane current of multilayer graphene can be ignored.

To study the relationship between the forming process and device area, we define the voltage at the maximum current slope as breakdown voltage (V_{BD}), and the maximum current before breakdown is the breakdown current (I_{BD}). Through the statistical analysis of the data results from 17 devices, I found that the smaller the device area, the higher V_{BD} needed, and the smaller I_{BD} needed. as shown in Figure 3.11b and Figure 3.11c. These observations are consistent with the percolation theory [1], that the breakdown event generates one or more conductive nanofilaments (CNFs) at the place/s with the electrically weaker locations. A larger device has a higher chance to encounter these areas with weak electrical performance, which leads to lower V_{BD} and higher I_{BD}. There is a high probability that these electrically weak locations are the defects indicated by the green arrow in the XTEM image and the high current region in the CAFM current map, which is also the preferred switching region for subsequent resistive switching.

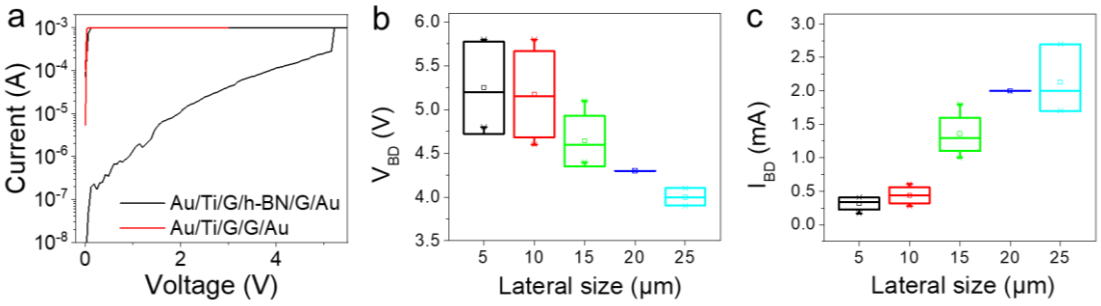


Figure 3.11 Analysis of forming of Au/Ti/graphene/h-BN/graphene/Au memristor

3.4.2 Tristate operation

After the device is electrically formed, I apply a cyclic RVS with opposite polarity (set process with CC, reset without CC). The device shows a stable bipolar RS with a device area of 5 μm × 5 μm. The memristor with Au/Ti/graphene/h-BN/graphene/Au structure usually displays two resistive states, while the LRS can be adjusted by the value of CC and the HRS can be tuned by V_{RESET} to realize multilevel storage. When a relatively high CC of 5 mA, the device switches between HRS and LRS. When using a relatively low CC of 0.5 mA, the device switches between the HRS and the LRS with low conductivity, which we name soft LRS (S-LRS). The V_{SET} of the device is stable at about 3.5 V, while V_{RESET} is about - 1 V under the high CC of 5 mA and about - 3.5 V under the low CC of 0.5 mA. Interestingly, in this experiment, when using CC 1 mA, these three resistive states are randomly shown in different cycles. Although it is a well-known behaviour to observe a higher current in the memristor when a higher CC is applied, which is related to the formation of wider CNF on the dielectric, metastable tristate RS has never been reported before for any 2D materials-based memristor. At the same time, two obvious current drops can be seen in the reset process under the 5 mA high CC, highlighted by the blue circle, indicating that LRS can be adjusted to S-LRS or HRS depending on the value of the V_{RESET}.

I drew the three-dimensional structure diagram of the device status under different resistive states and analyzed the structure and composition of the CNFs in the dielectric layer. Among them, the main components of the conductive filaments are the active Ti metal. When the device is in the HRS, there are only random defect states in the dielectric layer. When the device is set to LRS under high CC, a large number of Ti ions are drifted into the h-BN by the electric field. Then the Ti ions will be reduced to form one or more wide CNFs when they are approaching the bottom electrode. A limited number of boron vacancies (green ball) are driven by electron wind at the defective regions of h-BN. The LRS can be reset back to the HRS under a high I_{RESET} of value ~10 mA, which indicates the CNF is disrupted by the Joule heat. When the device is set to S-LRS under low CC, a nano-sized conductive filament is formed in the dielectric layer. The S-LRS can be reset to HRS under the influence of ion migration under the high electric field. In addition, LRS can also be reset to HRS or S-LRS by adjusting the V_{RESET} and S-LRS can be further set to the LRS by increasing the CC, to realize the switching between any two of these three resistive states.

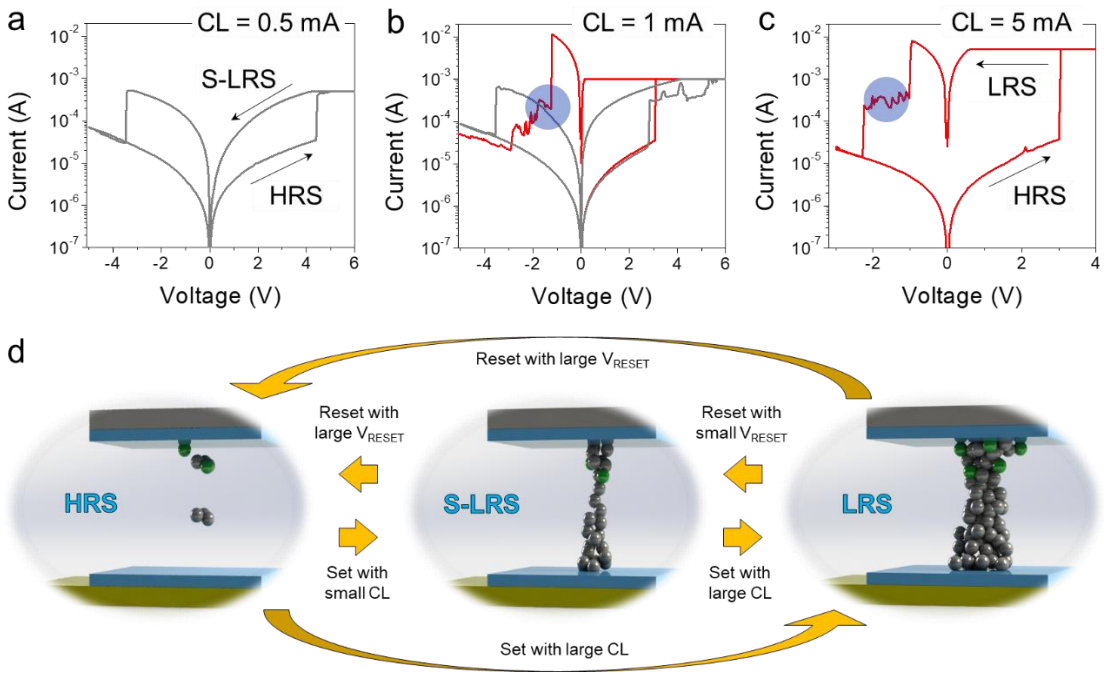


Figure 3.12 Tristate operation by controlling the CC and V_{RESET}

In the following three consecutive sections, I will further statistically analyze the detailed working condition, including the resistance level, simulation of the I-V curves and the device reliability.

3.4.3 Two-state operation between HRS and S-LRS under CC 0.5 mA

Figure 3.13a shows 50 consecutive I-V curves of typical two-state RS between HRS and S-LRS under CC of 0.5 mA. Figure 3.13b shows the cumulative probability plot of R_{HRS} and $R_{S\text{-}LRS}$ (read at 0.1 V). It can be observed that there is no overlap between the resistance of HRS and S-LRS, which indicates that the peripheral circuit can effectively and reliably distinguish the two resistive states. In addition, HRS showed a lower variability than S-LRS. In figure 3.13c, we statistically analyzed the V_{SET} and V_{RESET} of each cycle and found that the variability of the V_{SET} is larger than that of V_{RESET}.

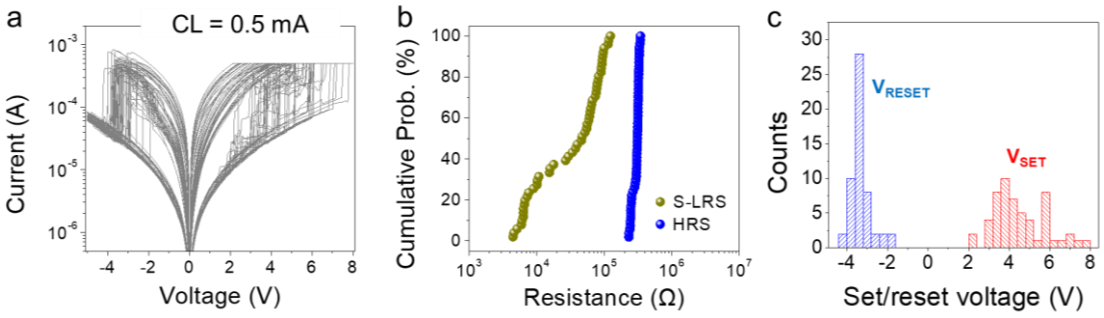


Figure 3.13 Two-state operation with CC 0.5 mA

3.4.4 Tristate operation under CC 1 mA

When the CC is increased to 1 mA, HRS can be randomly switched to S-LRS or LRS. Figure 3.14a shows the 150 I-V curves between HRS and S-LRS and Figure 3.14b shows 150 I-V curves between HRS and LRS, respectively. The two sets of curves are collected in the same device and separated based on the two different RS characteristics. Comparing the figure 3.14a and figure 3.14b, the reset process from LRS shows a more abrupt current drop, while the reset process of S-LRS is more gentle, indicating the different switching mechanism by Joule heat produced by high current or ion drift by the high electric field.

Figure 3.14c shows the experimental measurement curves and corresponding fitting curves of three different resistive states. The solid blue line, solid dark yellow line, and solid red line correspond to the experimental I−V curves measured in the HRS, S-LRS, and LRS, respectively.

By introducing the quantum conductance G_0, I qualitatively evaluated the three different resistance states. Quantum conductance refers to the conductance of a single-channel quantum wire under ballistic transport, that is, the conductance of a single atom-wide continuous conductive filament. G_0 is a constant value equal to 77.5 μS [177], which is plotted as the gray dash line in figure 3.14c. When the conductance of the device is much greater than G_0, we believe that the CNF in the dielectric layer has been completely formed. When the conductance of the device is less than G_0, the CNF between the top and bottom electrodes is not fully formed. When the conductivity of the device is close to G_0, we believe that one or more nanoscale conductive filaments are formed between the top and bottom electrodes [177]. With a reading voltage of 0.1 V, the resistance of S-LRS is calculated at ~ 10^4 Ω and that in LRS is ~ 10^2 Ω (see dark

yellow and red spheres). In the S-LRS, the conductance at low voltage is similar to the G_0 in some cycles slightly above and in others slightly below the G_0. This observation indicates that the width of the CNF is atomic scale. In the LRS, the conductance of the CNF is orders of magnitude higher than G_0, indicating that CNF has been fully formed and expanded laterally. The difference between each resistive state is sufficient to be distinguished by peripheral circuits (see Figure 3.14e, f).

To better understand the charge transport mechanisms, the I−V curves in each resistive state have been fitted to different conduction modes across the h-BN stack using the SIM^{2}RRAM simulator. SIM^{2}RRAM simulator is a physical model developed by Dr Marco A. Villena for memristor simulation. The model is obtained by combining a variety of electron conduction models on the premise of understanding the resistance variation mechanism. Although the model was initially used for HfOx-based memristors, with the help of Dr Marco, we extended the application of the simulator to the field of 2D materials.

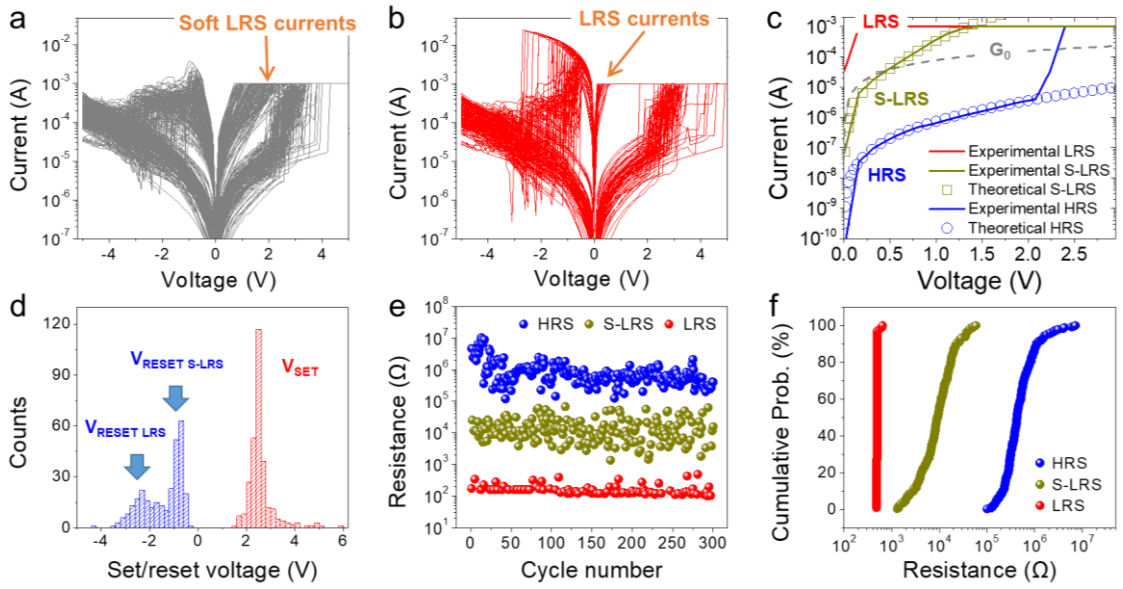


Figure 3.14 Tristate operation between HRS, LRS and S-LRS in Au/Ti/graphene/h-BN/graphene/Au cross point memristor.

Firstly, the relationship between current and voltage in the LRS is very consistent with Ohm's law, a linear relationship between current and voltage, which clearly shows that the electron conduction model in the state is based on the CNF.

Secondly, hybrid quantum point contact (QPC) model was used to fit the current in the S-LRS. The conduction of the CNF is described by the Ohmic conduction of the CNF (fully or partially formed) plus QPC conduction along the interface CNF/electrode. The fitting line of S-LRS is plotted as empty dark yellow squares, and the current follows the expression (1):

$$I_{S-LRS} = \frac{N}{R_0+NR_S}V + \frac{N-n}{e\alpha(R_0+NR_S)} ln\left[\frac{exp\{\alpha[\varphi-\beta e(V-IR_S)]\}+1}{exp\{\alpha[\varphi-(1-\beta)e(V-IR_S)]\}+1}\right] \qquad (1)$$

where N is the total number of partially formed CNF and n is the number of fully formed CNF. R_0 and R_S are quantum resistance (G_0^{-1}) and the series resistance, respectively. α Is the barrier thickness of the Fermi level, β represents the potential fraction drops in the barrier, and φ is the barrier height from the Fermi level. R_0 can be easily calculated, and its value is $1/G_0$ = 12.9 kΩ. R_S is calculated from the reset I-V curve with the highest conductivity before the reset process (i.e. the resistance of reading voltage 0.1 V in LRS), and the calculated value is 150 Ω. For $N \rightarrow \infty$, n = 0, α = 3.6, β = 1 and φ = 1.9 eV, the I-V curve in S-LRS is completely consistent with equation 1. The barrier parameters (α, β, and φ) used are in line with those used in previous RS studies, and the use of N → ∞ and n = 0 indicates that the CNF is almost formed but there is still a potential barrier at the CNF/electrode interface.

Finally, the I-V curves measured in the HRS fit with the Poole-Frenkel conduction model [178] as described in equation 2:

$$I_{HRS} = aEe^{\frac{-q\left(q\phi_B-\sqrt{\frac{q}{\pi\epsilon}E}\right)}{k_BT}} \qquad (2)$$

where q is the elementary charge, ε is the dielectric constant of the material (h-BN in our case), ϕ_B is the barrier height of the trap state (0.4 eV), T is the temperature, and K_B is the Boltzmann constant. The fitting parameter a is related to the diameter of partially formed CNF and electronic drift mobility. The non-measured parameter a has been obtained by fitting the I-V curve collected in the experiment with a value is 4.5× 10^{-8} C^2 N^{-1} s^{-1}. The fitting curve in the high resistance state is drawn in figure 3.14c using empty blue circles. The Poole Frenkel model in the HRS state can be explained by the defects remaining in the dielectric layer after the reset process.

With CC 1 mA, there will be the coexistence of three resistive states. The probability cumulative distribution plot is distinguished according to the distribution density of resistive states, which is indeed subjective. However, in actual operation, only when the CC is 5 mA and 0.5 mA, the resistive switching behaviour is highly controllable.

3.4.5 Typical two-state resistance under CC 5 mA

When the CC is further increased to 5 mA, the device shows a typical two-state resistance change between HRS and LRS. Figure 3.15a shows 100 continuous I-V cycle curves. The switching current ratio (I_{LRS}/I_{HRS}) read at 0.1 V is greater than one order of magnitude (Figure 3.15b), and its cycle-to-cycle variability allows reliable identification of each resistive state (Figure 4.6c). Further analysis of V_{set} and V_{reset} also shows that the variability is acceptable for memristive technologies, compared with memristors based on HfO_2. All I-V curves show reproducible LRS current after each set event. (see the orange circle in Figure 3.15a), indicating that the CNF formed in each cycle is stable. This observation is related to the use of a high CC of 5 mA, which forms a wide and stable conductive filament.

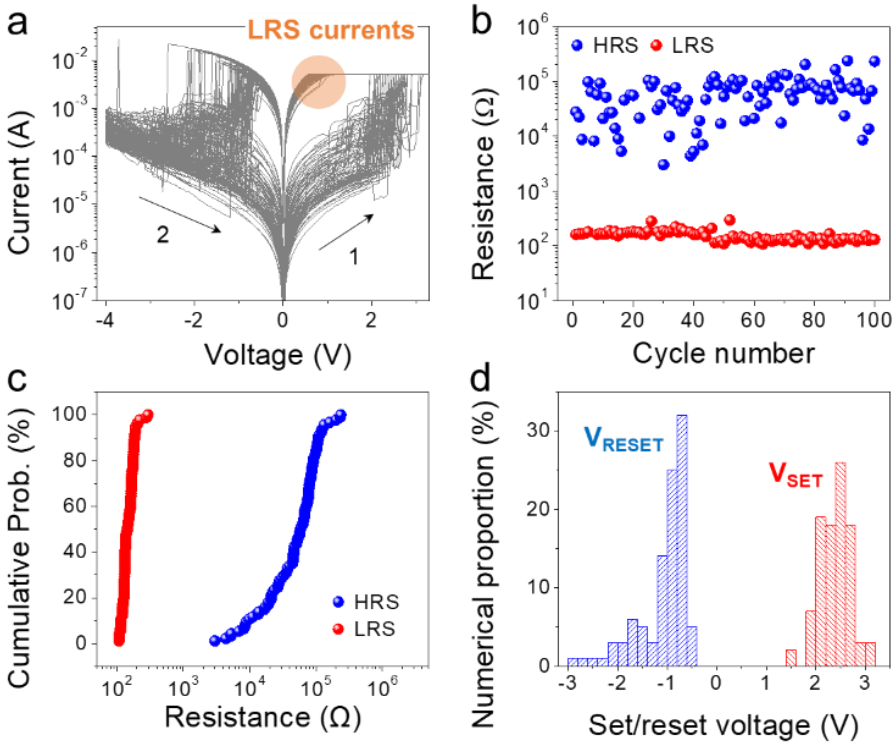


Figure 3.15 Two-state operation between HRS and LRS with CC 5 mA

It is worth noting that even if the device works under high CC, such as 5 mA, it can still return to the two-state resistance between HRS and S-LRS through the reset process, indicating that the device can switch between any two states, and each resistance direction is reversible.

It is worth noting that I designed two groups of comparative experiments to find the necessity of device area miniaturization and graphene intercalation for the realization of multilevel memristor, and listed the possible causes and as well as the subsequent suggestions for multilevel RS experiment in Table 3.2. In the first group of comparison experiments, I constructed a memristor with Au/Ti/ graphene/h-BN/graphene/Au van der Waals structure, but the size is larger of 100 μm × 100 μm. For any current limitation device, there is no metastable tristate operation. The reason may be that the size of those devices is similar to the grain size of the polycrystalline Cu foils on which the h-BN was grown [179]. Grain boundaries, resulting in more defects (e.g. pentagonal and heptagonal bonding, atomic deletion). In such case, the devices contain portions of h-BN that was grown on a Cu grain boundary (GB) and that results in a higher amount of defectives (i.e., pentagonal and heptagonal bonding, missing atoms). The existence of these defects makes the dielectric strength of the large-area h-BN insufficient to induce S-LRS. Graphene devices are more prone to breakdown at the grain boundary area. This observation also evidences that when using 2D materials, observations in large devices with areas $\geq 10^4$ μm^2 cannot be directly extrapolated to miniaturized (area ≈ 1 μm^2) devices and vice versa. In the second group of comparison experiments, I constructed a structure with the same small device area 5 μm × 5 μm metal/h-BN/metal devices (without graphene intercalation) also do not show tristate operation (even when similar CC is used), which means that the existence of graphene is very important and indispensable to reduce the penetration of metal into the h-BN dielectric layer and reduce the size of CNF to form S-LRS.

Table 3.2 Importance of device miniaturization and graphene insertion

	Large size (100 μm × 100 μm) with Graphene	Small size (5 μm × 5 μm) without graphene
Tristate operation	No	No
Possible reason	Device size is similar to the grain size of the polycrystalline Cu foils. 2D films grown on Cu grain boundary are electrically weaker to have enough dielectric strength to induce S-LRS.	Graphene layer is needed to constrain the ion diffusion to stabilize the S-LRS.
Suggestion	Crosspoint structure is recommended for RS.	Graphene insertion is recommended for stabilization.

3.4.6 Device failure analysis

Usually, the device will fail in the HRS with crosspoint area broken. After the device is set to LRS, the Joule heat generated by an excessive reset current (higher than 10 mA) will make the material thermally expand and permanently damage the device. This phenomenon is more likely to occur at the crosspoint region and the minimum line width region. As shown in Figure 3.16a, after several normal cycles of I-V curves, the memristor suddenly drops to an ultra-high resistance state (resistance is about 10^{10} Ω) after reaching a high current of 0.1 A in the fourth reset process, which is highlighted by the golden circle. The device cannot be set back to the LRS again in the subsequent setting process. Later, the SEM image at the crosspoint area shows the material fracture occurred between the bottom electrode and the top electrode, and some metal materials and 2D materials expanded and disappeared, thus forming a physical open circuit and behaving as an ultra-high resistance state electrically. It indicates that the Joule heat will damage the device and cause device failure.

The preliminary work of our group shows that the failure rate of the device will be effectively reduced by baking the device at 85 ℃ for 5 minutes before the test and removing the moisture that may be attached to the device after long-time exposure to the air [180].

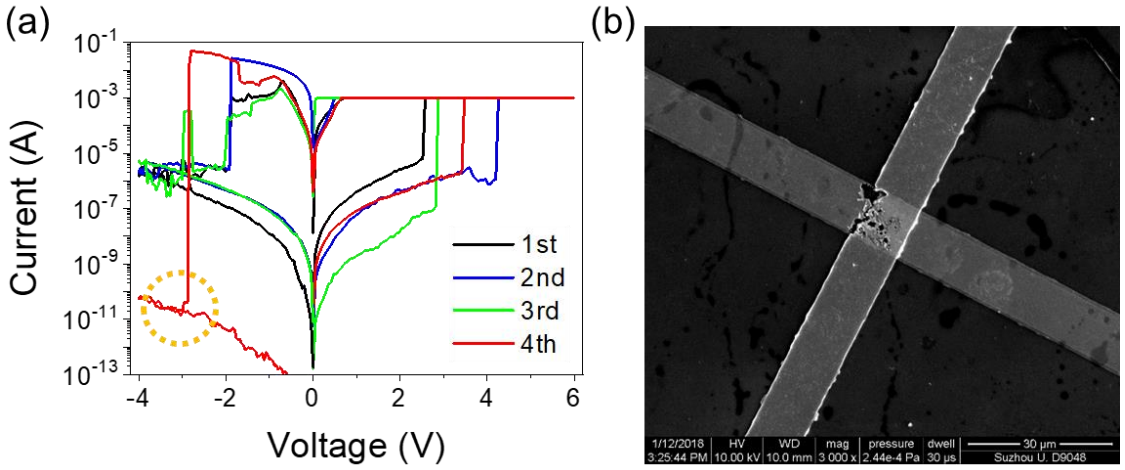


Figure 3.16 Device failure analysis

3.5 Conclusions and prospect

In conclusion, we have fabricated matrixes of 5 μm × 5 μm cross-point Au/Ti/G/h-BN/G/Au memristors that exhibit tristate operation depending on the CL and the reset voltage used. This operation has not been observed in similar Au/Ti/G/h-BN/G/Au devices with larger size (100 μm × 100 μm) nor in graphene-free 5 μm × 5 μm Au/Ti/h-BN/Au devices. These observations indicate (respectively) that miniaturization of 2D material-based memristors is essential to monitor the real performances of the devices (e.g., tristate operation) and that graphene can limit ionic exchange between the metallic electrode and the h-BN (RS medium) to allow the formation of an additional S-LRS. All the results have been confirmed statistically in more than 40 memristors with different device areas, which have been fabricated using a scalable and industry-compatible CVD method.

As Academician, Wu Hanming mentioned in his lecture at the information and electronic engineering frontier forum of the Chinese Academy of Engineering on May 16, 2020: "when the integrated circuit process node reaches 2 nm, the 2D and even lower dimensional materials will shine". 2D materials have a broad future, and 2D material-based multi-resistive state memristors have unlimited possibilities in the future.

Figure 3.15a shows an obvious current overshoot, which causes the current in the I-V curve under negative bias to be greater than the CC used during the set process. From a technical point of view, using a transistor in series with a memristor (1T1R) structure to reduce the current overshoot will provide better control of the size of the conductive wire and slow down the irreversible and irreducible lateral expansion of the conductive wire. Therefore, this will improve the durability of the equipment; Using the 1T1R structure is obviously the right way to avoid overshoot. However, combining a 2D material-based memristor with a transistor is a very complex task, and so far, only a few teams have been able to achieve it. Because the 1T1R structure based on 2D materials has higher complexity, I described my attempt in this direction in Chapter 4.

Chapter 4: Hybrid 2D/CMOS microchips for memristive technologies

4.1 Introduction

Two-dimensional (2D) layered materials could be employed to fabricate advanced integrated circuits, owing to their excellent electronic and thermal properties. However, most studies were limited to the fabrication and characterization of single devices, often fabricated with techniques that are incompatible with the omnipresent complementary metal oxide semiconductor (CMOS) technology (such as mechanical exfoliation). It is really difficult to implement the real applications, even just for a scientific demonstration, especially since the system is composed of several building blocks. For example, S. Wachter and his collaborators developed a 2D materials-based microprocessor, although the yield of each building block is high as 80%, the whole circuit only has a limited overall yield of a few percentage [181]. A hybrid system can take advantage of novel devices and materials and mature technology.

Here we present an industry-compatible approach to fabricate high-density high-performance CMOS/2D materials hybrid microchips by transferring the 2D material on an unfinished silicon wafer containing the CMOS components and wires and finalizing the circuits by patterning the top electrodes and wires. As an example, we use insulating multilayer hexagonal boron nitride (h-BN) to fabricate crossbar arrays of memristors that are connected in series to CMOS transistors. We observe that the CMOS transistors remarkably improve the performance of the h-BN memristors because they can completely remove the overshoot, allowing for the first time the observation of non-filamentary switching in h-BN. As a result, we present the longest endurance and best conductance controllability ever reported for any 2D materials based memristor. The hybrid CMOS/2D materials one-transistor-one-memristor cells also exhibit spike-timing-dependent-plasticity, which is suitable for the implementation of spiking neural networks with high accuracy and low power consumption during image classification. This reliable approach to build high-density hybrid CMOS/2D materials microchips paves the way toward future 2D electronics.

4.2 Fabrication of hybrid 2D/CMOS memristive microchips

Our silicon microchips have been designed via Synopsys software and fabricated in 200mm silicon wafers in an industrial clean room using a CMOS technology of the 180nm node (see Figure 4.1). Each microchip is 2cm×2cm and contains multiple circuits, ranging from 5×5 crossbar arrays of 1T1M cells to operational amplifiers and other peripheral hardware for the control of memristive neural networks.

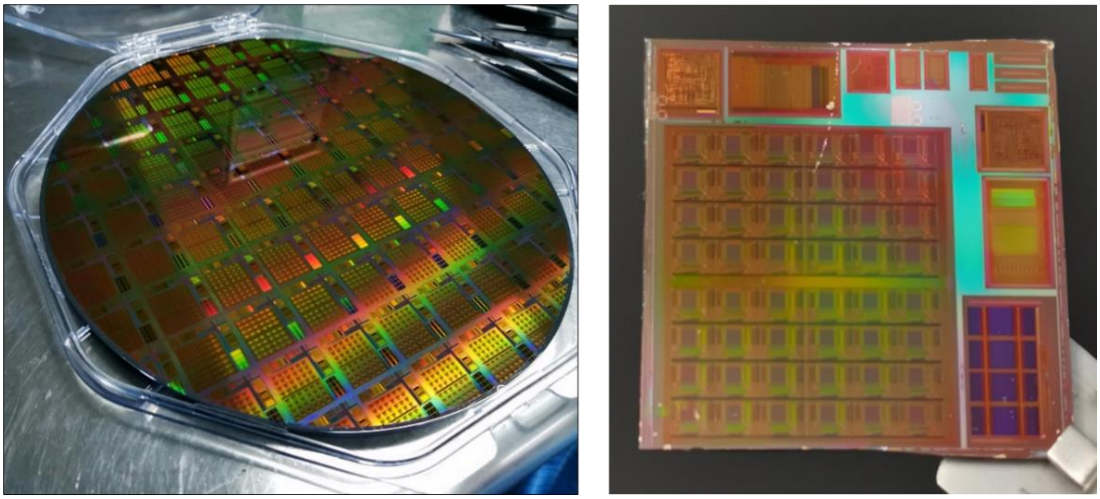


Figure 4.1 Photos of the wafer containing the CMOS circuitry and one identical functional chip

Columns 1 and 2 are for the single memristors. The blue box highlights one single memristor. Columns 3~8 are for the single 1T1M. The red box highlights one single 1T1M cell. The two matrixes at the right are for the 1T1M 5×5 crossbar array. The green box highlights one crossbar array.

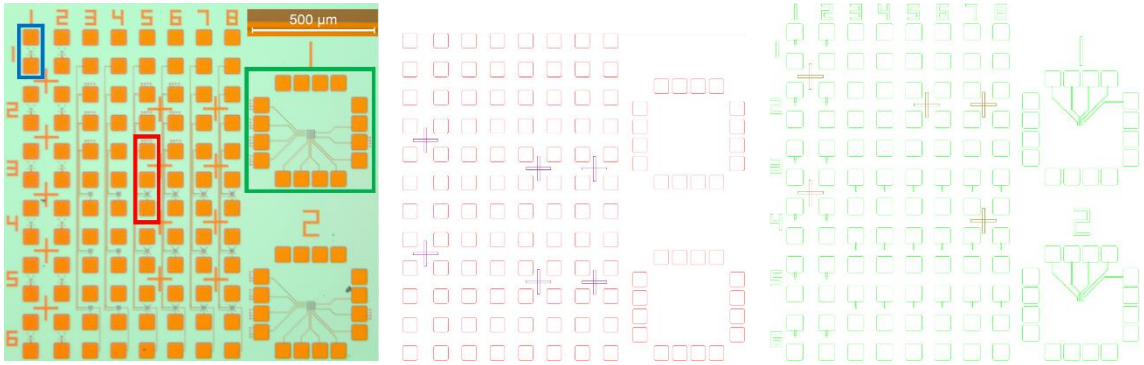


Figure 4.2 Optical microscope image of the interested area and mask design

The CMOS circuits have been designed to integrate the memristors at the BEOL interconnections, i.e., the microchips have been terminated at the fourth metallization layer and have been left without passivation. Hence, oxide naturally grows on the wafers when they are extracted from the industrial clean room (see Figure 4.3), which can be easily etched away using a hydrofluoric acid solution, which exposes the tungsten vias of the fourth metallization layer. Then, a ~18-layers-thick sheet of h-BN (~6nm) was transferred onto the microchip using a low-temperature process. Another photolithography was used to expose the h-BN on the bottom electrode with a dry etching method of Ar/O_2 plasma 300W for 10 minutes. Finally, top electrodes made of 3 nm Ti and 40 nm Au were deposited on the memristors to finalize the circuits, and the h-BN on the contact pads was etched. As the tungsten vias have a diameter of ~260nm, the lateral size of the resulting Au/Ti/h-BN/W memristors was, at most, $0.053\mu m^2$. For the as-received chip, only the vias (the bottom electrode highlighted in the white circle and the active via highlighted in the red circle) are exposed to the surface.

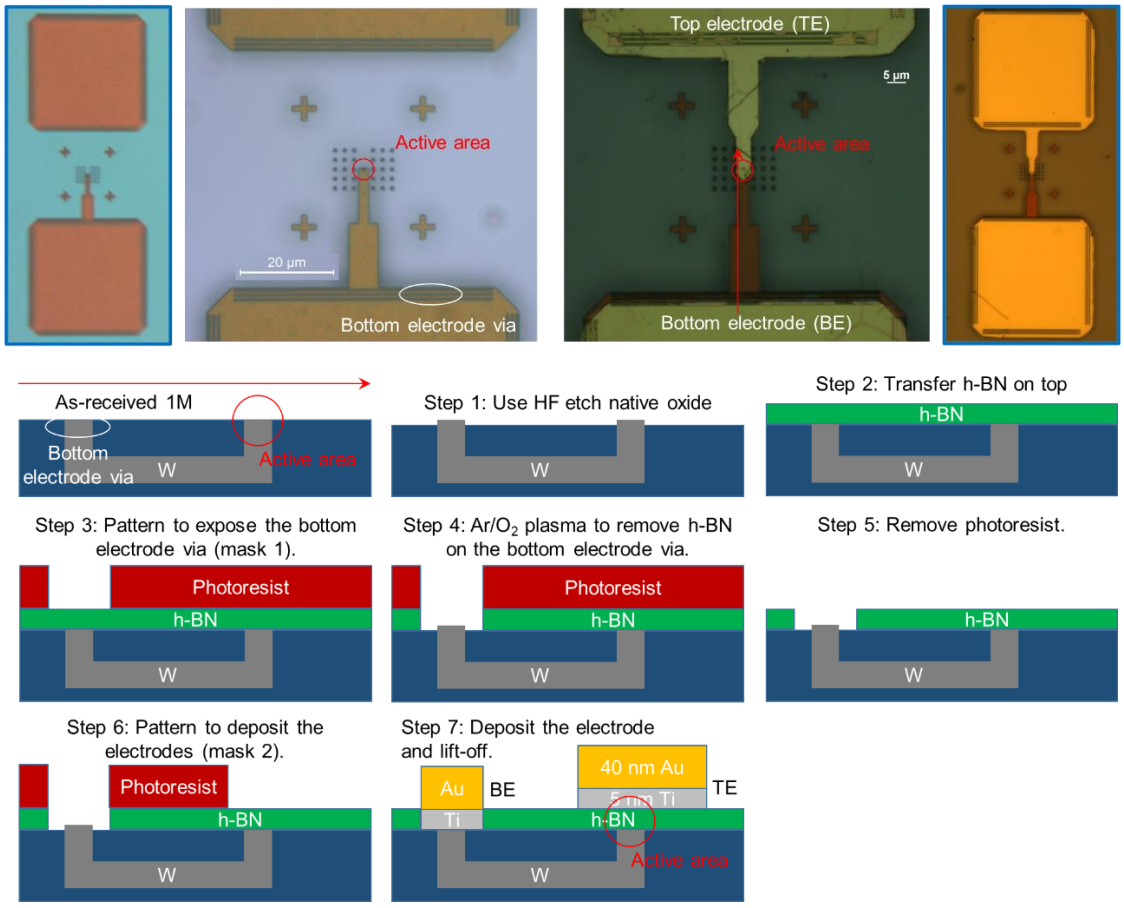


Figure 4.3 Optical images of as-received and fabricated 1M region with structure Au/Ti/h-BN/W, and detailed fabrication procedure.

Figure 4.4 shows the optical images of as-received and with memristros integrated 1T1M crossbar region. The optical microscope images reveal that the h-BN sheet does not crack during the transfer; this is an important advantage of using ~6-nm-thick 2D layered materials, and it remarkably increases the yield of the devices and circuits. Topographic maps are collected by AFM of the vias in the 5×5 crossbar arrays on the wafers as-received, after native oxide etching, and after the transfer of the h-BN sheet. The AFM image shows limited wrinkles after the h-BN transfer, and none of them passes through the active vias. The last optical image shows the zoon-in image of a finished 5×5 crossbar array of 1T1M.

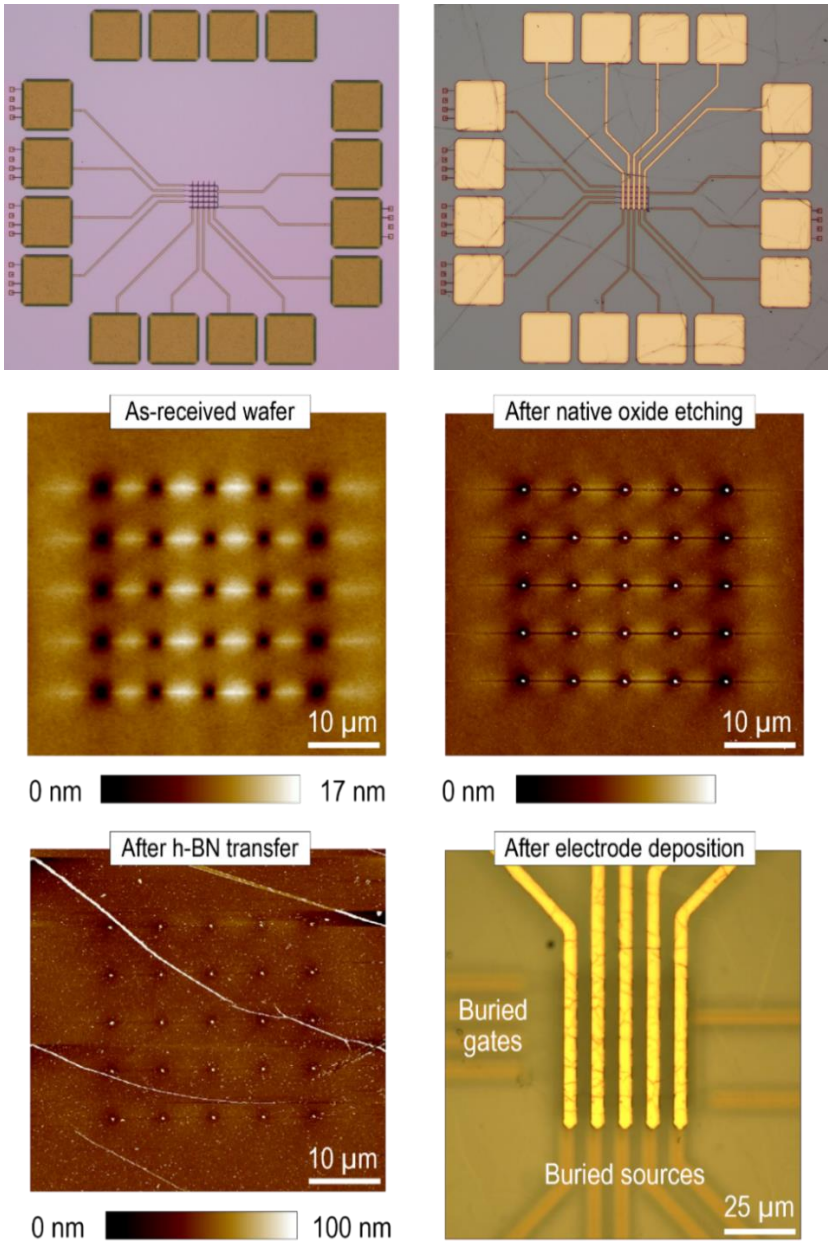


Figure 4.4 Optical images of as-received and integrated 1T1M crossbar array

Figure 4.5 shows a high-angle annular dark-field (HAADF) cross-sectional scanning transmission electron microscope (STEM) image of a 1T1M cell in the crossbar array. Although the technology node is 180 nm, the transistor size is specially designed with a width of 1 µm and length of 0.5 µm to achieve a high current for memristor operation. The correct layered structure of the h-BN stack is confirmed before and after transfer via cross-sectional TEM. Nano-chemical analyses via electron energy loss spectroscopy demonstrate the correct h-BN film composition, and also reveal that the 3-nm-thick Ti layer is rich in oxygen; this indicates that this layer may have probably formed TiO_X to some extent, which could have some effect on the electrical characteristics of the devices. The inset, which is 20nm×16nm, shows a cross-sectional transmission electron microscopy image of the Au/Ti/h-BN/W memristor on the via; the correct layered structure of h-BN can be seen. Some typical local defects of CVD-grown h-BN can be observed which contribute to the confined resistive switching region, as explained in Chapter 3.

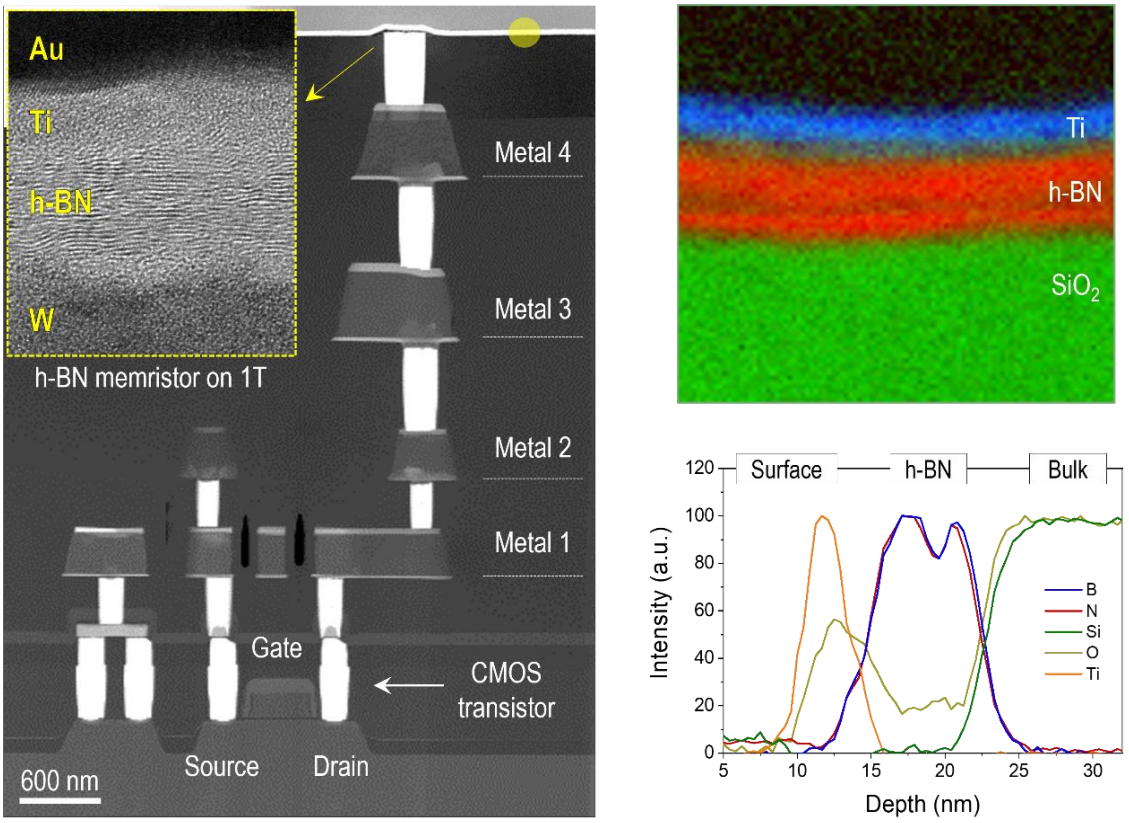


Figure 4.5 HAADF STEM image of a 1T1M cell in the crossbar array, and corresponding TEM, EELS mapping and profile.

4.3 Electrical characterization of the hybrid h-BN/CMOS-based chip

4.3.1 Isolated memristor

Next, we characterize the electronic properties of isolated memristors and compare them to the 1T1M cells. Figure 4.6 shows that, after a forming process at ~2V, the isolated Au/Ti/h-BN/W memristors exhibit non-volatile bipolar resistive switching (RS) with low reset voltage (V_{RESET}) -0.668±0.040V and set voltage (V_{SET}) 0.599±0.048V. The variability is quantified by calculating the coefficient of variance (C_V), i.e., the mean value (μ) divided by the standard deviation (σ), and it gives 8.05% for V_{SET} and 5.94% for V_{RESET}. Compared to other memristors of the same size these values are amongst the best[15-17], although similar values have been achieved in memristors of larger size. The switching mechanism in this type of device is filamentary, as demonstrated by: the high conductance of the device above the quantum conductance (G_0, 7.748×10^{-5}S), the abrupt reset, and the good fitting to the quantum point contact model. The conductive filament is formed by the penetration of Ti^{X+} ions from the top Au/Ti electrode into the h-BN when a positive bias is applied — note that in metal-oxides the movement of oxygen ions can form a metallic path, but this cannot happen in h-BN because it does not contain metallic atoms. However, the endurance of the isolated h-BN memristors is very limited (i.e., ~100 cycles), as in many other studies, in great part due to the poor controllability of the current across the memristor and the overshoot produced during the dielectric breakdown event.

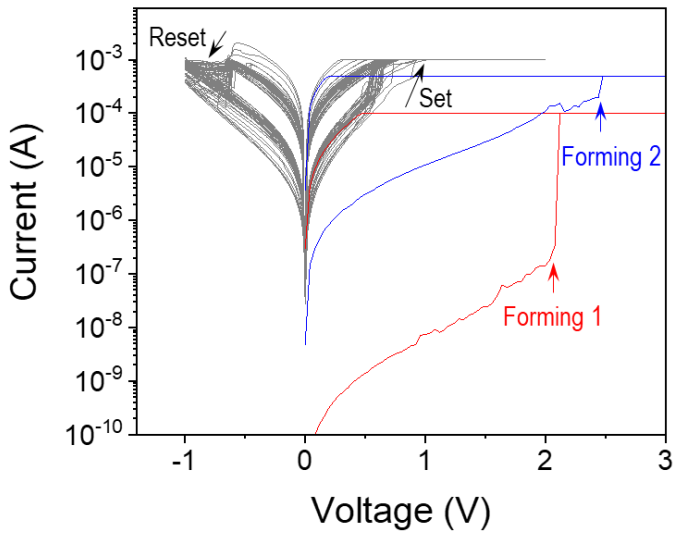


Figure 4.6 Non-volatile bipolar RS for Au/Ti/h-BN/Au standalone memristor.

4.3.2 Isolated 1T

The CMOS transistor in the 1T1M cell can precisely control the current across the Au/Ti/h-BN/W memristor and avoid the current overshoot caused during the dielectric breakdown, which results in outstanding performance. Figure 4.7 shows the optical image of a standalone transistor before (up) and after (down) its fabrication. First, we obtain the output characteristic of the standalone CMOS transistor by applying a constant voltage to the gate (V_G) and ramped voltage stresses (RVS) to the drain (V_{DS}), and measuring the drain-to-source current (I_{DS}); the CMOS transistor works correctly as expected with current level 600 µA under V_G 4 V.

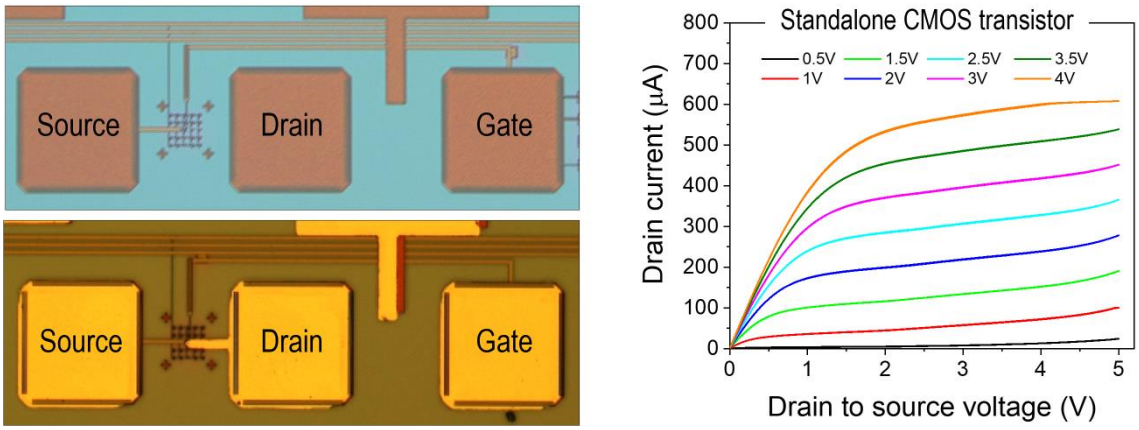


Figure 4.7 Optical image of a standalone transistor and its output characteristic

4.3.3 h-BN/CMOS based 1T1M cells

We measure the 1T1M cell by applying RVS at the top Au/Ti electrode of the memristor while keeping the source terminal of the transistor grounded and simultaneously applying a constant V_G. When using V_G=0.5V we observe volatile unipolar RS with V_{SET}=6.39±0.14V and V_{RESET}=4.13±0.08V (see Figure 4.8); hence, the cycle-to-cycle variability of the switching voltages is extremely low, i.e., the C_V of V_{SET} and V_{RESET} are 2.19% and 1.93% respectively. The resistances in high resistive state (R_{HRS}) and low resistive state (R_{LRS}) are high (~10^{10} and ~$10^6\Omega$, read at 5V) — which is good to reduce the power consumption —, the R_{HRS}/R_{LRS} ratio is >10^4 and the state transitions are highly nonlinear (i.e., the slope is ~13mV/decade). All these observations indicate the formation of an unstable conductive nanofilament (CNF)

across the h-BN that relaxes at low voltages, as similar behaviours have been readily observed in multilayer h-BN, but never in ultra-thin TiO_X. The currents in both states can be well fitted using the QSM model, further supporting the filamentary switching. This volatile unipolar RS performance is very attractive for the construction of leaky integrate-and-fire electronic neurons for spiking neural networks (SNNs).

When $V_G \geq 1.5V$ the 1T1M cell exhibit typical non-volatile bipolar RS (see Figure 2h) very similar to that of the standalone memristor, but with much higher state resistances (i.e., R_{HRS} ~20 GΩ and R_{LRS} ~100 kΩ). The good fitting to the QSM model in LRS indicates that a CNF is effectively formed across the h-BN, although the lack of an abrupt set transition suggests that the RS may be related to the modulation of the width of the CNF, rather than its entire disruption.

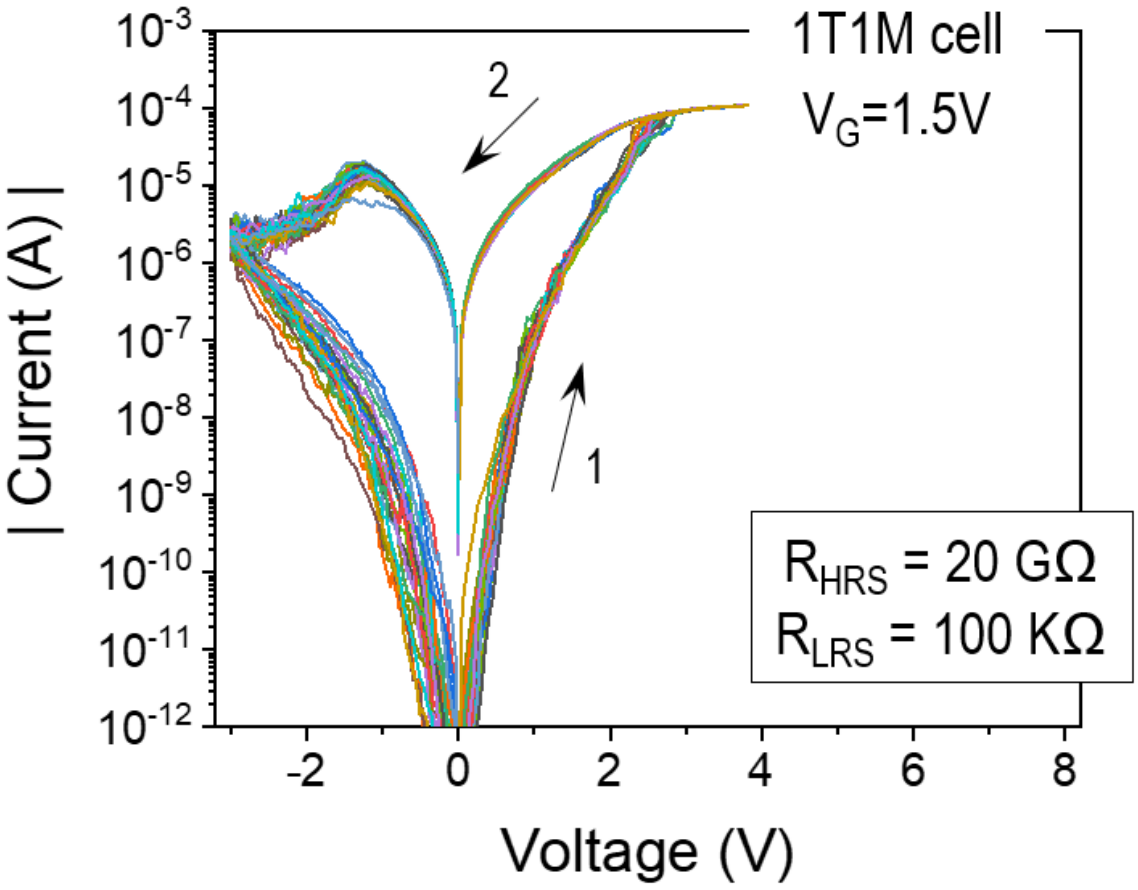


Figure 4.8 Electrical characteristics of a h-BN/CMOS based 1T1M cell

4.3.3.1 Endurance

Endurance is defined as the ability or number of cycles to rewrite data multiple times before the resistance fatigues to an unacceptable value.

The endurance plot can be extracted from I-V curves by selecting the read voltage (normally use at 0.1 V) and dividing it by the current in HRS and LRS. Hundreds of cycles are enough to demonstrate the idea.

The real endurance plot should be collected by pulse mode with minimum measurement points. Here I use B1500 as an example. One typical bipolar cycle has 4 pulses with the sequence of "Read-Write-Read-Erase" and another Read pulse from the next cycle. The first Read pulse is used to confirm the initial state of the device, and the second and third pulses are used to confirm whether the device is switched on or switched off correctly.

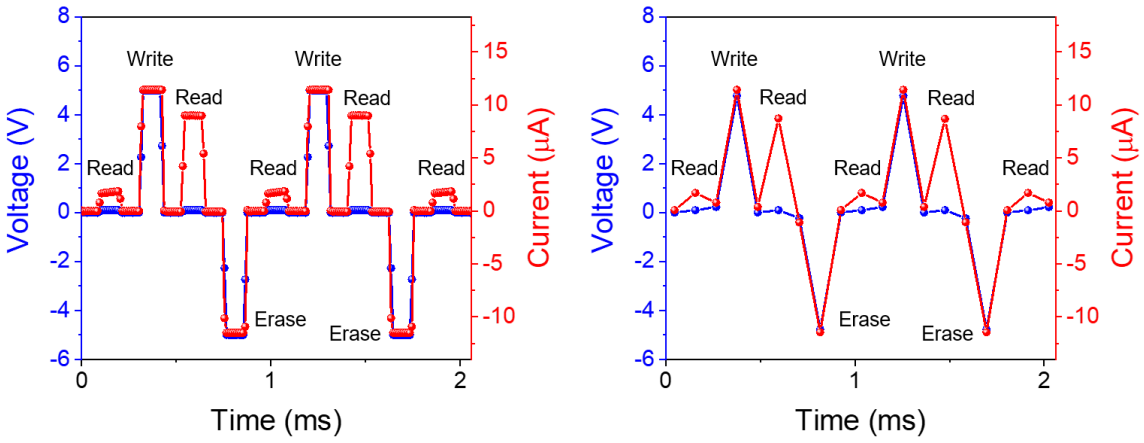


Figure 4.9 Reduce the measurement point to measure endurance

Figure 4.10 and Figure 4.11 show the pre-test and endurance results of 1T1M cells with memristor structure Au/Ti/h-BN/Au and Au/h-BN/Au, respectively. Both of them can show several millions of cycles.

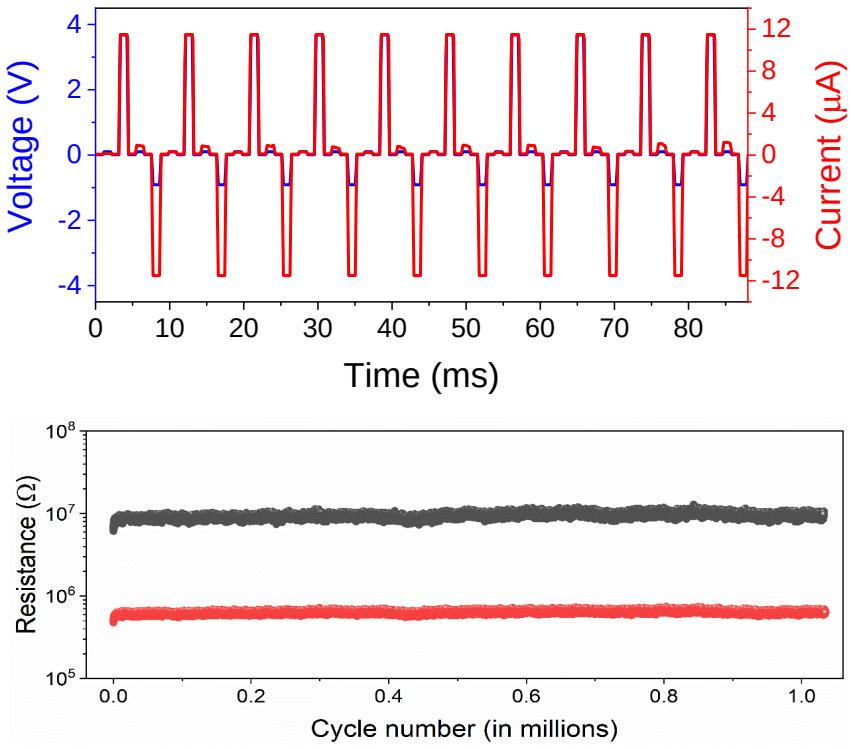


Figure 4.10 Endurance of Au/Ti/h-BN/W 1T1M cell

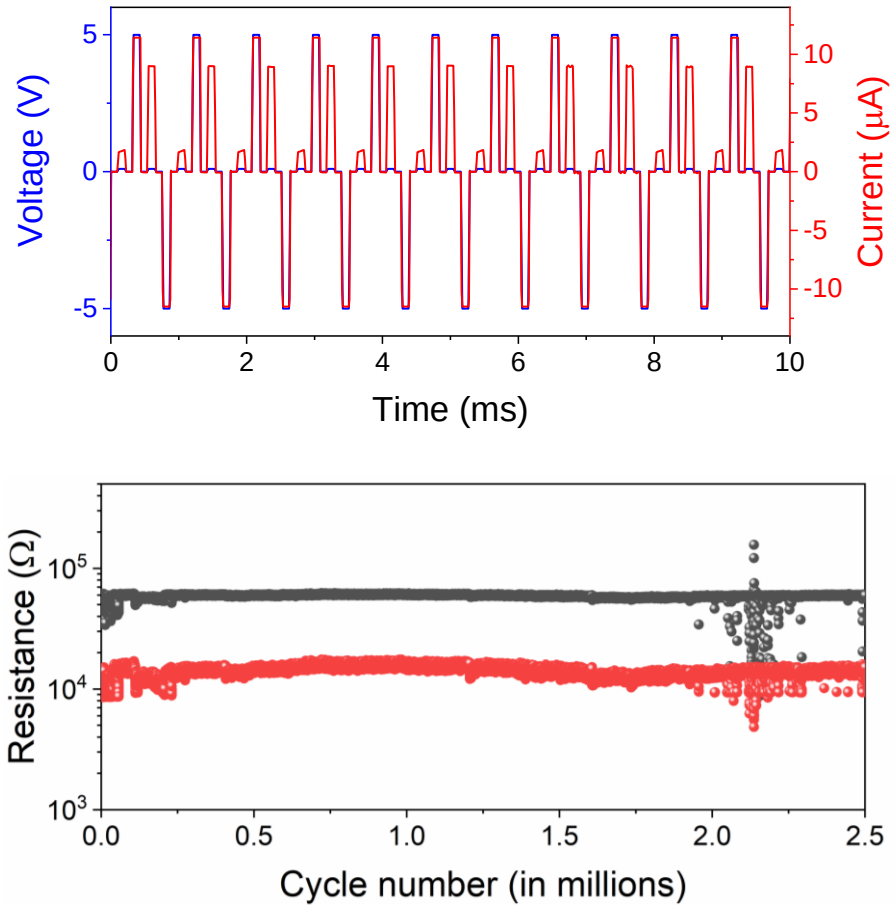


Figure 4.11 Endurance of Au/h-BN/W 1T1M cell

Note that the R_{HRS}/R_{LRS} ratio in this plot is lower than that measured in the I-V curves because in pulsed voltage stress mode the Keysight B1500A semiconductor parameter analyser only offers a current dynamic range of 3 orders of magnitude. The top plot has been collected by applying pulses of 6V (in LRS) and 0.1V (in HRS), and shows a R_{HRS}/R_{LRS} ratio of ~30. This value can be improved to ~1000 when applying 5V (in LRS) and 0.5V (in HRS). In both cases, the duration of the pulses and the interval time between them was 0.1ms, and V_G=0.5V.

If we take 0.2 million cycles as an example. If the pulse width (t_w) is 100 µs, the time for one complete cycle is 800 µs (4 pulses and 4 segments). For 0.2 million cycles, the total length of the pulse trains ($t_{0.2M}$) is 160s (200,000cycles × 8×10^{-4}s/cycle = 160s). However, the real measurement including the data saving/delay time is 2.15 hours for the old machine bought in 2015 and 0.65 hours for the new machine bought in 2021.

Table 4.1 Time considerations for endurance measurement

	t_w	$t_{0.2M}$	Real time	Real time/cycle	If 10^6	If 10^8	If 10^{10}
Old	100µs	160s	2.15h	38.7ms	10.75h	44.8days	12.3years
New	100µs	160s	0.65h	11.7ms	3.25h	13.5days	3.71years

One practical one to measure endurance is to measure all data points in millions of cycles. Beyond millions of cycles, we only apply pulse trains but don't measure them. This will save a huge amount of time since 93.2% of the measurement time (even for the most updated instrument) is wasted on the data saving and delay time. After a certain number of cycles, we measure several cycles to confirm whether they are still working or not. [182]

4.3.3.2 Switching time

Switching time is defined as the time needed to switch from HRS to LRS, or LRS to HRS.

For pulse width ≥ 1 μs, the WGFMU mode of B1500 has enough resolution to distinguish the resistance change in the middle of the write pulses. Confirmed by the first read pulse, the device starts from HRS and shows a sudden current increase after applying to write pulses. The set time (t_{SET}) is defined as the time from half of the write pulse to half of the current increase. More measurement points are suggested to see clearly the shift between states. Compared with Au/Ti/h-BN/W 1T1M cell, Ag/h-BN/W 1T1M cell shows a much faster switching time with a smaller operation voltage.

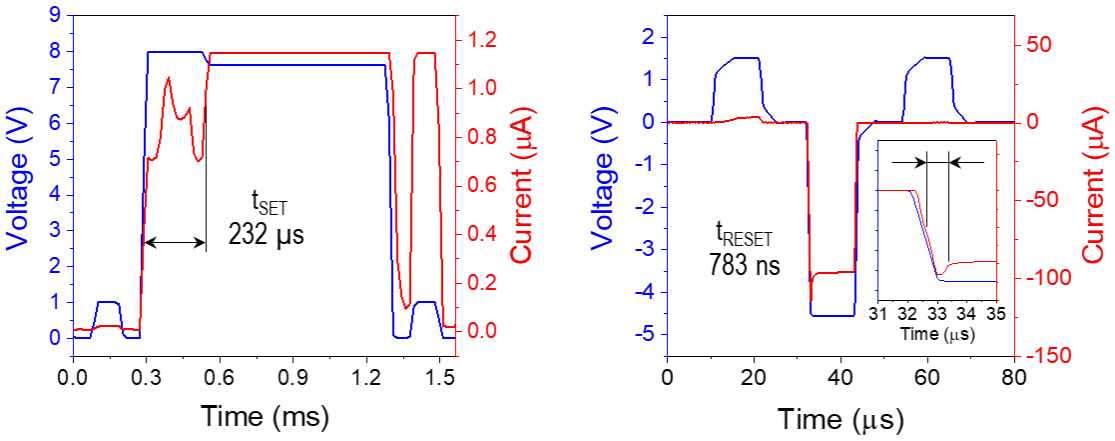


Figure 4.12 Switching time of Au/Ti/h-BN/W 1T1M cell

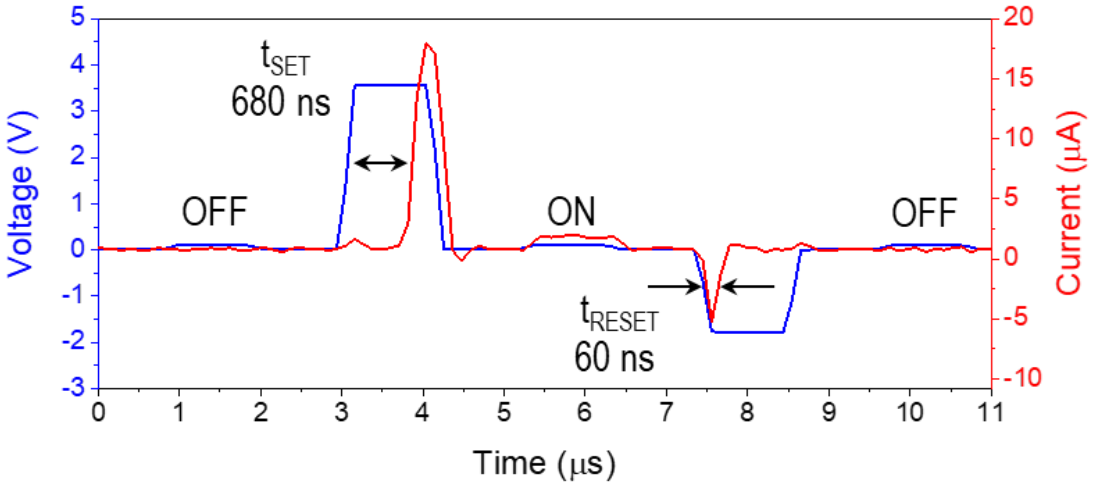


Figure 4.13 Switching time of Ag/h-BN/W 1T1M cell

For pulse width < 1 µs, the settling time problem get much more critical. Even for a high current measurement range of 1 mA, it takes the current 300 ns to be stable. One ex-situ method is used to evaluate the switching time. We first use I-t with a small constant bias V_{READ} to measure the initial resistance. Then apply a pulse with a short width, in this case, 10 ns. After the pulse operation, we measure the resistance again. If the device can be switched properly, then the switching time should be smaller than the applied pulse width.

Figure 4.14 shows an example of using ex-situ method to measure the switching time of Au/h-BN/Au 1M cell. From the I-t tests before and after a 10 ns pulse operation, we can see the device was successfully reset back to HRS, although we cannot see the in-situ transition from the step 2 measurement.

Also from the truth that the h-BN based 1M cell can be switched by a 10 ns wide pulse, the µs or hundreds of ns switching time of 1T1M cell could be slower by the transistor and memristor integration. Better industrial level integration could help to increase the operation speed.

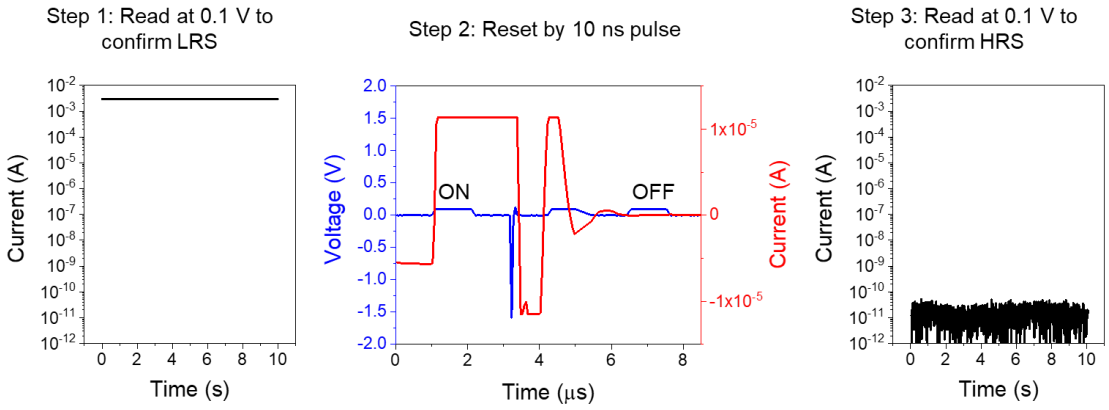


Figure 4.14 Switching time of Au/h-BN/Au single memristor cell using ex-situ method

It is worth noting that the minimum pulse width can be applied by B1500 with WGFMU is 10 ns. The characterization of shorter switching time needs special tools including a picosecond pulse generator to apply pulse stress and a high bandwidth oscilloscope (bandwidth at least 2 GHz. 20 GHz is recommended) to measure the switching time [183-185].

4.3.3.3 Retention

Retention is defined as the ability to store data or resistance after the operation stress was removed for a period of time at a specified temperature.

Current versus time pot collected at a constant voltage of 0.1V with room temperature after reset processes at different voltages, showing that multiple states can be programmed and that they are stable over time. V_G=1V.

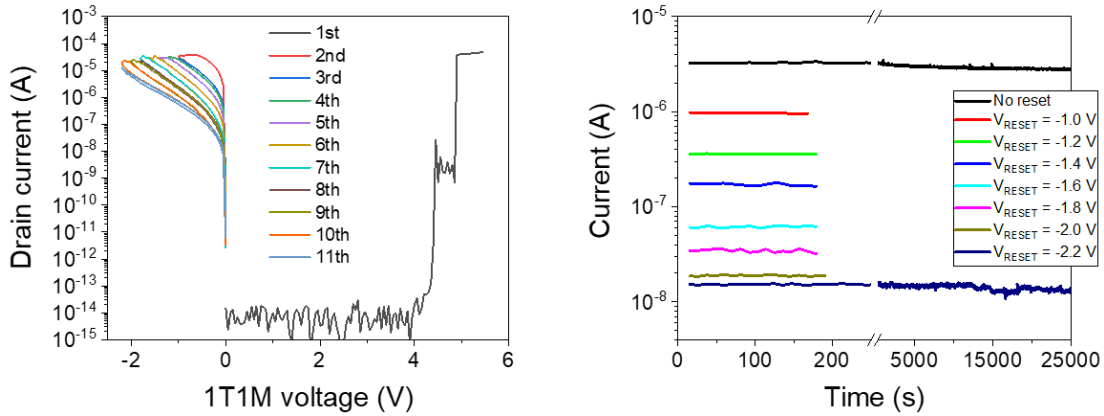


Figure 4.15 Retention of multilevel Au/Ti/h-BN/Au 1T1M cell

The desired data retention for RS-based NVM technologies is 10 years at 85 ℃ [44]. Here I also add the retention of the LRS read at 85 ℃. The device shows large read noise after 2,000 seconds. It is worth noting that even after 6,000 seconds, the device just shows large noise without fully degrading to HRS. The right plot is measured immediately after the left measurement, further confirm the LRS is not degraded. The noise could originate from the tip engagement or integration between the transistor and the memristor.

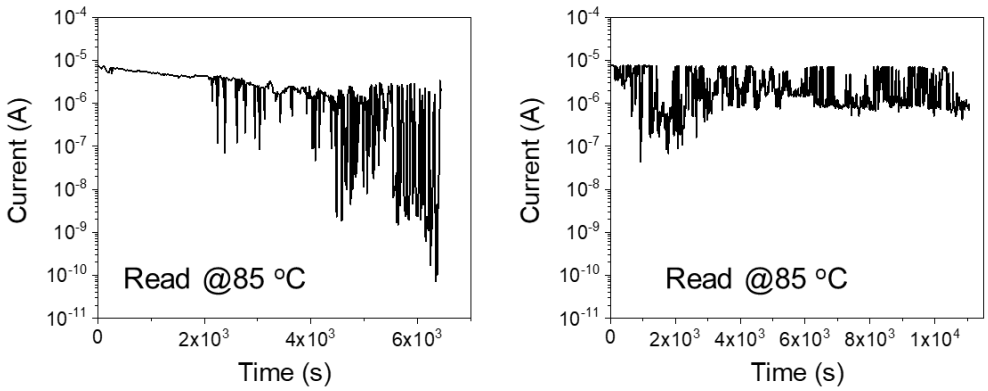


Figure 4.16 Retention of LRS read at 85 ℃

4.3.3.4 Threshold switching behaviour

Here we also measure the threshold endurance data of the Au/Ti/h-BN/Au 1T1M cell. Figure 4.17 shows the volatile switching behavior in both dc and pulse mode. With V_{READ} of 0.1V, the window is cleaner, but the on-off ratio is smaller. However for the high resistive state. we are just measuring the noise level. With V_{READ} of 0.5V the window will be enlarged.

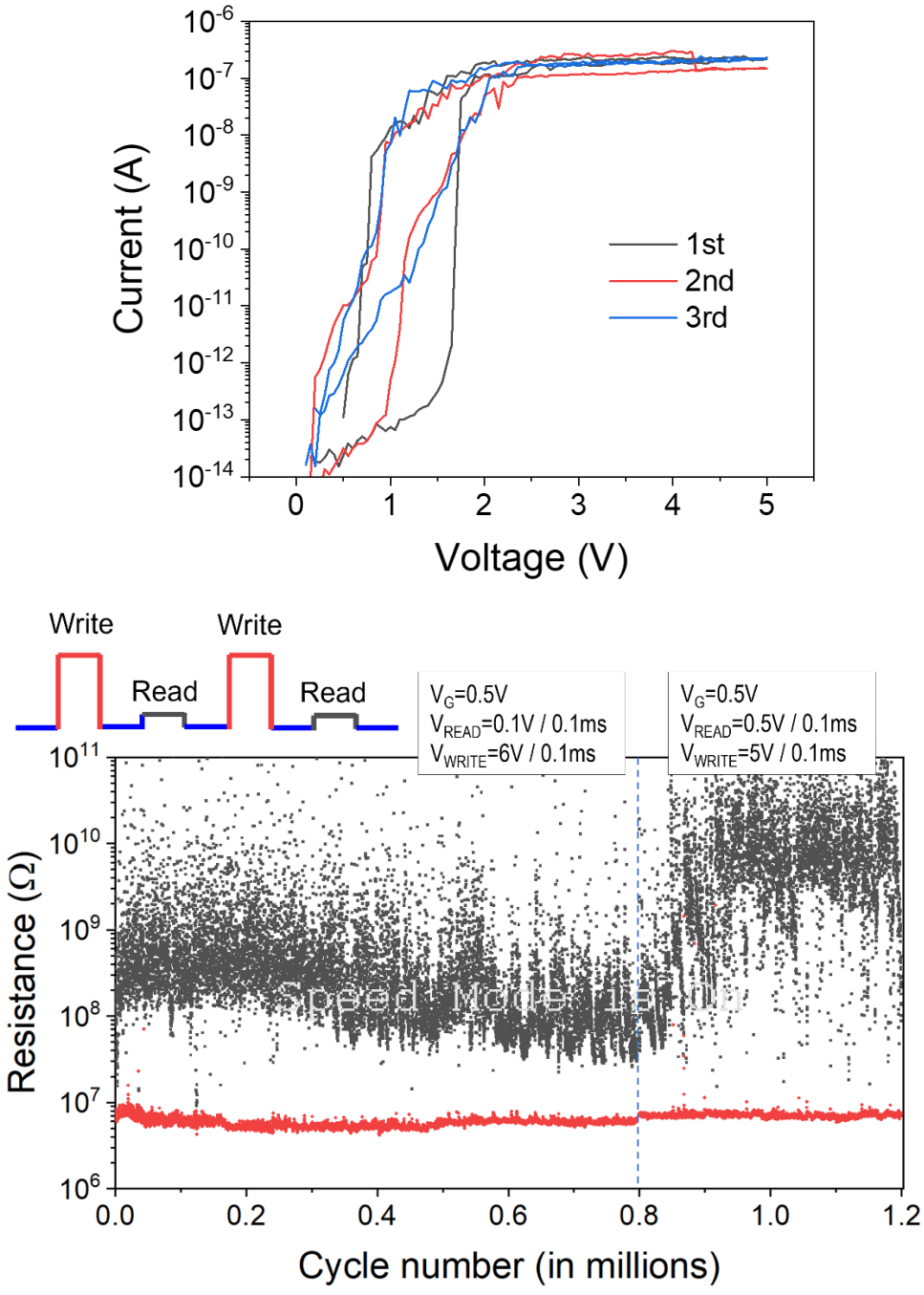


Figure 4.17 Volatile RS in the h-BN based 1T1M cell under the V_G 0.5 V

4.3.3.5 Multilevel switching capability

For Au/Ti/h-BN/W 1T1M cell, the values of R_{HRS}, R_{LRS} and R_{LRS}/R_{HRS} can be accurately controlled in three different ways: by tuning the duration of the write pulse (see Figure 4.18), by tuning the amplitude of the write pulse (Figure 4.19 left), and by tuning the amplitude of the erase pulse (see Figure 4.19 right).

The first two methods adjust the value of R_{LRS}, while the third one adjusts the value of R_{HRS}. This allows us to program multiple resistive states between the most conductive R_{LRS} and the most resistive R_{HRS} and to achieve a nearly analogue transition between them. Each state can have a high endurance of millions of cycles, and in total, 6.5 million cycles have been measured from the same Au/Ti/h-BN/W 1T1M cell between different resistance levels.

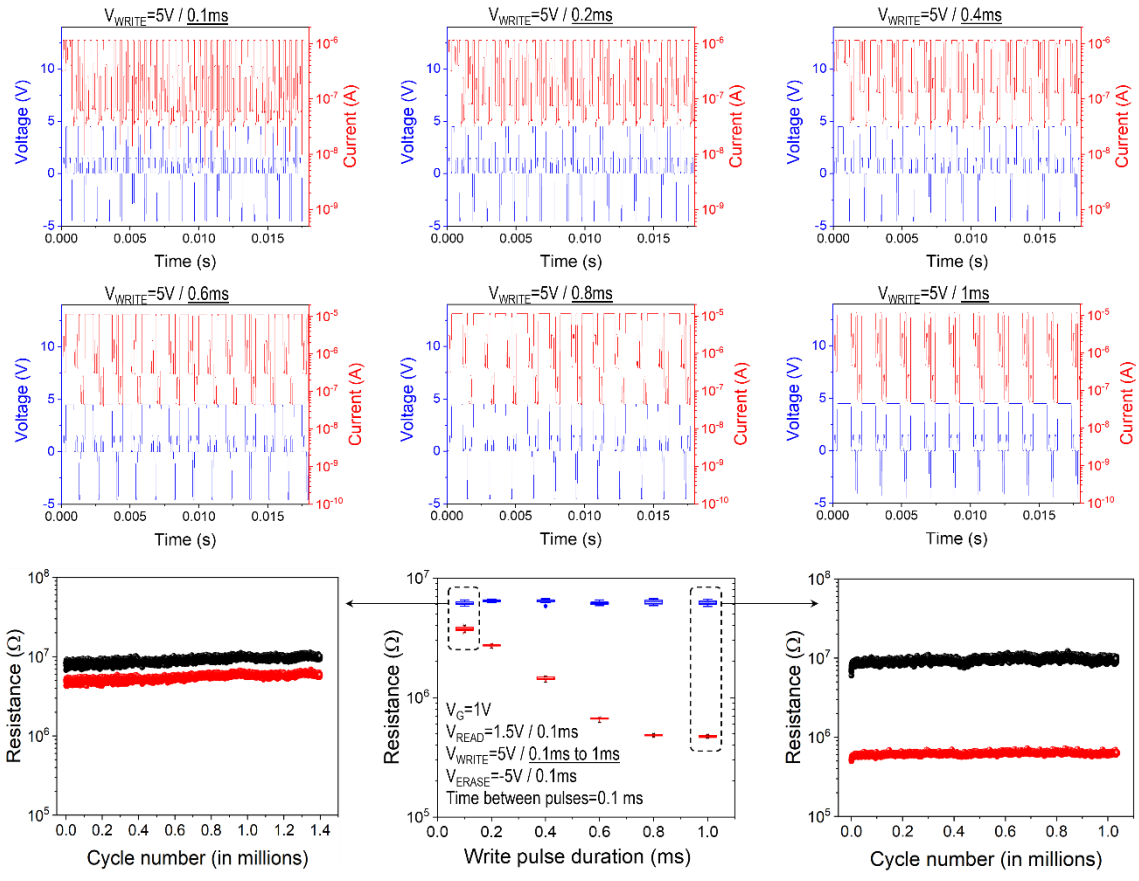


Figure 4.18 Different write pulse widths are selected to enlarge the on-off ratio and corresponding endurance

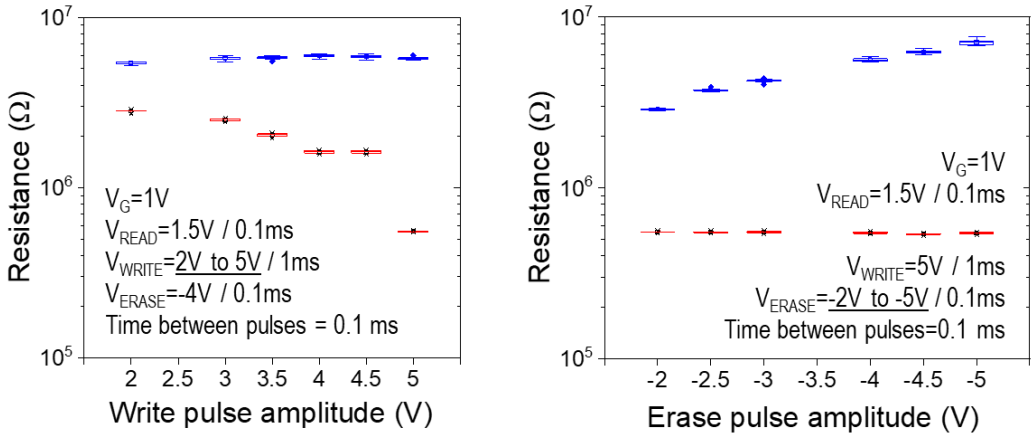


Figure 4.19 The resistance vs. the pulse amplitude of write and erase operation

4.3.3.6 Spike-timing-dependent-plasticity (STDP)

Furthermore, the 1T1M cells with Au/Ti/h-BN/W memristors exhibit spike-timing dependent plasticity (STDP) when applying PVS displaced in time at the input and output. This non-volatile RS performance is very attractive to construct electronic synapses for SNNs; this type of networks has become very attractive because they consume less energy than traditional deep neural networks, due to the method in which information is codified (i.e., as short voltage spikes instead of constant voltages).

The measurement of the STDP adopts the methodology published by M. Maestro-Izquierdo [187]. Here I show the detailed voltage programming sequence to induce STDP functionality in the CMOS/h-BN based 1T1M cell. The device is first initialized to the same conductance level, which is recorded as G_{pre}. Then apply the combined pulses shown in Figure 4.20 starting from Δt = -1 ms. The grey, red and blue lines represent pre-, post and combined spikes. Pre- means the spike is applied on the top electrode of the memristor, while post- means the spike is applied on the source electrode of the transistor. For simplicity, we applied the combined spike on the top electrode, and keep the bottom electrode ground. The time delay Δt is defined by $\Delta t = t_{post} - t_{pre}$. When $\Delta t > 0$, pre-spike comes before post-spike. When $\Delta t < 0$, pre-spike comes after post-spike. Then the final state or conductance will be read and recorded as G_{post}. Then repeat the initiation-program-read process described above but with an increased Δt with a step of 0.1 ms or 0.2 ms. The conductance change ΔG is defined as $\Delta G = (G_{post} - G_{pre})/G_{pre}$.

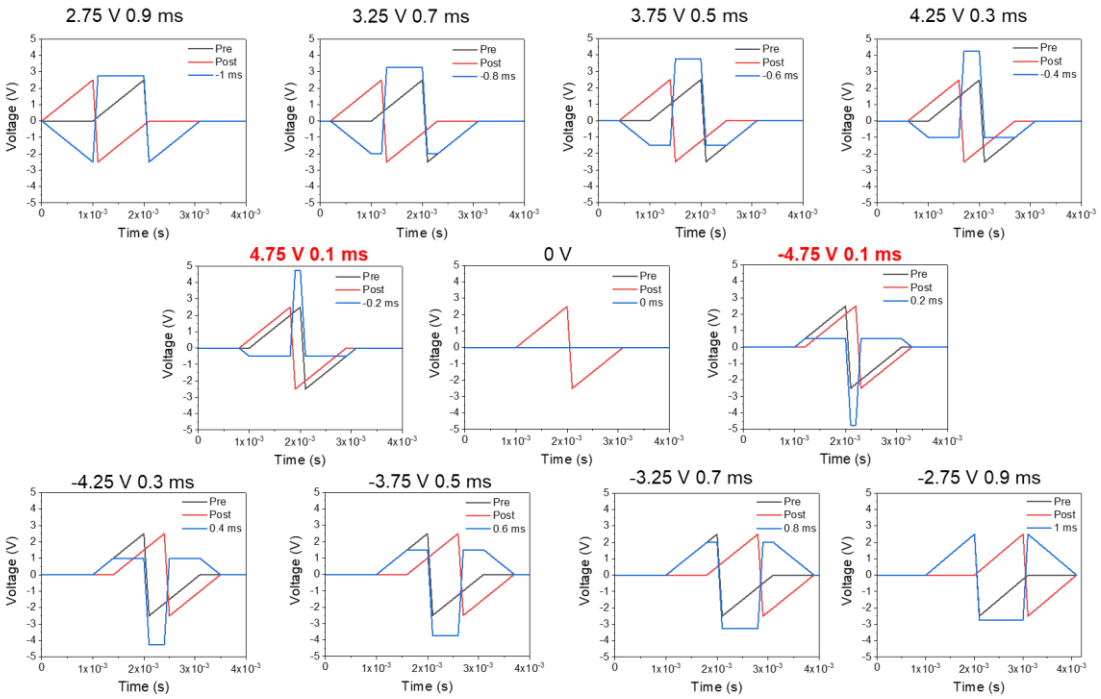


Figure 4.20 Detailed voltage programming sequence to induce STDP functionality in the h-BN based 1T1M cell

Figure 4.21 shows the STDP of the Au/Ti/h-BN/W 1T1M cell. The same initial conductances are plotted under the right y-axis. The asymmetric STDP characteristic of the h-BN based 1T1M cell is fitted with the STDP model presented in the inset. The fitting parameters are: A_-=-0.03x10^{-6}, A_+=0.21x10^{-6}, τ_+=0.35 msec. and τ_-=0.5 msec. To account for the device-to-device variability, we included a 20% variability in the previously mentioned A_+ and τ_+ parameters.

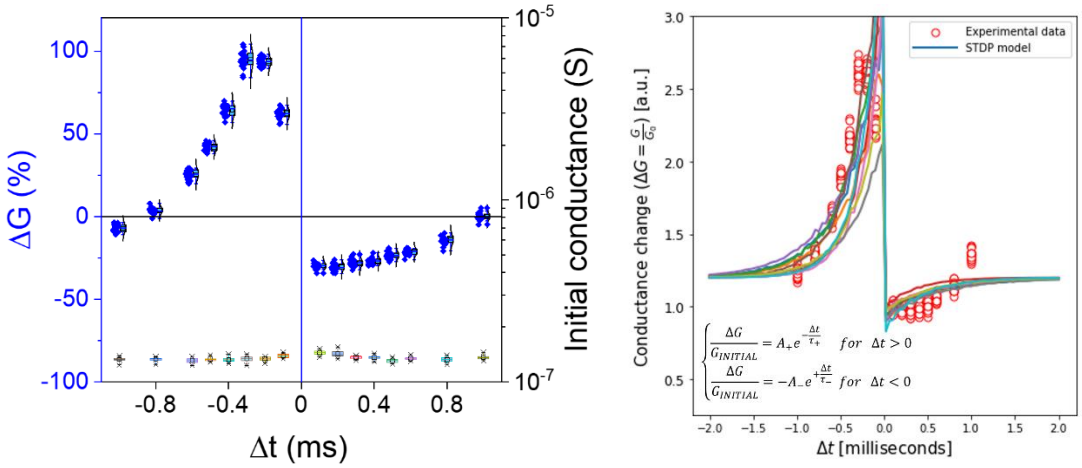


Figure 4.21 STDP of the Au/Ti/h-BN/W 1T1M cell

Based on the above-measured performance metrics, we emulate the performance of a memristive SNN for image classification. To do so, first we fit the measured STDP (including the device-to-device variability) data from Figure 3i considering a piecewise exponentially decaying model to implement the learning rule (i.e., STDP), and to enable an unsupervised learning scheme to evaluate its main figures-of-merit as it classifies the images from the Modified National Institute of Standards and Technology (MNIST) database of handwritten digits. A schematic representation of the SNN is displayed in Figure 4.25. To study the impact of the device variability we considered a Monte-Carlo approach and repeated the training 50 times for each SNN. For each Monte-Carlo run, as the training progresses, the synapses connected to each of the excitatory neurons learn the general features of a given pattern. This is shown in Figure 4b, where each slice presents the 313,600 synapses arranged in 400 groups (20×20, i.e., the numeric patterns) of 784 synapses each (28×28, i.e., the pixels that form each numeric pattern); the synapses connect the input layer to the excitatory layer (green and red spheres in Figure 4a, respectively). The red square in Figure 4.25b indicates a group containing the 784 synapses that connect the input neurons to the first neuron of the excitatory layer. For instance, this particular neuron progressively learns the representation of the digit '1'. Due to the random initialization of the synaptic conductance and the device variability, the patterns learned by each neuron change among the different Monte-Carlo runs. The accuracy was evaluated for every 1000 images by presenting the MNIST test dataset. The behaviour of the SNN is verified by the corresponding confusion matrix obtained after the complete training dataset (the case of 400 excitatory neurons is presented in Figure 4.25c). The evolution of the classification accuracy during training (averaged over 50 Monte Carlo runs and indicating the standard deviation) is presented in Figure 4.25d, and it shows that the inference accuracy increases as the number of excitatory-inhibitory neurons increases. The best average accuracy reaches ~90%, which is a very high value considering the simplicity of the SNN, its similarity to biological neural networks (i.e., the learning rule employed is STDP), and the unsupervised training protocol.

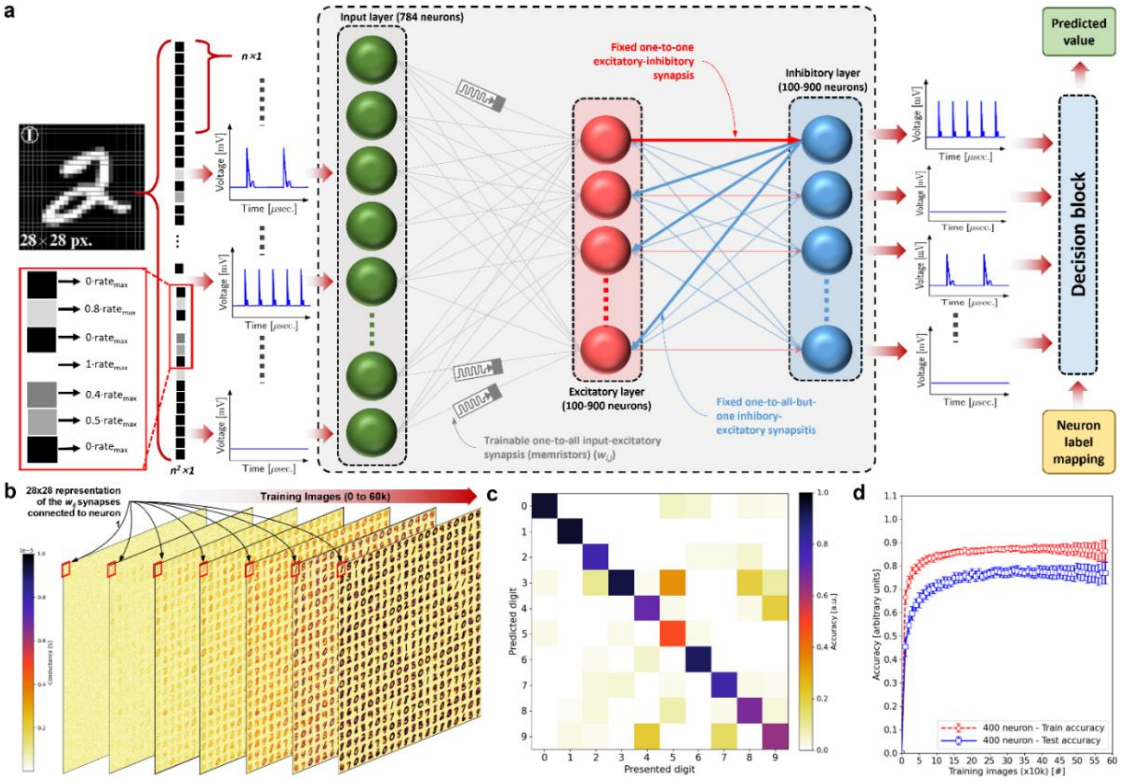


Figure 4.22 Implementation of a spiking neural network using h-BN based 1T1M cells

a, structure of the considered SNN. Each MNIST image is reshaped as a 784×1 column vector, and the intensity of the pixels is encoded in terms of the firing frequency of the input neurons. The only trainable synapses are those connecting the input layer with the excitatory layer, and they are modelled with the STDP characteristic of the CMOS/h-BN based 1T1M cells. The learning is unsupervised, and the neurons are labelled only after the training. These label-neuron assignments are then feed to the decision block altogether with the firing patterns of the neurons, to infer the class of the image presented in the input. b, evolution of the synaptic connections between the input and excitatory layers during training for the case of 400 excitatory/inhibitory neurons. The red square identifies 784 synapses arranged in a 28×28 representation. c, Confusion matrix indicating the classification accuracy for each class from the dataset. d, Classification test and training accuracy as a function of the number of presented training images for the neural network comprising 400 excitatory/inhibitory neurons. The error bars show the standard deviation for 50 Monte Carlo simulation runs for every accuracy point.

4.4 Matrix operation

The logic “OR” is implemented in Figure 4.23. Here I apply ramped voltage stress on the Drain 1 (D1) or top electrode side, and keep the Source 1 (S1) or bottom electrode grounded to control the resistive state of cell 1 and cell 2. “0” represents HRS, while “1” represents LRS. Gate 1 and Gate 2 are always on with constant a bias of 2 V. The output current “I” is the current measured at the drain or source.

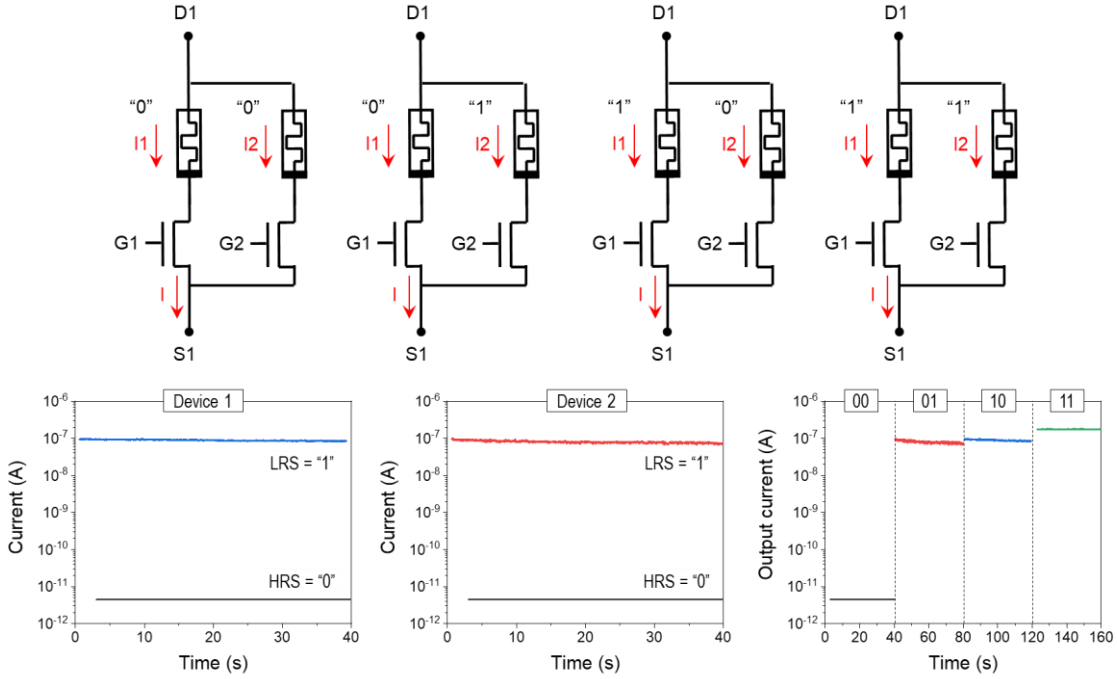


Figure 4.23 “OR” logic of 2 1T1M cells

To implement the logic “OR” and “IMP” operation, all sources are collected together by dry transferred Au thin strip (250 μm long, 50 μm wide), highlighted by the red box in red. A pre-test has been done to confirm the good electrical connections between different sources.

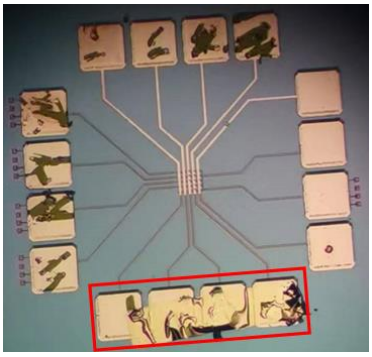


Figure 4.24 Connect the sources to fulfil matrix operation

The logic "IMP" is implemented in Figure 4.23. For IMP logic 2 parallel connected memristors are needed. The voltage applied on D2 (V_{D2}) is large than V_{SET}, while the voltage applied on D1 (V_{D1}) is slightly smaller than V_{SET}, and $V_{D2} - V_{D1} < V_{SET}$. Device 2 can only be set to LRS when both Device 1 and Device 2 are in HRS. The output is the resistance of D2 after the IMP operation. If Device 2 is LRS and Device 1 is HRS, device 2 will stay in LRS. If Device 2 is HRS and Device 1 is LRS, the voltage drop in Device 2 equals $V_{D2} - V_{D1}$, which is smaller than the V_{SET}. So Device 2 stays in HRS. When both Device 1 and Device 2 are LRS, Device 2 will stay in LRS. [189]

For the real experiment, the pre-test shows the V_{SET} is around 7V, so I applied ramped voltage sweep from 0 to 10V for Drain 2, and I chose ramped voltage sweep from 0 to 5V for Drain 1. The connected common source is grounded. Gate 1 is always on by connecting a DC power supply with a constant voltage of 2V.

The output current "I" is the current flows isolated in Device 2, measured after the IMP operation, which also represents the resistance level of Device 2.

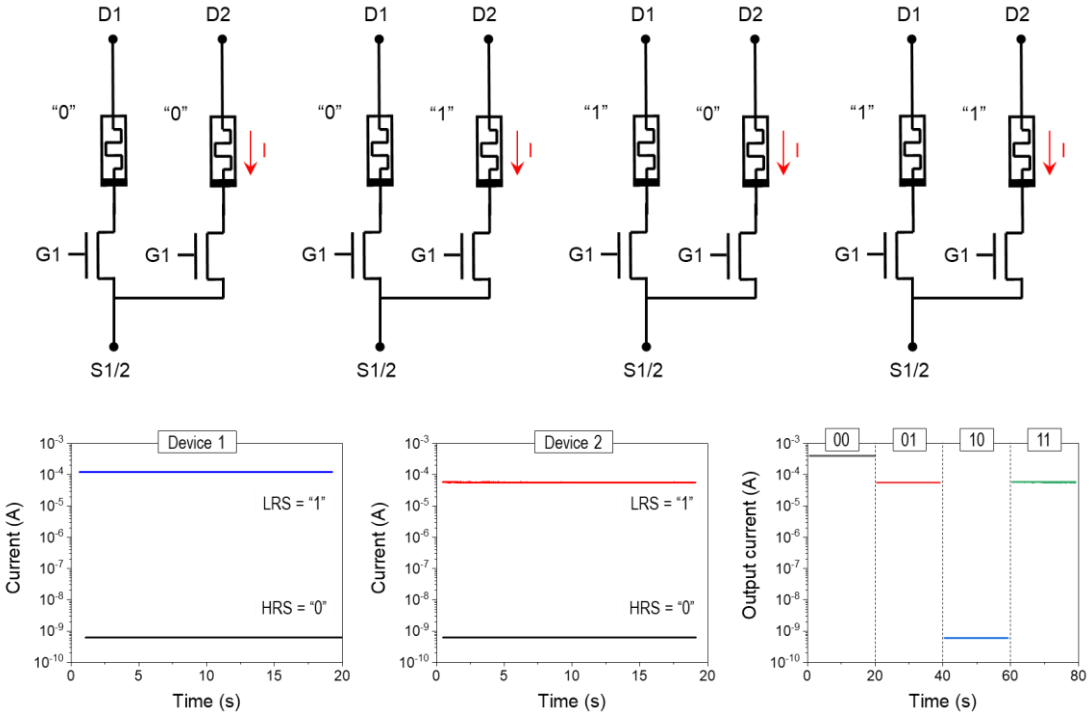


Figure 4.25 "OR" logic of 2 1T1M cells

4.5 Proposed circuit implementation of a neuron-synapse-neuron block combining h-BN based 1T1M cells and CMOS circuitry.

Aiming to implement a hardware-based SNN accelerator exploiting the capabilities of our h-BN/CMOS based 1T1M cells, we propose the CMOS circuit shown in Figure 4.26a for emulating the electrical response of a biological neuron, which is capable of accounting for the adaptative firing threshold and the refractory period after firing. The entire circuit is simulated in SPICE and its response is calculated — SPICE is the industry-standard software, which considers the response of the electronic devices plus the non-idealities of the circuit (i.e., parasitic capacitances, line resistances)[47]. The correct response of the circuit is demonstrated by the presynaptic and postsynaptic traces and the evolution of the membrane potential, presented in Figure 4.26b-c, respectively. Since the accuracy increase when going from 400 to 900 neurons is not too large, a trade-off between improving the inference accuracy and reducing the circuit complexity should be considered to limit the hardware implementation of this network to 400 excitatory-inhibitory neurons.

Overall, the electrical characteristics of the h-BN memristors connected to a CMOS transistor are by orders of magnitude superior to those of standalone h-BN memristors and h-BN memristors connected to 2D-materials-based transistors. Compared to commercial metal-oxide and phase-change memristors, the h-BN/CMOS based 1T1M cells exhibit a similar endurance, but with the advantage of a much higher controllability of the current, up to the point that we can switch between operation regimes (volatile unipolar and non-volatile bipolar) and select the switching mechanism (charge trapping and de-trapping or filament formation and disruption). Such versatility is beneficial because it could considerably simplify the fabrication of memristive neural networks, as the h-BN/CMOS based 1T1M cells can be employed as both electronic neurons and synapses depending on the value of V_G used. Moreover, the values of R_{HRS} and R_{LRS} are higher than in standalone memristors (reducing the power consumption when integrated in high-density circuits), and the crossbar arrays of 1T1M cells do not suffer from sneak path currents because they can be completely supressed by fixing

V_G=0V. The fact that we solved the problem of endurance in miniaturized 2D-materials-based memristors enables their use in other applications, such as radiofrequency switches and data encryption. The electrical behaviours here presented have been observed in 32 out of 40 devices, with an overall the yield of 80%. The 8 devices that did not work showed a very high resistance and could not be set even after applying 20V, indicating the presence of contamination, probably polymers from the transfer. Fabrication in an industrial facility using optimized processes should help to improve the yield and the device-to-device variability. Our study is an important step towards the integration of 2D materials in silicon microchips, and could inspire other scientists to study 2D-materials-based devices integrated in a real microchip with functional circuitry, not isolated on a SiO_2/Si or Al_2O_3/Si substrate.

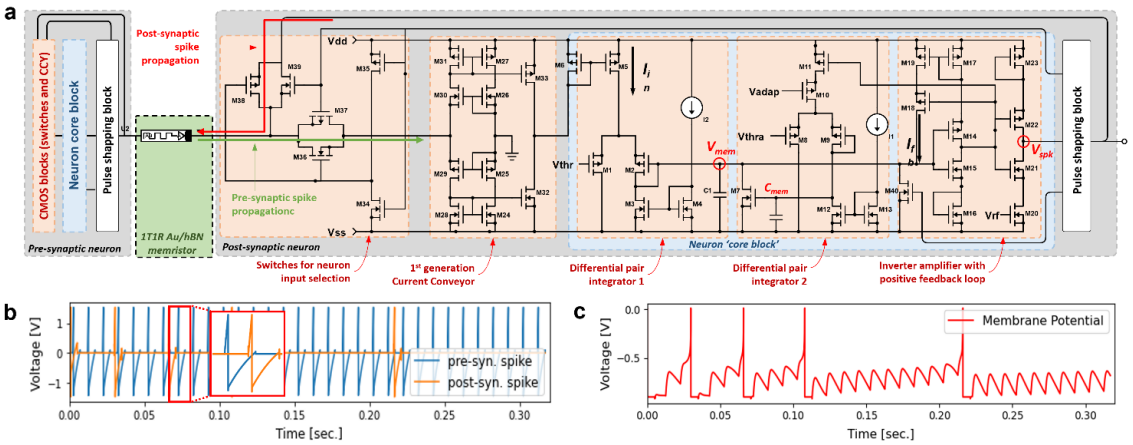


Figure 4.26 Proposed circuit implementation of a neuron-synapse-neuron block combining h-BN based 1T1M cells and CMOS circuitry.

a, Circuit schematic of the proposed CMOS neuron and how it is connected to the synapse. The colors indicate the complete neuron (grey surrounding box), the core block (light-blue box) and the individual building blocks (light-red boxes). b, SPICE simulation of the pre- and post- synaptic signals applied to the CMOS/h-BN based 1T1M. c, SPICE simulation of the neuron's membrane potential. The firing events progressively separate from each other due to the adaptative firing threshold.

4.6 Conclusions and prospect

Overall, the electrical characteristics of the h-BN memristors connected to a CMOS transistor are by orders of magnitude superior to those of standalone h-BN memristors and h-BN memristors connected to 2D-materials-based transistors. Compared to commercial metal-oxide and phase-change memristors, the h-BN/CMOS based 1T1M cells exhibit a similar endurance, but with the advantage of a much higher controllability of the current, up to the point that we are able to switch between operation regimes (volatile unipolar and non-volatile bipolar) and select the switching mechanism (charge trapping and de-trapping or filament formation and disruption). Such versatility is beneficial because it could considerably simplify the fabrication of memristive neural networks, as the h-BN/CMOS based 1T1M cells can be employed as both electronic neurons and synapses depending on the value of VG used. Moreover, the values of R_{HRS} and R_{LRS} are higher than in standalone memristors (reducing the power consumption when integrated into high-density circuits), and the crossbar arrays of 1T1M cells do not suffer from sneak path currents because they can be completely supressed by fixing V_G=0V. The fact that we solved the problem of endurance in miniaturized 2D-materials-based memristors enables their use in other applications, such as radiofrequency switches and data encryption. The electrical behaviours here presented have been observed in 32 out of 40 devices (see Extended Data Figure 10), with an overall yield of 80%. The 8 devices that did not work showed a very high resistance and could not be set even after applying 20V, indicating the presence of contamination, probably polymers from the transfer. Fabrication in an industrial facility using optimized processes should help to improve the yield and the device-to-device variability. Our study is an important step towards the integration of 2D materials in silicon microchips, and could inspire other scientists to study 2D-materials-based devices integrated into a real microchip with functional circuitry, not isolated on a SiO_2/Si or Al_2O_3/Si substrate.

Chapter 5: Memristors with Initial Low-Resistive State for Efficient Neuromorphic Systems

5.1 Introduction

Developing new hardware for artificial intelligence systems is a major goal for the government of technologically advanced countries and leading companies in the field of nanoelectronics, and investments in this field are expected to keep growing exponentially in the next years. Recent studies demonstrated that crossbar arrays of metal/insulator/metal (MIM) nanocells —often called memristors— can be used to construct artificial neural networks (ANN) with unsupervised learning capability [191], and they have been already employed to accomplish sophisticated operations, such as image classification, encryption and position detection [192-193]. In an ANN, memristors are used as electronic synapses, i.e., elements that connect two neurons and that show the property of changing their conductance between different levels [194-195] —often called resistive switching (RS).

The initial conductance of most MIM-like memristors is in the range of few pico/nano-siemens with a larger set voltage or a higher resistance compared to the following cycles [43] (namely *forming memristors*), and it can be increased by applying electrical stimuli between the two metallic electrodes; this generates local defects at random locations in the microstructure of the insulator (i.e., intrinsic vacancies and/or impurities from adjacent metallic electrodes) that promote charge transport. If the density of defects in the insulator is high enough a conductive nanofilament (CNF) can be effectively formed, which drives all the current flowing between the two electrodes and results in an overall linear conductance in the range of few millisiemens. In most MIM nanocells, the defects in the insulator and the CNF are stable after the bias is removed, and the conductance of the MIM cell can only be reduced by applying additional electrical stimuli (normally with opposed polarity).

When the CNF is only partially formed, it is said that the device is operated in a high resistive state (HRS), and controlling its conductance in this regime is very

challenging because the current flows across a random number of defects spread along the volume of the insulator, and the formation of new defects results in exponential increases of conductance [196]. On the contrary, when the CNF is completely formed, it is said that the device is operated in a low resistive state (LRS), and controlling its conductance is much easier because all the current flows along the CNF, which can be accurately widened and narrowed by applying electrical stimuli of different polarities [197]. MIM nanocells operating at high conductances above quantum conductance (77.5 μS), with high linearity between current and voltage, can lead to accurate analogue computing and stable multilevel states [198].

As the weight (i.e., conductance) of electronic synapses needs to be increased and decreased progressively and in a linear manner, MIM-like memristors for ANNs are normally operated in LRS and the typical maximum/minimum conductance ratio (I_{MAX}/I_{MIN}) is ~10 (see Supplementary Table 1). Consequently, most memristive electronic synapses need an initialization step to increase their initial conductance, which is an important nuisance for the fabrication of ANNs because it increases the complexity of the programming algorithms and peripheral hardware [199]. One common solution to mitigate this problem is to fabricate forming-free memristors by introducing impurities in the insulator during the manufacturing process via ion implantation [200-201] and/or reducing the thickness of the insulating film [166], which increases the initial conductance of the devices to the HRS (i.e. an effective CNF is not completely formed). However, the device-to-device variability of forming-free memristors with initial conductance in HRS (namely *HRS memristors*) is high (see Supplementary Table 2), due to the inhomogeneous distribution of impurities in the insulating film. Such increased device-to-device variability of the initial conductance represents a problem for the training of the ANN, as it may slow down the learning process and even present convergence problems [202].

Here we report the first fabrication of memristors whose initial conductance is already in the LRS regime (namely *LRS memristors*) by using ultra-thin TiO_2 films synthesized by magnetron sputtering. We have selected TiO_2 because, compared to other metal-oxides often used as dielectric in memristors (i.e., HfO_2, TaO_X, Al_2O_3), this material has the lowest bandgap. This makes that the initial conductance is higher even if the material might not be completely shorted. For all the devices tested (more than 20), the initial electrical resistance falls within a window narrower than one order of magnitude (i.e., between 100 Ω and 714 Ω), which is much smaller than that of *HRS*

memristors and *forming memristors*. The switching thresholds are uniform enough to enable analog tuning of individual devices in a crossbar integrated circuit. More importantly, the compact design of peripheral circuits is feasible as devices are forming-free, and the maximum switching threshold is compatible with the core voltage of standard CMOS processes.

5.2 Device fabrication

Figure 5.1a shows the optical microscope image of an array of Au/TiO_2/Au cross-point memristors with lateral sizes ranging from 5 μm × 5 μm (left column) to 50 μm × 50 μm (right column), and Figure 5.1b shows a three-dimensional (3D) topographic map of a 5 μm × 5 μm Au/TiO_2/Au memristor at the cross point area, collected via atomic force microscopy (AFM). Height colour scale: black = 0 nm, white = 300 nm. Scale bar: 250 nm.

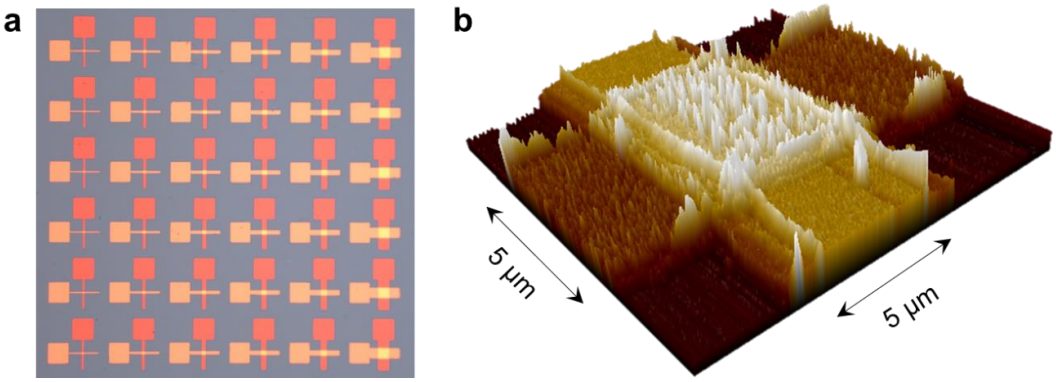


Figure 5.1 Fabrication of Au/TiO_2/Au synapses

Figure 5.2 shows an AFM topographic map of the surface of the sputtered TiO_2 film (on a 300 nm SiO_2/Si substrate). The root mean square (RMS) surface roughness is 4.6 nm, much larger than that of the SiO_2/Si substrate (<0.3 nm) [203], and much larger than that of TiO_2 films grown by atomic layer deposition (ALD) on SiO_2/Si substrates (~0.2 nm) [204]. The surface of the sputtered TiO_2 thin film is composed of tall grains with diameters 131.2 ± 58.9 nm (see Figure 5.2 and Supplementary Figure 1) separated by deep phase boundaries. Statistical analysis of TiO_2 phase size from 5 μm × 5 μm large area scan. The original image and data processing method are detailed in Supplementary Figure 3. Figure 5.2 also shows the surface profile in **c** from A to A' crossing two phases.

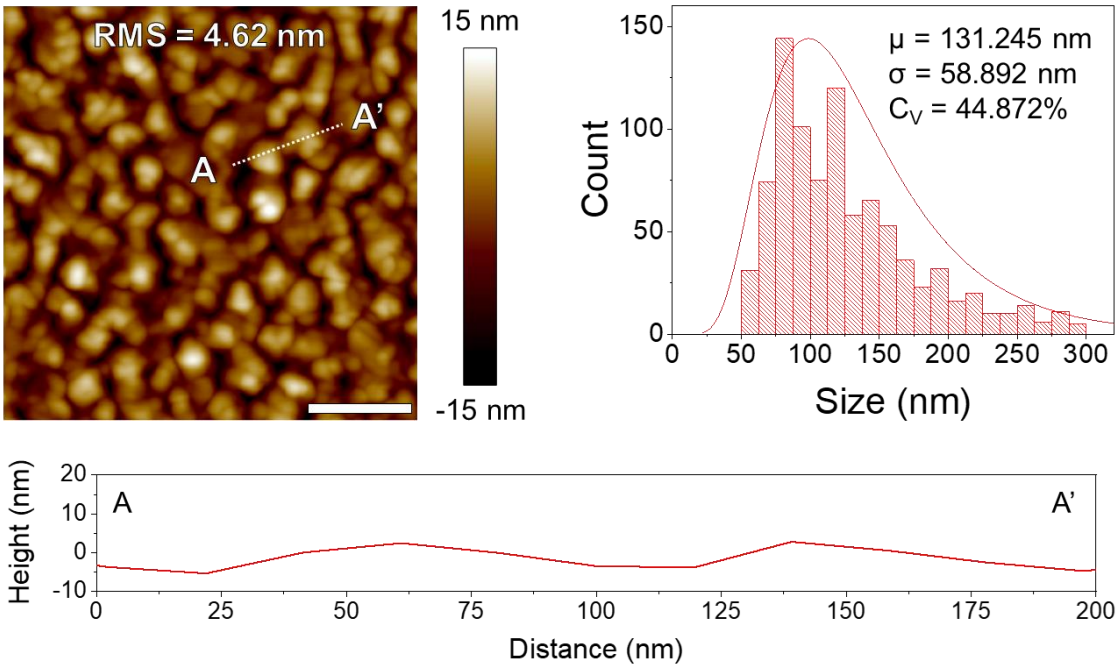


Figure 5.2 AFM topography map of the sputtered TiO_2 on 300 nm SiO_2/Si substrate

To further characterize these features, we use focused ion beam (FIB) to cut the MIM cells and fabricate thin lamellas, and analyse them using transmission electron microscopy (TEM). The cross-sectional TEM images (Figure 5.3) reveal that the thickness of the sputtered TiO_2 film is ~4.5 nm (as expected), and that it contains few-nanometres-wide regions rich in Au every ~125 nm, which correlates to the deep phase boundaries observed in the topographic AFM maps. Scale bar: 25 nm

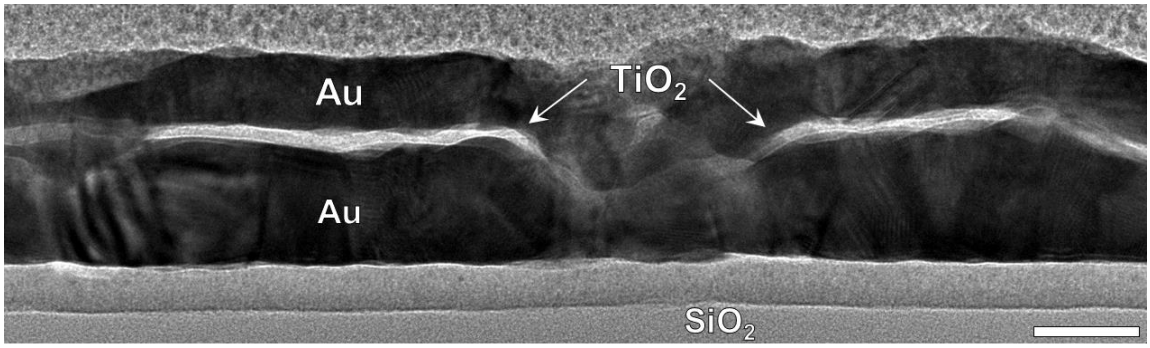


Figure 5.3 Cross-sectional TEM image of an Au/TiO_2/Au device showing two phases, for which the phases size is in the same scale with Figure 5.2.

5.3 Electrical characterization

5.3.1 Bipolar switching

When applying sequences of ramped voltage stresses (RVS), the devices always show initial conductance of ~5 mS (typical of LRS regime) without the need of a forming process, followed by a conductance decrease to HRS regime at ~0.8 V (i.e. reset process). This behaviour is observed independently of the polarity of the first RVS (see Figure 5.4a-b), which is consistent with the symmetric structure of the MIM cell (i.e. Au/TiO_2/Au). Figure 5.4c shows typical I-V curves collected on an Au/TiO_2/Au memristor device. Figure 5.4d and 5.4e show the resistance versus cycle and the cumulative distribution of the resistance per cycle in both HRS and LRS during 154 cycles under 0.1 V read voltage. Figure 5.4f shows the statistical analysis of set/reset voltage of 15 different devices. The device-to-device variability of the Au/TiO_2/Au memristors is quantified by measuring RVS in 15 devices and calculating the coefficient of variance (C_V), defined as the standard deviation (σ) divided by the mean value (μ) [205], at 0.1 V. Figure 5.4g shows the histogram of set/reset voltage by combining all the 15 devices. In total 300 counts are collected for both set and reset. The μ and C_V are -1.285 V and 0.156 for set voltage, and 0.811 V and 0.153 for reset voltage, respectively. Figures 5.4d-g demonstrate that the bipolar RS is stable over multiple switching cycles, and that the device-to-device variability of the switching voltages is reasonably low. Current versus time curves measured under 0.1 V at room temperature in HRS and LRS prove long state/data retention of more than a day (Figure 5.4h).

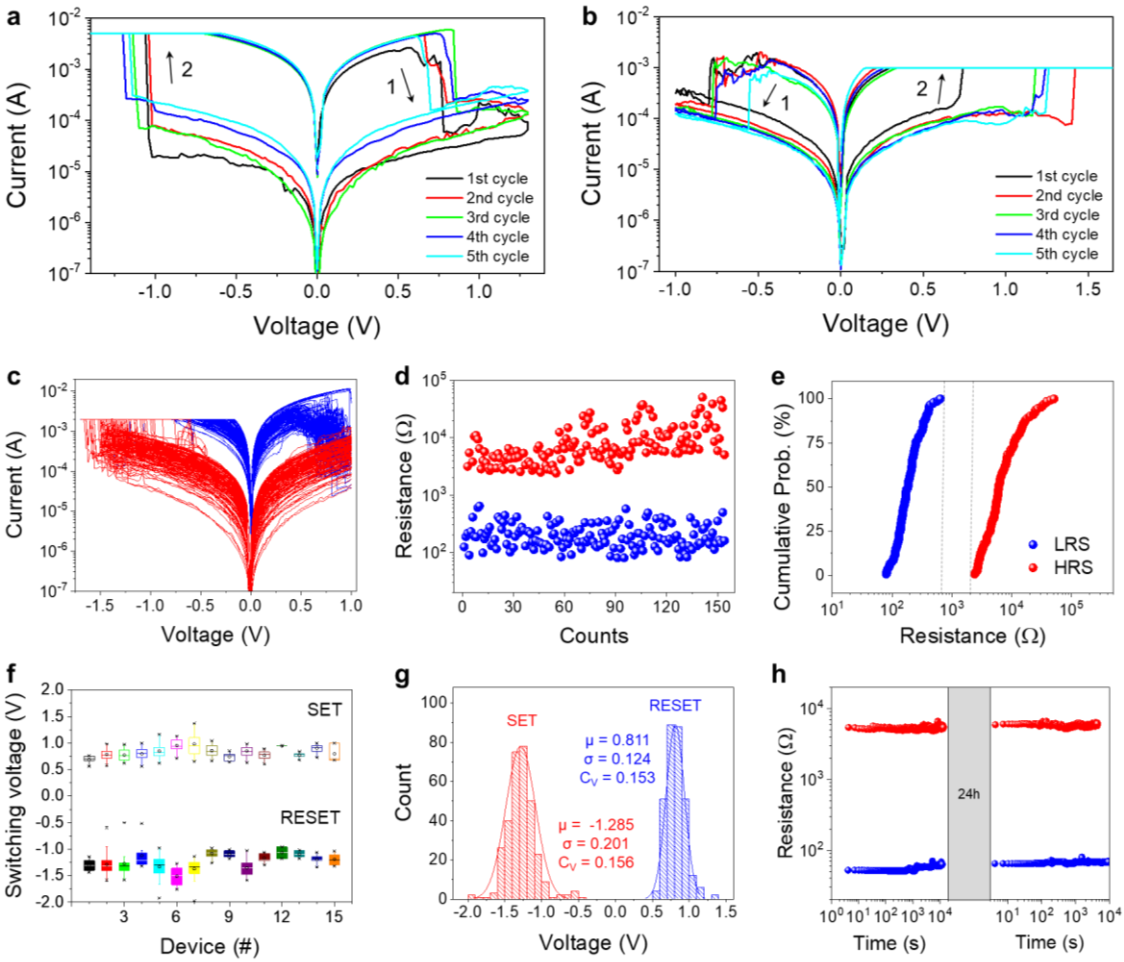


Figure 5.4 Resistive switching behaviour in Au/TiO_2/Au synapses

a and b show the typical resistive switching behaviour initializing in LRS with polarity independence. The numbered arrows indicate the order of the reset (first) and set (second) transitions. a starts with a positive bias, while b starts with negative bias. c, 154 typical I-V curves collected on an Au/TiO2/Au memristor device. d and e show the resistance versus cycle and cumulative probability versus resistance plots obtained under 0.1 V read voltage. f shows the device-to-device variability during cycling by statistical analysis of set/reset voltage of 15 devices with more than 5 endurance cycles, respectively. g, A histogram of set and reset voltage by combining all 15 devices. The normal distribution fits are provided as a guide to the eye. h shows the retention test read at 0.1 V at room temperature for 20,000 s, then stop for one day and measure again for 10,000 s, which confirms excellent state retention.

5.3.2 Forming analysis

The variability of the initial conductance of the Au/TiO_2/Au memristors is quantified by measuring RVS in 20 devices (Figure 5.5a) and plotting the value of the resistance at 0.1 V, which indicates that the initial resistance of all the devices ranges between 10^2 Ω and 10^3 Ω. The device-to-device variability of the initial conductance of LRS memristors (Figure 5.5a) is compared with that of HRS memristors (in Figure 5.5b) and forming memristors (see Figure 5.5c) also fabricated during this investigation (see Methods section). As Figures 5.5d and Supplementary Figure 2 show, the device-to-device variability of the LRS memristors is the lowest, indicating that these Au/TiO_2/Au devices present a competitive advantage (compared to HRS memristors and forming memristors) for the realization of ANNs.

Supplementary Figure 3 confirms that the relationship between current and voltage is linear, indicating that the devices are initially in LRS probably due to the formation of Au-rich regions in the TiO_2 film during the fabrication process. Considering that the CNFs have the shape of a truncated cone (one of the radii is assumed to be four times greater than the other), which is the most typically observed in MIM-like RS devices [206], we can estimate its size as follows:

$$r = \sqrt{\frac{\rho_{CNF} t_{OX}}{4\pi R}} = \frac{1}{2}\sqrt{\frac{\rho_{CNF} t_{OX}}{\pi R}} \qquad (1)$$

where R is the device resistance for a certain applied voltage (extracted from the RVS, see Figure 5.5a), ρ_{CNF} is the electrical resistivity of the CNF (the conductivity is assumed to be $\sigma_{CNF} = 1/\rho_{CNF} = 5\text{x}10^5\ \Omega^{-1}\text{m}^{-1}$, an average value of several previously employed conductivities in different models; in particular, this value was chosen from Ref. [206]), t_{OX} is the thickness of the insulator (4.5 nm according to Figure 5.3), and r is one of the truncated cone radii. By repeating the calculation for all the devices measured, it can be observed that the deviation on the cone radii is relatively low (see Figure 5.5e). By calculating the radii of the CNF at different reading voltages it is possible to discern the CNF narrowing at the onset of the reset process (Figure 5.5f), which is consistent with the observations in previous reports [207-208]. It is worth noting that the CNF radii calculated in Figures 5.5e-f are much narrower than that of the

GBs observed in AFM scans and TEM images. The reason behind this observation is that, despite the GBs are regions rich in Au ions (in these zones the CNF formation is favoured), only few locations may be included in the percolation path that constitutes the CNF that shortens the electrodes.

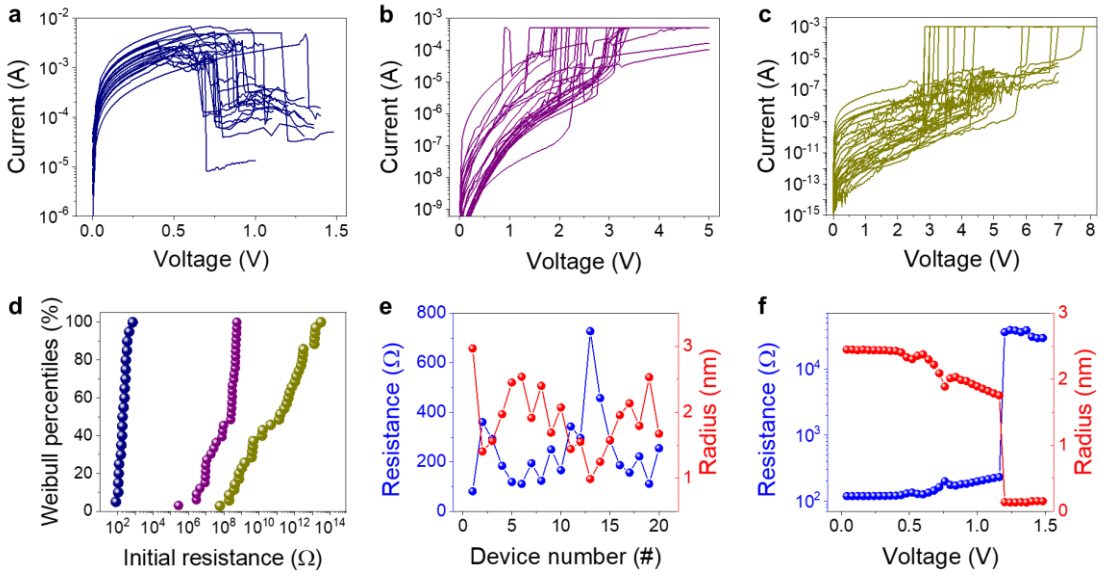


Figure 5.5 Variability of the initial conductance of LRS, HRS and forming memristors

a-c show the first I-V curves measured when applying a RVS in 20 LRS memristors (Au/TiO_2/Au), HRS memristors (Au/Ti/h-BN/Cu) and forming memristors (Au/h-BN/Au), respectively. **d**, Weibull percentiles versus initial resistance of the memristors shown in panels **a-c**, with read voltage 0.1 V. **e**, Statistical analysis of initial resistance and CNF radius of the 20 devices in **a**. **f**, Evolution of the resistance and CNF radius with the voltage from a device from panel **a**. The CNF starts to shrink at ~0.4 V, with radius decreasing from 2.3 nm to 1.7 nm. Then the filament is ruptured at ~1.2 V.

5.3.3 Pulse set and reset

The switching of the Au/TiO_2/Au devices has been also confirmed by applying sequences of pulsed voltage stresses. First, binary bipolar RS has been confirmed. Figure 5.6 shows 4 consecutive read-write-read-erase-read operations of the device under pulse mode. The device starts in HRS with 1 nA read by a pulse of 0.5 V 1 ms. Then the device is switched on by a write pulse of -4 V 1 ms. Another read pulse shows the device is switched on with a current higher than 1 μA. Then the device is switched back to HRS by an erase pulse of 5 V 1 ms. A relative high read pulse of 0.5 V is used to enable the reading of HRS over the noise level ~ 0.2 nA. Write pulses: -4 V 1 ms. Erase pulses: 5 V 1 ms. Read pulses: 0.5 V 1 ms. A higher read pulse of 0.5 V enables to read HRS over the noise level ~ 0.1 nA.

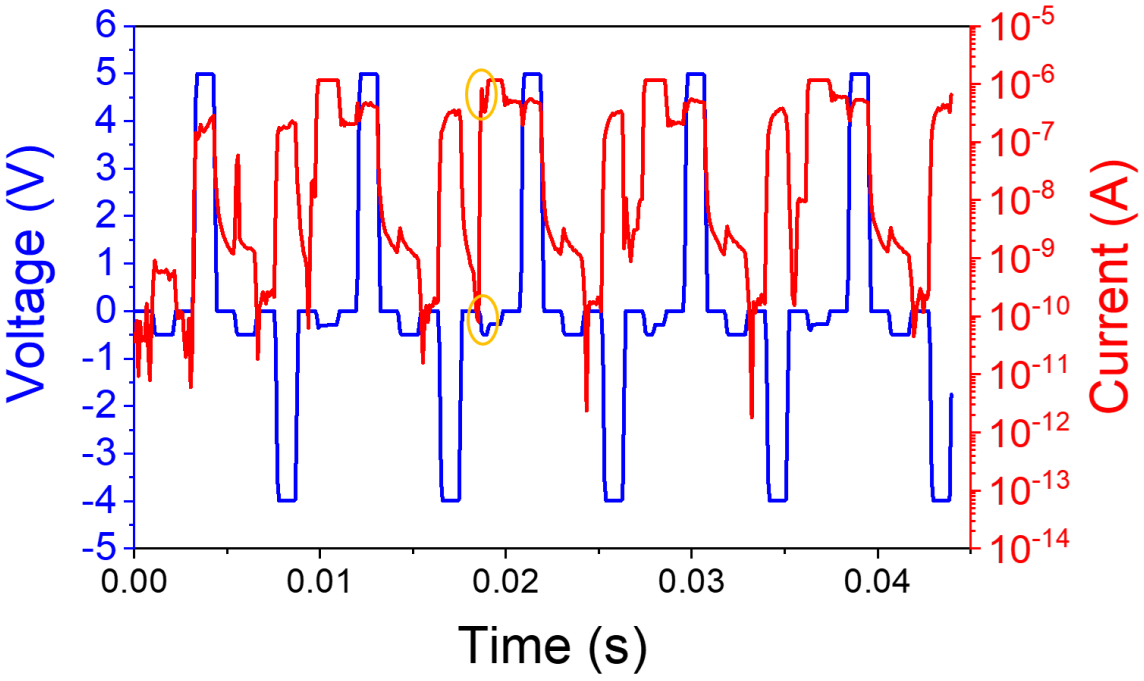


Figure 5.6 4 consecutive read-write-read-erase-read operation cycles of the device under pulse mode.

Pulses with short widths are needed to optimize the switching time and energy since the resistive change is much shorter than the pulse width, as highlighted by the golden cycles.

5.3.4 Potentiation and depression

Second, we analyze the ability of the devices to exhibit analog potentiation, which is a fundamental function required for memristive electronic synapses. To do so, the devices are exposed to potentiation pulses of 0.4 V and 1 ms, depression pulses of -0.4 V and 1 ms, and read pulses of 0.1 and V 1 ms. Figure 3h shows 10 consecutive analogue potentiation and depression cycles. Each potentiation process contains 40 pulses, while each depression process contains 40 pulses. Figure 3h only shows the current collected during the read process.

It should be highlighted that other works on HRS memristors and forming memristors may have achieved a device-to-device variability of the initial conductance lower than that of the HRS memristors and forming memristors presented here (Figures 3b-c), but we are not aware of any work reporting MIM-like memristors (of any type) with an initial variability of the conductance as low as that of the LRS memristors presented here (Figure 3a). Potentiation pulses: 0.4 V and 1 ms. Depression pulses: -0.4 V and 1 ms. Read pulses: 0.1 V and 1 ms.

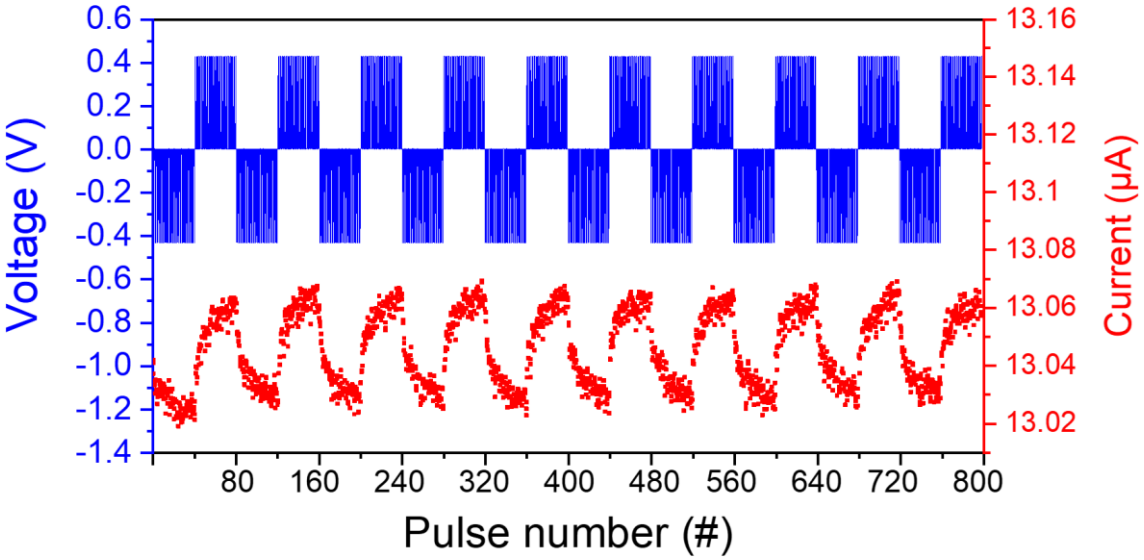


Figure 5.7 Potentiation and depression under trains of pulses

The on-off ratio can be further enlarged by applying higher switching voltage or width pulses.

5.4 Neuromorphic circuits with forming-free memristors

The most common operation in most practical neuromorphic networks is vector-by-matrix multiplication (VMM). The practical approach to implement such massive VMM blocks is to divide them into smaller VMMs (e.g., 64×64 arrays) and integrate responses in the analog/digital domain. Figure 5.8a shows the emulation of a mixed-signal VMM block using adjustable memristive devices. Due to their integration density, VMMs based on passive crossbars are promising; however, the fabrication of passively-integrated memristive neuromorphic circuits is challenging due to the required uniformity and low switching threshold distribution. The uniformity allows the individual tuning of devices and mitigating half-select disturbance. Figure 5.4 shows low variability in switching thresholds, particularly in reset operation, which grants us to implement the V/2 tuning scheme (in a crossbar integrated circuit) without employing partitioning/isolation/selector circuits that impose large area overheads. We study this issue in our recent work by extensive simulations and demonstrate excellent prospects for devices with similar variability characteristics.

Here, we focus on the second challenge, i.e., low-voltage switching characteristics and forming-free operation. The implications of low-voltage switching on inference accelerators' performance are significant, even though we only perform switching operations during the tuning phase (note that our primary focus is on ex-situ trained classifiers). The forming-free function and low switching voltage of memristors are crucial to designing CMOS-compatible, (and hence) compact, and energy-efficient neuromorphic circuits. For our reported devices, the switching voltage in both reset and set regimes are very low. The $\mu + 3\sigma$ (where μ is the average and σ is the standard deviation) of the set and reset switching threshold distributions are very close to 1.8 V and 1.2 V, respectively. Unlike most previous devices, which either require high voltage electroforming processes or feature large switching voltages, our devices are initially in the LRS state (and hence do not require high-voltage electroforming) and are switchable with CMOS-compatible input/output voltages (e.g., 1.8 V, which is the core voltage of a standard 180 nm process). Therefore, we may design all auxiliary CMOS circuitries that are inevitable to implement the desired functionality using thin-oxide MOSFETs.

Figure 5.8b shows the detailed yet simplified circuit schematic of one input and one output channel of VMM in Figure 5.8a. The analog switch matrix (ASM) block connects external pulse signals, peripheral circuits (data converter and local sensing), and

the crossbar. It is crucial to use thin-oxide MOS devices (core-voltage devices) in ASM. This stems from the fact that voltage drop on analog switches could be highly detrimental to the VMM accuracy. Thus, CMOS switches in ASM could be as large as 10× to 50× of the minimum size, depending on the block size, network, etc. In crossbars with large tuning voltages, the use of thick-oxide MOSFETs in ASM is indispensable to avoid junction breakdown and punch through. This not only makes the ASM circuit large but also would necessitate the inclusion of level-shifter circuitries (~10× minimum size thick-oxide device) to deliver the voltage from the core-voltage decoders to the ASM at a proper level. As a second side effect, the larger the ASM, the higher inter- and intra-VMM parasitics. The sizeable parasitic capacitance associated with ASM affects the VMM speed and reduces the overall throughput.

To quantitatively assess our claims, we study a representative neuromorphic architecture: a fully-analog massive multiplier perceptron with seven hidden layers and ~100M parameters (1024- 16384-16384-4096-1024-1024-512-512-256-256-10). Note that as of 2017, >60% of the Google machine learning workload was two multi-layer perceptron networks. We evaluate the performance of this network with two benchmarks: i) our forming-free devices that feature low switching voltage, and ii) passive devices that share the same specifications but require forming and high switching voltages. We design the classifier targeting high throughput and use physical layout and simulation of critical components (64×64 VMM blocks, peripheral circuits, digital blocks, ASMs, etc.) in Silterra's 180 nm CMOS process.

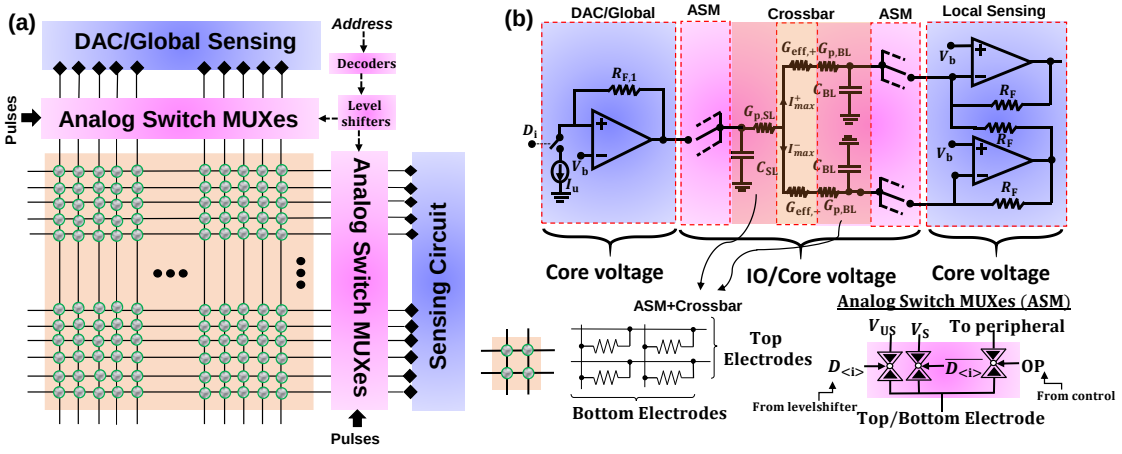


Figure 5.8 The circuit schematic of a VMM block and related components

Panel **a** shows the top-level schematic of a VMM block based on two-terminal memristive crossbars consisting of a crossbar, switch matrices, decoders, and sensing

circuits. The analog input signal is supplied either by a digital-to-analog converter (DAC) (e.g., in time-multiplexed architectures) or the global sensing circuit from the previous stage (full-analog circuits). The tuning process is implemented via analog multiplexers (MUXes). Panel **b** shows the detailed circuit diagram of one channel that performs two-quadrant multiplication with highlighted lumped parasitic from crossbar and ASM. It is imperative to design the analog peripheries (DAC and local sensing) in core voltage to ensure efficient design.

Let us emphasize several vital points before discussing the results. Though the actual size of the fabricated device is much larger, we assume $4F^2$ passive devices (where F is the half-pitch metal feature size, 250 nm in this CMOS technology) for meaningful results. Note the grain sizes in Figure 1 are ~ 100 nm, which supports device miniaturization prospects. Since the device area shrinkage reduces the leakage current as well, it is expected (and presumed in this work) that the average conductance scale linearly with device size. We simulate a critical path from one input to one output in all layers and properly model the intra-block parasitic capacitance on electrodes and global lines connecting the local sensing blocks in every layer. The total parasitic capacitance on electrodes consists of line-to-line capacitance in crossbar structure (M5/M4/M3) that includes coupling and fringing capacitors between conductors and the parasitics associated with ASM. We use the same design of critical analog components for both benchmarks; however, we optimize the settling time in each scenario, e.g., by fine-tuning the compensation circuits within amplifiers. The neurons in the last layer are loaded with a 100 fF capacitor.

Supplementary Table 3 shows relevant process parameters and simulations results. Our results indicate that the classifier with our forming-free memristors offers 1.83× better throughput per area and consumes ×0.85 less energy to classify a single image of the CIFAR-10 dataset, which stems from ~63% better density and ~17% faster operation. We explain the superior density by arguing that our forming-free devices are compatible with a core-voltage design of ASM and operate without level shifters resulting in ~35×10^3 μm^2 VMM block versus ~56×10^3 mm^2 required otherwise. The improved speed is also attributed to the fact that the total parasitic capacitance on every local sensing input lines (top electrodes or bitlines) is only 71 fF (far less than 289 fF obtained for the memristors that require forming) beside the larger parasitics on global lines as well.

5.5 Conclusions and prospect

Finally, we would like to point out that our LRS memristors could also be used to build reliable fixed-resistance physically unclonable function circuits efficiently, a building block for many cryptographic systems. These circuits are normally designed using HRS pristine memristors, and previous works experimentally demonstrate significant performance improvement in part because of the elimination of tuning circuits. The downside of this approach is higher bit-error-rate due to the reduced noise margin and large temperature dependency of HRS pristine devices. Our proposed forming-free LRS memristor technology offers a clear advantage here: forming-free LRS memristors have the most benefits of previous designs, and additionally, provides a far better noise margin and significantly better temperature dependency.

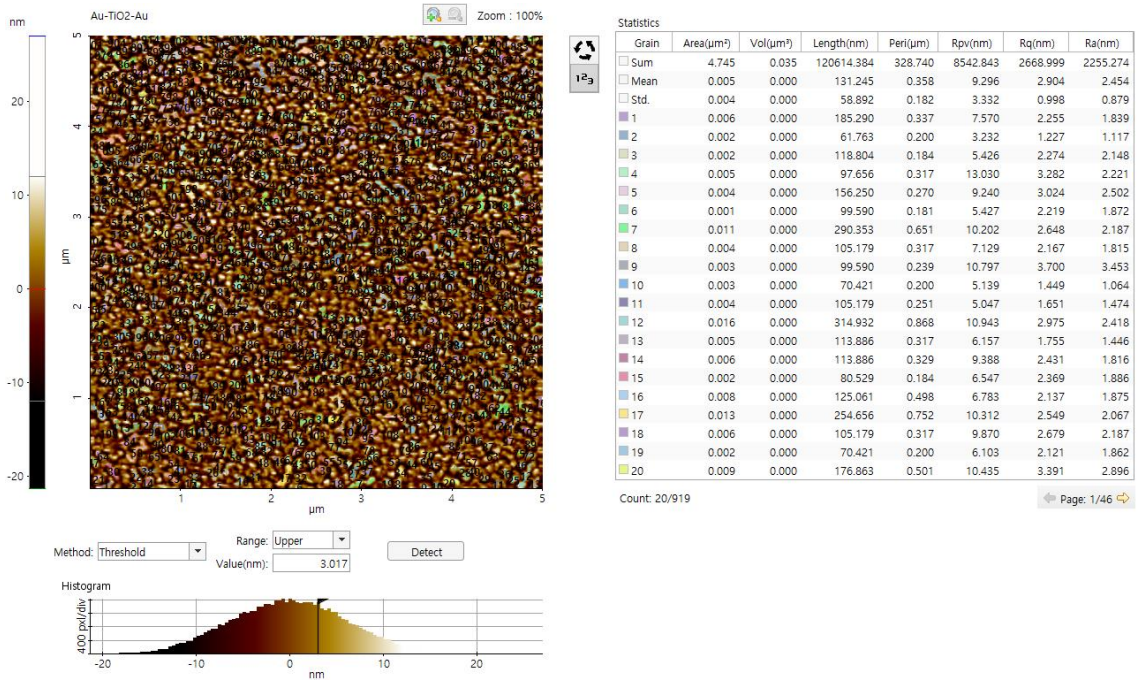
5.6 Support information

Device structure (left top / right bottom)	Pre-forming conductance	Lowest conductance	Highest conductance	Ref.
W/Al/PCMO/Pt	-	0.005 nS	0.15 nS	[1]
Pt/α-IGZO bilayer/Pt	-	0.1 nS	1 nS	[2]
Pt/Cr/Ag:Si/α-Si/W	-	1 nS	40 nS	[3]
Pt/AlO_x/N-TiN/PCMO/Pt	-	8 nS	30 nS	[4]
Ni/HfO_x/p^{++}Si	0.02 nS	13 nS	21 nS	[5]
Ta/TaO_x/TiO_2/Ti	-	80 nS	130 nS	[6]
TiN/HfO_x/Pt	10 pS	100 nS	100 μS	[7]
TiN/HfO_x/AlO_x/Pt	40 nS	1 μS	100 μS	[8]
TiN/Ti/Al:HfO_2/TiN	0.1 nS	4 μS	10 μS	[9]
TiN/TiO_x/Ta_2O_5/TiN	<1 μS	5 μS	70 μS	[10]
TiN/Ti/HfO_2/TiN	0.3 nS	20 μS	200 μS	
Ag/NiO/Pt	10 μS	32 μS	35μS	[11]
Ti/TiO_{2-x}/Pt	0.3 nS	1 mS	7 mS	[12]
Al/AlO_y/SnO_x/FTO	0.2 mS	1 mS	15 mS	[13]
Au/Ti/MoS_2/Au	-	350 nS	700 nS	[14]
Ag/MoS_2/Pt	5 nS	6 μS	8.5 μS	[15]

Supplementary Table 1 | Distribution of initial conductance, conductance range of potentiation and depression, and the ratio between lower conductance boundary and initial conductance.

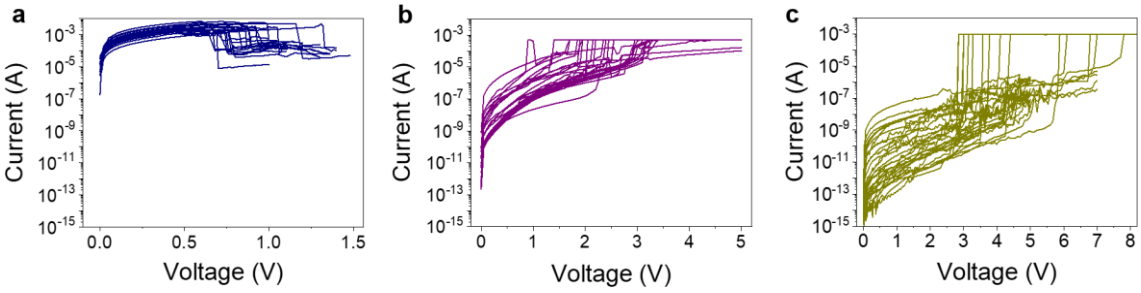
Device structure (left top / right bottom)	Forming Variability (decades)	HRS Variability (decades)	LRS variability (decades)	Ref.
TiN/Ti/HfO_2/TiN	Initial LR 20 μS - 1 mS	10 μS - 0.1 mS	0.5 mS - 1 mS	[16]
Cu/TaO_x/TiN	-	-	-	[17]
Ag/PMMA/$MA_3Sb_2Br_9$/ITO	-	0.2 mS - 0.5 mS	3 mS - 10 mS	[18]
Al/Cu_xO/Cu	-	1 μS - 500 μS	3 mS - 50 mS	[19]
TiN/TiO_x/HfO_x/TiN	-	50 nS - 1 μS	0.3 mS - 1 mS	[20]
Pt/NiO_y/Pt	-	-	-	[21]
TiN/HfO_x/TiO_x/HfO_x/TiO_x/Pt	-	2 μS - 3 μS	2 mS - 10 mS	[22]
Al/N-AlO_x/Al	-	0.3 nS - 1 nS	100 nS - 300 nS	[23]
Al/Ni doped HfO_x/Pt	-	20 μS - 200 μS	2 mS - 5 mS	[24]
Al/$BiFeO_3$/FTO	-	200 μS - 300 μS	2 mS	[25]
Al/SnO_x/FTO	-	0.2 mS - 2 mS	20 mS	[26]
Ta/HfO_2/RuO_2	50 μS - 0.1 mS	80 μS - 0.2 mS	0.5 mS	[27]
Pt/TaO_x/TiO_x/Pt	-	-	-	[28]
Al/nc-Si:H/a-SiNx:H/P^+-Si	-	20 nS - 20 μS	10 μS - 0.2 mS	[29]
Pt/Ti/GeSe/TiN	-	2 μS - 10 μS	0.5 mS - 1 mS	[30]

Supplementary Table 2 | Distribution of device-to-device variability of forming-free memristors with initial conductance in HRS

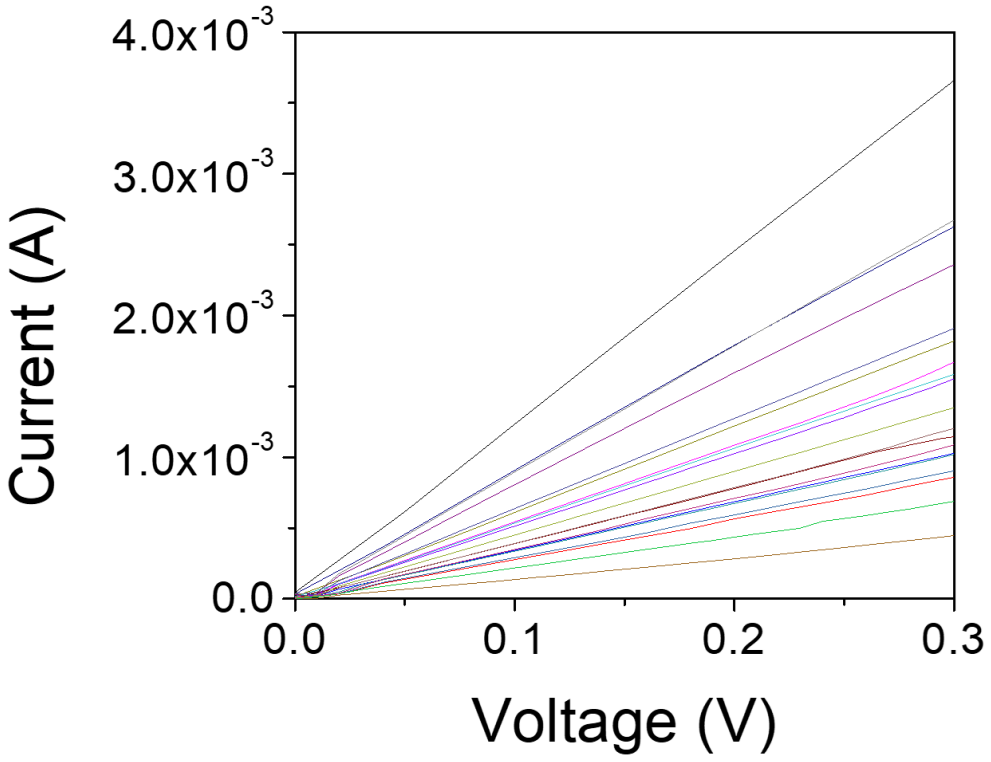


Grain	Area(µm²)	Vol(µm³)	Length(nm)	Peri(µm)	Rpv(nm)	Rq(nm)	Ra(nm)
Sum	4.745	0.035	120614.384	328.740	8542.843	2668.999	2255.274
Mean	0.005	0.000	131.245	0.358	9.296	2.904	2.454
Std.	0.004	0.000	58.892	0.182	3.332	0.998	0.879
1	0.006	0.000	185.290	0.337	7.570	2.255	1.839
2	0.002	0.000	61.763	0.200	3.232	1.227	1.117
3	0.002	0.000	118.804	0.184	5.426	2.274	2.148
4	0.005	0.000	97.656	0.317	13.030	3.282	2.221
5	0.004	0.000	156.250	0.270	9.240	3.024	2.502
6	0.001	0.000	99.590	0.181	5.427	2.219	1.872
7	0.011	0.000	290.353	0.651	10.202	2.648	2.187
8	0.004	0.000	105.179	0.317	7.129	2.167	1.815
9	0.003	0.000	99.590	0.239	10.797	3.700	3.453
10	0.003	0.000	70.421	0.200	5.139	1.449	1.064
11	0.004	0.000	105.179	0.251	5.047	1.651	1.474
12	0.016	0.000	314.932	0.868	10.943	2.975	2.418
13	0.005	0.000	113.886	0.317	6.157	1.755	1.446
14	0.006	0.000	113.886	0.329	9.388	2.431	1.816
15	0.002	0.000	80.529	0.184	6.547	2.369	1.886
16	0.008	0.000	125.061	0.498	6.783	2.137	1.875
17	0.013	0.000	254.656	0.752	10.312	2.549	2.067
18	0.006	0.000	105.179	0.317	9.870	2.679	2.187
19	0.002	0.000	70.421	0.200	6.103	2.121	1.862
20	0.009	0.000	176.863	0.501	10.435	3.391	2.896

Supplementary Figure 1 | Grain size analysis with XEI software on the sputtered TiO_2 on 300 nm SiO_2/Si substrate. Here we choose 3 nm as threshold value to detect the grain.



Supplementary Figure 2 | Variability of the conductance of LRS, HRS and forming memristors. This figure corresponds to the data in Figure 3a-c, but using the same minimum and maximum values in the current scale. It can be easily seen that the variability of the initial currents in LRS is the lowest.



Supplementary Figure 3 | IV curves of LRS memristors linear scale. The linear relation between current and voltage indicates the conductive filaments are initially formed. These data are the same than those in Figure 3a of the main text.

	description	value	
Process features	CMOS Process feature size (nm)	180	
	Parasitic cap of a minimum metal width (fF/μm) (crossbar structure, M5/M4/M3)	0.226	
	Parasitic cap of a minimum size of a thick oxide device (fF)	1.2	
	Parasitic cap of a minimum size of a thin oxide device (fF)	0.58	
	Global and Local sensing voltage swing (V)	0.1	
	Maximum drop on tuning switches (V)	0.5	
	IO Supply voltage (V)	3.3	
	Core Supply voltage (V)	1.8	
	The on-resistance of a minimum-size thick-oxide device (kΩ)	4.4	
	The on-resistance of a minimum-size thin-oxide device (kΩ)	2.7	
	Minimum area of a thick oxide device (μm^2)	1.46	
	Minimum area of a thin oxide device (μm^2)	1.12	
	MIM Capacitor (fF/μm^2)	4	
	Memory block Size	64×64	
		0T1R (this work, low Vth)	0T1R High Vth
Memory features	Cell area (μm^2)	0.25	0.25
	Midrange conductance (μS)	10	10
	Cell parasitic capacitance (fF)	0.113	0.113
	VMM parasitic capacitance (fF)	71	289
	Nominal switching current (μA)	200	200
	Maximum local Sensing input-current (μA)	10	10
VMM Design features	Array (μm^2)	1,024	1,024
	Level shifters (μm^2)	*NA*	4,608
	Local sensing circuits (μm^2)	17,280	17,280
	Digital circuits (μm^2)	3,340	3,340
	Tuning Switches (μm^2)	14,192	30,087
	Analog VMM block (μm^2)	35,836	56,339
	DACs (μm^2)	600	600
	Global Sensing circuits (μm^2)	270	270
	Analog VMM Power (mw)	1.152	1.152
7-layer Fully analog MLP architecture (~105M weights)	Throughput (Mfps)	14.61	12.46
	Total Area (cm^2)	18.47	29
	Throughput per area (Mfps/cm^2)	0.79	0.43
	Energy Consumed per frame (μJ/f)	13.63	15.98

Supplementary Table 3 | Simulation parameters and results for the fully-analog neuromorphic classifier.

Supplementary references

[S1] S. Park et al., *Tech. Dig. - Int. Electron Devices Meet. IEDM* **2012**, *231*.

[S2] Z. Wang et al., *Adv. Funct. Mater.* **2012**, *22*, 2759.

[S3] S. Jo, et al., *Nano Lett.* **2010**, *10*, 1297.

[S4] S. Park et al., *Tech. Dig. - Int. Electron Devices Meet. IEDM* **2013**, 625.

[S5] J. Chen et al, **2018**, *404*, 1.

[S6] I. Wang et al., *Nanotechnology* **2016**, *27*, 1.

[S7] B. Gao et al, *ACS Nano* **2014**, *8*, 6998.

[S8] S. Yu et al., *IEEE Trans. Electron Devices* **2011**, *58*, 2729.

[S9] S. Chandrasekaran et al., *Nanotechnology* **2019**, *30*.

[S10] J. Woo et al., *IEEE Electron Device Lett.* **2017**, *38*, 1220.

[S11] Y. Li et al., *ACS Appl. Mater. Interfaces* **2018**, *10*, 24598.

[S12] S. Chen et al, *Chem. Mater.* **2019**, *31*, 8394.

[S13] J. Woo et al., *IEEE Electron Device Lett.* **2017**, *38*, 1220.

[S14] D. Li et al., *ACS Nano* **2018**, *12*, 9240.

[S15] S. Yin et al., *Phys. Status Solidi Appl. Mater. Sci.* **2019**, *216*, 1.

[S16] Y. Chen et al., *International Symposium on VLSI Technology, Systems, and Applications* **2009**, *29*, 37.

[S17] Y. Li et al., *Jpn. J. Appl. Phys.* **2020**, *59*, 044004.

[S18] J. Yang et al., *Nanoscale* **2019**, *11*, 6453.

[S19] H. Lv et al., *IEEE Electron Device Lett.* **2008**, *29*, 47.

[S20] H. Lee, et al., *Tech. Dig. - Int. Electron Devices Meet. IEDM* **2008**, 3.

[S21] K. Kinoshita et al., International Conference on Solid State Devices and Materials **2006**, 570-571.

[S22] Z. Fang et al., *IEEE Electron Device Lett.* **2011**, *32*, 566.

[S23] W. Kim et al., *Dig. Tech. Pap. - Symp. VLSI Technol.* **2011**, *30*, 22.

[S24] M. Akbari, J. S. Lee, *RSC Adv.* **2016**, *6*, 21917.

[S25] C. Kumari et al., *AIP Adv.* **2020**, *10*.

[S26] S. Ng et al., *ACS Appl. Electron. Mater.* **2020**, *2*, 817.

[S27] G. Kim et al., *ACS Appl. Mater. Interfaces* **2019**, *11*, 47063.

[S28] S. Srivastava et al., *Nanoscale* **2019**, *11*, 18159.

[S29] Y. Sun et al., *Nanotechnology* **2019**, *30*.

[S30] W. Kim et al., *ACS Appl. Mater. Interfaces* **2019**, *11*, 38910.

Chapter 6: Conclusions and perspectives

In conclusion, in this PhD thesis, I had a deep analysis of the h-BN based memristor, from a single device to integrated circuits.

The main objectives of this thesis are to design, fabrication, characterization and optimization of h-BN based memristors, with the main goal of achieving a circuit-level integration with the mature Si-based CMOS technology. Comparing the main objectives, I conclude my PhD thesis as below:

1- Chemical vapour deposited (CVD) and liquid phase exfoliated (LPE) (See Appendix D: Inkjet-printed h-BN based memristor) produced h-BN can be used as a reliable resistive switching medium for 2D materials-based memristors. Both of them are scalable methods for large area integration.

2- The correct layered structure of the multilayer h-BN has been confirmed by the cross-sectional TEM, and the thickness is around 6 nm (18 layers). The 2D nature of h-BN and graphene has been confirmed by the Raman spectrum statistically.

3- The local defects are electrically weak spots which can work as the confined switching region for memristors, confirmed by the conductive atomic force microscope (CAFM).

4- With crosspoint structure and photolithography, h-BN memristors with a size of 25 μm^2 have been fabricated. With via hole structure, h-BN memristors with a size of 0.053 μm^2 have been achieved.

5- h-BN memristors with graphene insertion show stable and highly controllable tristate operation, by controlling the current compliance (CC) and reset voltage (V_{RESET}).

6- One-transistor-one-memristor (1T1M) cell structure can greatly improve the performance of the h-BN memristor and provide more precise controllability of resistance level, with inert bottom electrode tungsten (W).

7- 1T1M cell with top electrode Ti can achieve low current operation and multilevel switching. 1T1M cell with top electrode Au can achieve high endurance of 2.5 million cycles. 1T1M cell with top electrode Ag can achieve fast operation and low energy consumption with reset energy 1.41 pJ.

8- 1T1M cell with top electrode Ti can show both volatile and non-volatile switching behaviour based on the gate voltage (V_G) used. High V_G means high current compliance (by the transistor, instead of the instrument semiconductor parameter analyser) and tends to show non-volatile switching and vice versa.

9- 1T1M cell with top electrode Ti can show synaptic behaviour of spike-time-dependent-plasticity (STDP).

10- By using the 1T1M cell structure, the sneak path current can be efficiently suppressed.

11- With two 1T1M cells, we can accomplish logic operations including “OR” and “IMP”.

12- Based on the STDP characteristic, a spiking neural network (SNN) was modelled with the MNIST dataset and the best average accuracy reaches ~90%.

We have demonstrated that fabricating hybrid CMOS/2D-LMs microchips is possible, and that it can provide superior electrical performance. In particular, we have integrated multilayer h-BN on the back-end-of-line interconnections of a silicon chips containing CMOS transistors arranged in a crossbar array configuration. The resulting crossbar array of 1T1M cells shows a performance much superior to that of standalone h-BN memristors.

Future work should focus on the fully hardware implementation of the SNN and achieve in-situ classification; vector-matrix multiplication for artificial neural network (ANN), and fully 2D materials-based electronics. Although the high operation speed can be achieved in a single h-BN memristor (switching time less than 10 ns), the switching time of 1T1M cell is much long of hundreds of ns or even few µs. The integration should be optimized. High-temperature retention should be improved with better connection or integration between the 2D materials and CMOS via.

References

[1] M. Hilbert, P. Lopez, Science 2011, 332, 60.

[2] D. Reinsel et al., (13 April 2017). "Data Age 2025: The Evolution of Data to Life-Critical" (PDF). seagate.com.

[3] Website of Zoom. https://zoom.us/

Transistors

[4] The International Roadmap for Devices and Systems 2020 Edition (IRDS, IEEE, accessed 24 December 2020); https://irds.ieee.org/editions/2020

[5] J. Kilby, Miniaturized electronic circuits. US patent 3,138,743 (1964).

[6] R. Noyce, Semiconductor device-and-lead structure. US patent 2,981,877 (1961).

[7] D. Kahng, Electric field controlled semiconductor device. US patent 3,102,230 (1963).

[8] G. Moore. Cramming more components onto integrated circuits. Electronics 38, 8, 1965. **_Moore's Law**

[9] Website of wiki. Wikipedia.org/wiki/transistor_count

[10] D. Patel, Apple's A14 packs 134 million transistors/mm2, but falls short of TSMC's density claims. SemiAnalysis (27 October 2020).

[11] R. Dennard. Design of ion-implanted MOSFET's with very small physical dimensions. Proc. IEEE. 87, 4, 1999. **_Dennard's Law**

Memristors

[12] M. Lanza et al., Recommended Methods to Study Resistive Switching Devices. Adv. Electron. Mater. 2019, 5, 1800143. **_Recommended methods to study memristors**

[13] S. Pi et al., Memristor crossbar arrays with 6-nm half-pitch and 2-nm critical dimension. Nature Nanotechnology 14, 35, 2019.

[15] B. Choi et al., High-speed and low-energy nitride memristors. Adv. Funct. Mater. 2016, 26, 5290–5296.

[16] M. Lee. et al., A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta2O5−x/TaO2−x bilayer structures. Nature Materials. 10, 625 (2011). **_Endurance**

[17] Retention.

[18] L. Chua, IEEE Trans. Circuit Theory CT18, 507–519 (1971).

[19] D. Strukov et al., The missing memristor found. Nature 453, 80 (2008).

[20] A. Beck et al., Reproducible switching effect in thin oxide films for memory applications. Applied Physics Letters 77, 1 (2000). **_IBM**

[21] I. Baek et al., Highly scalable non-volatile resistive memory using simple binary oxide driven by asymmetric unipolar voltage pulses. IEDM 04-587 (2004). **_Samsung**

[22] I. Baek et al., Multi-layer cross-point binary oxide resistive memory (OxRRAM) for post-NAND storage application, IEEE IEDM 750-753 (2005). _**Samsung**

[23] Y. Kim et al., Bi-layered RRAM with unlimited endurance and extremely uniform switching. Symposium on VLSI Technology 52-53 (2011). _**Samsung**

[24] B. Govoreanu et al., 10x10nm2 Hf/HfOx crossbar resistive RAM with excellent performance, reliability and low-energy operation. IEDM 11-729 (2011). _**Imec**

[25] A. Fantini et al., Lateral and vertical scaling impact on statistical performances and reliability of 10nm TiN/Hf(Al)O/Hf/TiN RRAM devices. Symposium on VLSI Technology (2014). _**Imec**

[26] C. Chen et al., Stack optimization of oxide-based RRAM for fast write speed (<1 ls) at low operating current (<10 lA). Solid-State Electronics 125, 198 (2016). _**Imec**

[27] O. Golonzka et al., Non-volatile RRAM embedded into 22FFL FinFET technology. Symposium on VLSI Technology. T230 (2019). _**Intel**

[28] A. Bessonov et al., Layered memristive and memcapacitive switches for printable electronics. Nature Materials 14, 199 (2015). _**Nokia**

[29] Y. Tseng et al., High density and ultra small cell size of contact ReRAM (CR-RAM) in 90nm CMOS logic technology and circuits. IEDM 09-109 (2009). _**TSMC**

[30] C. Chou et al., A 22nm 96KX144 RRAM macro with a self-tracking reference and a low ripple charge pump to achieve a configurable read window and a wide operating voltage range. IEEE Symposium on VLSI Circuits 1, (2020). _**TSMC**

[31] K. Tsunoda et al., Low power and high speed switching of Ti-doped NiO ReRAM under the unipolar voltage source of less than 3 V. IEEE IEDM 767 (2007). _**Fujitsu**

[32] Fujutsu Semiconductor Data Sheet. Memory ReRAM 8M (2014K×8) Bit SPI MB85AS8MT. _**Fujitsu**

[33] T. Liu et al., A 130.7mm2 2-Layer 32Gb ReRAM memory device in 24nm technology. IEEE International Solid-State Circuits Conference (IEEE ISSCC) (2013). _**SanDisk**

[34] A. Chen et al., Non-volatile resistive switching for advanced memory applications. IEEE IEDM 746 (2005). _**Spansion**

[35] A. Chen. Accessibility of nano-crossbar arrays of resistive switching devices. IEEE International Conference on Nanotechnology. 1767 (2011). _**GlobalFoundries**

[36] J. Yang et al., A 28nm 1.5Mb embedded 1T2R RRAM with 14.8 Mb/mm2 using sneaking current suppression and compensation techniques, 2020 IEEE Symposium on VLSI Circuits, 1, (2020). _**IMECAS**

[37] J. Yang et al., 24.2 A 14nm-FinFET 1Mb Embedded 1T1R RRAM with a 0.022μm2 Cell Size Using Self-Adaptive Delayed Termination and Multi-Cell Reference, 2021 IEEE International Solid- State Circuits Conference (ISSCC), 336 (2021). _**IMECAS**

[38] Probe station tip

[39] Y. Zuo et al., Effect of the pressure exerted by probe station tips in the electrical characteristics of memristors. Adv. Electron. Mater. 1901226 (2020). **_Common bottom electrode**

[40] P. Yao et al., Face classification using electronic synapses. Nature Communications, 8, 15199 (2017).

[41] The website of Muse Semiconductor. https://www.musesemi.com/shared-block-tapeout-pricing

[42] Z. Wang et al., Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. Nature Materials. 16, 101 (2017). **_Planar structure for in-situ TEM _Ag**

[43] D. Ielmini, R. Waser, Resistive switching: from fundamentals of nanoionic redox processes to memristive device applications. Wiley-VCH Book, Wiley-VCH, Weinheim, Germany 2016.

[44] M. Prezioso et al., Training and operation of an integrated neuromorphic network based on metal-oxide memristors. Nature 512, 61 (2015).

[45] Y. Li et al., Improvement of resistive switching characteristics in ZrO2 film by embedding a thin TiOx layer. Nanotechnology 22, 254028 (2011).

[46] S. Rahaman et al., Scalability and reliability issues of Ti/HfOx-based 1T1R bipolar RRAM: Occurrence, mitigation, and solution. Applied Physics Letters 110, 213501 (2017).

[47] Y. Liu et al., Inkjet Printing: A cheap and easy-to-use alternative to wire bonding for academics. Crystal Research & Technology 57, 2100210 (2022).

[48] H. Kim et al., Engineering chemically abrupt high-k metal oxide/silicon interfaces using an oxygen-gettering metal overlayer. Journal of Applied Physics 96, 6 (2004).

[49] S. Wang et al., Effects of Ti top electrode thickness on the resistive switching behaviors of rf-sputtered ZrO2 memory films. Applied Physics Letters 95, 112904 (2009).

[50] W. Kim et al., Electroforming-free bipolar resistive switching in GeSe thin films with a Ti-containing electrode. ACS Appl. Mater. Interfaces 11, 38910 (2019).

[51] T. Wang et al., Improving HfO2-based resistive switching devices by inserting a TaOx thin film via engineered in situ oxidation. ACS Appl. Mater. Interfaces 14, 21, 24565 (2022).

[52] Z. Wang et al., Threshold switching of Ag or Cu in dielectrics: materials, mechanism, and applications. Advanced Functional Materials 28, 1704862 (2018).

[55] T. Becker et al., An electrical model for trap coupling effects on random telegraph noise. IEEE Electron Device Letters 41, 10 (2020).

[56] X. Li et al., Random telegraph noise in metal-oxide memristors for true random number generators: a materials study. Advanced Functional Materials 2102172 (2021).

[58] The website of imec. https://www.imec-int.com/en/200mm-GaN-on-Si-technology

[59] I. Valov et al., Electrochemical metallization memories—fundamentals, applications, prospects. Nanotechnology 22, 254003 (2011).

[64] D. Kwon et al., Atomic structure of conducting nanofilaments in TiO2 resistive switching memory. Nature Nanotechlogy 2010, 5, 148.

[65] X. Yan et al., Superior resistive switching memory and biological synapse properties based on a simple TiN/SiO2/p-Si tunneling junction structure. Journal of Materials Chemistry C 5, 2259 (2017).

[66] A. Mehonic et al., Silicon oxide (SiOx): a promising material for resistance switching? Advanced Materials 1801187 (2018).

[67] Y. Wu et al., Al2O3-based RRAM using atomic layer deposition (ALD) with 1-μA RESET current. IEEE Electron Device Letters 31, 1449 (2010).

[69] B. Yalagala et al., V2O5 Nanosheets for Flexible Memristors and Broadband Photodetectors. ACS Applied Nano Materials 2, 937−947 (2019).

[70] S. Kumar et al., Chaotic dynamics in nanoscale NbO2 Mott memristors for analogue computing. Nature 548, 318 (2017).

[72] S. Goswami et al., Robust resistive memory devices using solution-processable metal-coordinated azo aromatics. Nature Materials 16, 1216 (2017).

[73] T. Wang et al., Electroforming in metal-oxide memristive synapses. ACS Applied Materials Interfaces 12, 11806 (2020).

[76] M. Prezioso et al., Training and operation of an integrated neuromorphic network based on metal-oxide memristors. Nature 521, 61 (2015).

2D materials

[81] K. Novoselov et al., Electric field effect in atomically thin carbon films. Science 306, 666 (2004).

[82] The Nobel Prize in Physics 2010: https://www.nobelprize.org/prizes/physics/2010/summary/

[83] N. Li et al., Large-scale flexible and transparent electronics based on monolayer molybdenum disulfide field-effect transistors. Nature Electronics 3, 711 (2020).

[84] Z. Lin et al., Solution-processable 2D semiconductors for high performance large-area electronics. Nature 562, 254 (2018).

[85] M. Wang et al., Robust memristors based on layered two dimensional materials. Nature Electronics 1, 130 (2018).

[86] A. Bessonov et al., Layered memristive and memcapacitive switches for printable electronics. Nature Materials 14, 199 (2014).

[87] Y. Yang et al., The role of geometric sites in 2D materials for energy storage. Joule 2, 1075 (2018).

[88] X. Zhang et al., Two-dimensional MoS2-enabled flexible rectenna for Wi-Fi-band wireless energy harvesting. Nature 566, 368 (2019).

[89] J. Yao et al., Highly transparent nonvolatile resistive memory devices from silicon oxide and graphene. Nature Communications 3, 1101 (2012).

[90] D. Akinwande et al., Two-dimensional flexible nanoelectronics. 5, 5678 (2014).

[91] S. Bae et al., Roll-to-roll production of 30-inch graphene films for transparent electrodes. Nature Nanotechnology. 5, 574 (2010).

[92] N. Wang et al., Replacing copper interconnects with graphene at a 7-nm node. IEEE International Interconnect Technology Conference (IITC) 1 (2017).

[93] L. Li et al., BEOL compatible graphene/Cu with improved electromigration lifetime for future interconnects. IEDM 9.5.1 (2016).

[94] Z. Wang et al., Oxide thin-film electronics using all-MXene electrical contacts. Advanced Materials 30, 1706656 (2018).

[95] B. Lyu et al., Large-area MXene electrode array for flexible electronics. ACS Nano 13, 11392 (2019).

[96] B. Radisavljevic et al., Single-layer MoS2 transistors. Nature Nanotechnology 6, 147 (2011).

[97] K. Kang et al., High-mobility three-atom-thick semiconducting films with wafer-scale homogeneity. Nature 520, 656 (2015).

[98] Z. Lin et al., Solution-processable 2D semiconductors for high performance large-area electronics. Nature 562, 254 (2018).

[99] K. Smithe et al., Low Variability in Synthetic Monolayer MoS2 Devices. ACS Nano 11, 8456 (2017).

[100] Q. Smets et al., Ultra-scaled MOCVD MoS2 MOSFETs with 42nm contact pitch and 250μA/μm drain current. IEDM 23.2.1 (2019).

[101] D. Lin et al., Dual gate synthetic WS2 MOSFETs with 120μS/μm Gm 2.7μF/cm2 capacitance and ambipolar channel. IEDM 3.6.1 (2020).

[102] Z. Ahmed et al., Introducing 2D-FETs in Device Scaling Roadmap using DTCO. IEDM 22.5.1 (2020).

[103] Q. Smets et al., Sources of variability in scaled MoS2 FETs. IEDM 3.1.1 (2020).

[104] I. Asselberghs et al., Wafer-scale integration of double gated WS2transistors in 300mm Si CMOS fab. IEDM 40.2.1 (2020).

[105] X. Wu et al., ALD Encapsulation of CVD WS2 for Stable and High-Performance FET Devices. IEEE Electron Devices Technology & Manufacturing Conference (EDTM) 1 (2021).

[106] X. Wu et al., Dual gate synthetic MoS2 MOSFETs with 4.56μF/cm2 channel capacitance, 320μS/ μm Gm and 420 μA/ μm Id at 1V Vd/100nm Lg.

[107] Y. Shi et al., Superior electrostatic control in uniform monolayer MoS2 scaled transistors via in-situ surface smoothening.

[108] Q. Smets et al., Scaling of double-gated WS2 FETs to sub-5nm physical gate length fabricated in a 300mm FAB. IEDM 34.2.1 (2021).

[109] K. Brien et al., Advancing 2D Monolayer CMOS Through Contact, Channel and Interface Engineering. Symposium on VLSI Technology 1 (2021).

[110] C. Dorow et al., Advancing monolayer 2D NMOS and PMOS transistor integration from growth to van der Waals interface engineering for ultimate CMOS scaling. Symposium on VLSI Technology 1 (2021).

[111] Y. Lin et al., Contact Engineering for High-Performance N-Type 2D Semiconductor Transistors. IEDM 37.2.1 (2021).

[112] A. Chou et al., Antimony semimetal contact with enhanced thermal stability for high performance 2D Electronics. IEDM 7.2.1 (2021).

[113] K. Shin et al., Study of selective graphene growth on non-catalytic hetero-substrates. 2D Materials 7, 011002 (2019).

[114] J. Lee et al., Wafer-Scale Growth of Single-Crystal Monolayer Graphene on Reusable Hydrogen-Terminated Germanium. Science 344, 286 (2014).

[115] S. Hong et al., Ultralow-dielectric-constant amorphous boron nitride. Nature 582, 511 (2020).

[116] D. Pacile et al., The two-dimensional phase of boron nitride: Few-atomic-layer sheets and suspended membranes. Applied Physics Letters 92, 133107 (2008).

[117] N. Alem et al., Atomically thin hexagonal boron nitride probed by ultrahigh-resolution transmission electron microscopy. Physical Review B 80, 155425 (2009).

[118] J. Meyer et al., Selective sputtering and atomic resolution imaging of atomically thin boron nitride membranes. Nano Letters 9, 2683 (2009).

[119] C. Dean et al, Boron nitride substrates for high-quality graphene electronics. Nature Nanotechnology 5, 722 (2010).

[120] J. Xue et al., Scanning tunnelling microscopy and spectroscopy of ultra-flat graphene on hexagonal boron nitride. Nature Materials 10, 282 (2011).

[121] G. Lee et al., Electron tunneling through atomically flat and ultrathin hexagonal boron nitride. Applied Physics Letters 99, 243114 (2011).

[122] L. Britnell., Electron tunneling through ultrathin boron nitride crystalline barriers. Nano Letters 12, 1707 (2012).

[123] K. Kim et al., Synthesis of monolayer hexagonal boron nitride on Cu foil using chemical vapor deposition. Nano Letters 12, 161 (2012).

[124] K. Kim et al., Synthesis and characterization of hexagonal boron nitride film as a dielectric layer for graphene devices. ACS Nano 6, 8583 (2012).

[125] The Website of company Graphene Supermarket. https://www.graphene-supermarket.com/products/single-layer-h-bn-boron-nitride-film-grown-on-copper-foil-2-x-1

[126] The Website of company Graphene Supermarket. https://www.graphene-supermarket.com/products/multilayer-h-bn-boron-nitride-film-grown-on-copper-foil-2-x-1

[127] J. Lee et al., Wafer-scale single-crystal hexagonal boron nitride film via self-collimated grain formation. Science 362, 817 (2018).

[128] L. Wang et al., Epitaxial growth of a 100-square-centimetre singlecrystal hexagonal boron nitride monolayer on copper. Nature 570, 91 (2019).

[129] T. Chen et al., Wafer-scale single-crystal hexagonal boron nitride monolayers on Cu (111). Nature 579, 219 (2020).

[130] K. Ma et al., Epitaxial single-crystal hexagonal boron nitride multilayers on Ni (111). Nature 606, 88 (2022).

2D materials based memristor

[141] J. Yao et al., Highly transparent nonvolatile resistive memory devices from silicon oxide and graphene. Nature Communications 3, 1101 (2012).

[142] H. Tian et al., Monitoring oxygen movement by Raman Spectroscopy of Resistive Random Access Memory with a Graphene-inserted electrode. Nano Letters 13, 651 (2013).

[143] X. Zhao et al., Breaking the current-retention dilemma in cation-based Resistive Switching Devices utilizing graphene with controlled defects. Advanced Materials 30, 1705193 (2018).

[144] J. Lee et al., Tuning ionic transport in memristive devices by Graphene with engineered nanopores. ACS Nano 10, 3571 (2016).

[145] B. Liu et al., Programmable synaptic metaplasticity and below Femtojoule spiking energy realized in Graphene-based neuromorphic memristor. ACS Applied Materials & Interfaces 10, 20237 (2018).

[146] T. Aziz et al., Transparent hybrid ZnO-graphene film for high stability switching behavior of memristor device. Materials Science in Semiconductor Processing 89, 68 (2019).

[147] H. Tian et al., High performance 2D perovskite/graphene optical synapses as artificial eyes. IEDM 18-891 (2018).

[148] Q. Chen et al., Low power parylene-based memristors with a graphene barrier layer for flexible electronics applications. Advanced Electronic Materials 5, 1800852 (2019).

[149] M. Wang et al., Robust memristors based on layered two dimensional materials. Nature Electronics 1, 130 (2018).

[150] X. Wang et al., Interface engineering with MoS2–Pd nanoparticles hybrid structure for a low voltage resistive switching memory. Small 14, 1702525 (2018).

[151] R. Ge et al., Atomristor: nonvolatile resistance switching in atomic sheets of transition metal dichalcogenides. Nano Letters 18, 434 (2018).

[152] M. Kim et al., Zero-static power radio-frequency switches based on MoS2 atomristors. Nature Communications 9, 2524 (2018).

[153] F. Wu et al., Interface engineering via MoS2 insertion layer for improving resistive switching of conductive-bridging random access memory. Advanced Electronic Materials 5, 1800747 (2019).

[154] R. Xu et al., Vertical MoS2 double-Layer memristor with electrochemical metallization as an atomic-scale synapse with switching thresholds approaching 100 mV. Nano Letters 19, 2411 (2019).

[155] A. Krishnaprasad et al., Electronic synapses with near-linear weight update using MoS2/graphene memristors. Applied Physics Letters 115, 103104 (2019).

[156] H. Kalita et al., Artificial Neuron using Vertical MoS2/Graphene Threshold Switching Memristors. Scientific Reports 9, 53 (2019).

[157] S. Yin et al., Emulation of learning and memory behaviors by memristor based on Ag migration on 2D MoS2 Surface. Phys. Status Solidi A 1900104 (2019).

[158] D. Li et al., MoS2 Memristors Exhibiting Variable Switching Characteristics toward Biorealistic Synaptic Emulation. ACS Nano 12, 9240 (2018).

[159] A. Bessonov et al., Layered memristive and memcapacitive switches for printable electronics. Nature Materials 14, 199 (2015).

[160] D. Son et al., Colloidal synthesis of uniform-sized molybdenum disulfide nanosheets for wafer-scale flexible nonvolatile memory. Advanced Materials 28, 9326 (2016).

[161] Y. Hattori et al., Layer-by-layer dielectric breakdown of hexagonal boron nitride. ACS Nano 9, 916 (2015).

[162] H. Zhao et al., Atomically thin Femtojoule memristive device. Advanced Materials 29, 1703232 (2017).

[163] Z. Ding et al., Atomically thin CBRAM enabled by 2-D Materials: scaling behaviors and performance limits. IEEE Transactions on Electron Devices 65, 4160 (2018).

[164] Y. Ji et al., Boron nitride as two-dimensional dielectric: Reliability and dielectric breakdown. Applied Physics Letters 108, 012905 (2016).

[165] C. Pan et al., Coexistence of grain-boundaries-assisted bipolar and threshold resistive switching in multilayer hexagonal boron nitride. Advanced Functional Materials 27, 1604811 (2017).

[166] Y. Shi et al., Electronic synapses made of layered two-dimensional materials. Nature Electronics 1, 458 (2018).

[167] C. Wang et al., 3D monolithic stacked 1T1R cells using monolayer MoS2 FET and hBN RRAM fabricated at low (150 °C) temperature. IEDM 18-528 (2018).

[171] L. Jiang et al., Dielectric breakdown in chemical vapor deposited hexagonal boron nitride. ACS Applied Materials & Interfaces 9, 39758 (2017).

[172] X. Li et al., Transfer of Large-Area Graphene Films for High-Performance Transparent Conductive Electrodes. Nano Letters 9, 4359 (2009).

[173] W. Xie et al., Clean graphene surface through high temperature annealing. Carbon 94, 740 (2015).

[174] J. Lin et al., Smoothening of wrinkles in CVD-grown hexagonal boron nitride films. Nanoscale 10, 16243 (2018).

[175] A. Ferrari, D. Basko. Raman spectroscopy as a versatile tool for studying the properties of graphene. Nature Nanotechnology 8, 235 (2013).

[176] S. Son et al., Graphene hot-electron light bulb: incandescence from hBN-encapsulated graphene in air. 2D Materials 5, 011006 (2018).

[177] S. Long et al., Voltage and power-controlled regimes in the progressive unipolar RESET transition of HfO2-based RRAM. Scientific Reports 3, 2929 (2013).

[178] J. Frenkel. On pre-breakdown phenomena in insulators and electronic semi-conductors. Physical Review 54, 647 (1938).

[179] M. Lanza et al., Electrical and mechanical performance of graphene sheets exposed to oxidative environments. Nano Researches 6, 485 (2013).

[180] B. Wang et al., Experimental observation and mitigation of dielectric screening in hexagonal boron nitride based resistive switching devices. Crystal Research & Technology 53, 1800006 (2018).

1T1M

[181] S. Wachter et al., A microprocessor based on a two-dimensional semiconductor. Nature Communications 8, 14948 (2017).

[182] M. Lanza et al., Standards for the characterization of endurance in resistive switching devices. ACS Nano 15, 17214 (2021). _**Recommended methods to study endurance**

[183] H. Lee et al., Evidence and solution of over-RESET problem for HfOX based resistive memory with sub-ns switching speed and high Endurance. IEDM 19.7.1 (2010).

[184] B. Choi et al., High-speed and low-energy nitride memristors. Advanced Functional Materials 26, 5290 (2016).

[185] A. Torrezan et al., Sub-nanosecond switching of a tantalum oxide memristor. Nanotechnology 22, 485203 (2011).

[186] S. Lee et al., Multi-level switching of triple-layered TaOx RRAM with excellent reliability for storage class memory. Symposium on VLSI Technology (VLSIT), 71 (2012). _**Retention @ 85 °C**

[187] M. Izquierdo et al., Mimicking the spike-timing dependent plasticity in HfO2-based memristors at multiple time scales. Microelectronic Engineering. 215, 111014 (2019). **_STDP**

[189] J. Borghetti et al., 'Memristive' switches enable 'stateful' logic operations via material implication. Nature 464, 873 (2010).

Initial LRS

[191] P. Yao et al., Fully hardware-implemented memristor convolutional neural network. Nature 577, 641 (2020).

[192] C. Wen et al., Advanced data encryption using 2D materials. Advanced Materials 33, 2100185 (2021).

[193] I. E. Ebong, P. Mazumder, CMOS and memristor-based neural network design for position detection. Proceedings of IEEE 100, 2050 (2012).

[194] G. Burr et al., Neuromorphic computing using non-volatile memory. Advanced Physics X 2, 89 (2016).

[195] D. S. Jeong, C. S. Hwang, Nonvolatile memory materials for neuromorphic intelligent machines. Advanced Materials 30, 1704729 (2018).

[196] F. Chiu, A review on conduction mechanisms in dielectric films. Advances in Materials Science and Engineering 2014, 578168 (2014).

[197] K. Zhu et al., Graphene–Boron Nitride–Graphene cross-point memristors with three stable resistive states. ACS Appl. Mater. Interfaces 11, 37999 (2019).

[198] F. Hui et al., In situ observation of low-power nano-synaptic response in graphene oxide Using Conductive Atomic Force Microscopy. Small 17, 2101100 (2021).

[199] H. Wong et al., Metal–oxide RRAM. Proc. IEEE 100, 1951 (2012).

[200] W. Kim et al., VLSI Technol. (VLSI-Technology) Honolulu, HI, USA 22 (2011).

[201] M. Akbari, J. S. Lee, Reproducible and reliable resistive switching behaviors of AlO X/HfO X bilayer structures with Al electrode by atomic layer deposition. RSC Adv. 2017, 6, 21917.

[202] M. R. Mahmoodi et al., Versatile stochastic dot product circuits based on nonvolatile memories for high performance neurocomputing and neurooptimization. Nature Communications 10, 5113 (2019).

[203] Sil'tronix Silicon Technologies, https://www.sil-tronix-st.com/en/silicon-wafer/silicon-thermal-oxide-wafer-sio2 (accessed: November 2020).

[204] I. Yu et al., Surface passivation of c-Si by Atomic Layer Deposition TiO2 thin films deposited at low temperature. IEEE Photovoltaic Specialist Conference (PVSC), 1271 (2014).

[205] S. Chen et al., Wafer-scale integration of two-dimensional materials in high-density memristive crossbar arrays for artificial neural networks. Nature Electronics 3, 638 (2020).

[206] G. González-Cordero et al, In-depth study of the physics behind resistive switching in TiN/Ti/HfO2/W structures. Journal of Vacuum Science & Technology B 35, 01A110 (2017).

[207] N. Xiao et al., Resistive Random Access Memory Cells with a Bilayer TiO2/SiOX Insulating Stack for Simultaneous Filamentary and Distributed Resistive Switching. Adv. Funct. Mater. 2017, 27, 1700384.

[208] S. Aldana et al., Resistive switching in HfO2 based valence change memories, a comprehensive 3D kinetic Monte Carlo approach. Journal of Physics D: Applied Physics 2020, 53, 225106.

Appendix A: Scientific curriculum vitae

PERSONAL INFORMATION

Name: Kaichen Zhu (朱凯晨)
Gender: Male
Date of birth: July 7th, 1995
Nationality: China
Languages: Chinese (native), English
Phone/Email: +86 18862344928 / kczhu1995@126.com
Google Scholar: https://scholar.google.com/citations?user=s-Z-z2QAAAAJ&hl=zh-CN

EDUCATION BACKGROUND

Jun. 2021	PhD in Nanoscience Electronic Engineering Department, University of Barcelona
Sep. 2017 - Jun. 2020	Master in Physics Professor Mario Lanza's research group Institute of Functional Nano & Soft Materials Soochow University
Sep. 2013 - Jun. 2017	Bachelor of Engineering Soochow University

RESEARCH EXPERIENCE

Apr. 2021 - Dec. 2021	Research consultant King Abdullah University of Science and Technology
Sep. 2020 - Mar. 2021	Research assistant Soochow University

RESEARCH EXPERIENCE

❖ Research interest and skills

2D materials and transition metal oxide based memristor for memory and neuromorphic computing; perovskite based light emitting diode.

1- 2D materials handling: High quality 2D materials by mechanical exfoliation, chemical vapor deposition (CVD) and liquid phase exfoliation (LPE), including

graphene, h-BN and MoS_2.

2- Device fabrication: Photolithography/direct laser writer (pattern), electron beam evaporator, magnetron sputtering, atomic layer deposition, inkjet printer (thin film deposition), plasma cleaner (surface modification).
3- Materials and device characterization: Scanning electron microscope, conductive atomic force microscope (C-AFM) (morphology and electrical information), Probe-station with pulse mode, Raman spectroscopy, energy dispersive X-ray spectroscopy (EDS).

❖ Technical software:

1. Familiar with technical software like Origin (for curves and histograms), NanoScope (for AFM data processing), L-edit (for mask pattern design) and Spice for (circuit simulation).
2. Familiar with Solidworks (for basic diagram drawing), and Matlab (for data processing and artificial neural network simulation).
3. with Power Point presentations and Excel data processing.

JOURNAL PAPERS

19. **Kaichen Zhu**, Mohammad Reza Mahmoodi, Zahra Fahimi, Yiping Xiao, Tao Wang, Kristýna Bukvišová, Miroslav Kolíbal, Juan B. Roldan, David Perez, Fernando Aguirre, Mario Lanza*, "Memristors with Initial Low-Resistive State for Efficient Neuromorphic Systems", ***Advanced Intelligent Systems*** 2200001, 2022.

18. Pratik Kumar, **Kaichen Zhu**, Xu Gao, Sui-Dong Wang, Mario Lanza*, Chetan Singh Thakur*, "Hybrid architecture based on two-dimensional memristor crossbar array and CMOS integrated circuit for edge computing", ***npj 2D Materials and Applications*** 8, 2022.

17. Yingwen Liu, **Kaichen Zhu**, Fei Hui, Bin Yuan, Chenhui Zhang, Yinchang Ma, Xixiang Zhang, and Mario Lanza*, "Inkjet Printing: A Cheap and Easy-to-Use Alternative to Wire Bonding for Academics", ***Crystal Research & Technology*** 57, 2100210, 2022.

16. Jiang Liu, Erkan Aydin, Jun Yin, Michele De Bastiani, Furkan H. Isikgor, Atteg Ur Rehman, Emre Yengel, Esma Ugur, George T. Harrison, Mingcong Wang, Yajun Gao, Jafar Iqbal Khan, Maxime Babics, Thomas G. Allen, Anand S. Subbiah, **Kaichen Zhu**, Xiaopeng Zheng, Wenbo Yan, Fuzong Xu, Michael F. Salvador, Osman M. Bakr, Thomas D. Anthopoulos, Mario Lanza, Omar F. Mohammed, Frederic Laquai, and Stefaan De Wolf*, "28.2%-efficient, outdoor-stable perovskite/silicon tandem solar cell", ***Joule***. 5, 1-18, 2021.

15. **Kaichen Zhu**, Chao Wen, Areej A. Aljarb, Fei Xue, Xiangming Xu, Vincent Tung, Xixiang Zhang, Husam N. Alshareef, Mario Lanza*, "The development of integrated circuits based on two-dimensional materials", ***Nature Electronics***. 4, 775-785, 2021.

14. Wenwen Zheng, Fernan Saiz, Yaqing Shen, **Kaichen Zhu**, Yingwen Liu, Clifford McAleese, Ben Conran, Xiaochen Wang, Mario Lanza*, "Defect-free Metal Deposition on Two-dimensional Materials via Inkjet Printing Technology", ***Advanced Materials***. 2104138, 2021.

13. Mario Lanza*, Rainer Waser, Daniele Ielmini, J. Joshua Yang, Ludovic Goux, Jordi Suñe, Anthony Joseph Kenyon, Adnan Mehonic, Sabina Spiga, Vikas Rana, Stefan Wiefels, Stefan Menzel, Ilia Valov, Marco A. Villena, Enrique Miranda, Xu Jing, Francesca Campabadal, Mireia B. Gonzalez, Fernando Aguirre, Felix Palumbo, **Kaichen Zhu**, Juan Bautista Roldan, Francesco Maria Puglisi, Luca Larcher, Tuo-Hung Hou, Themis Prodromakis, Yuchao Yang, Peng Huang, Tianqing Wan, Yang Chai, Kin Leong Pey, Nagarajan Raghavan, Salvador Dueñas, Tao Wang, Qiangfei Xia, and Sebastian Pazos, "Standards for the Characterization of Endurance in Resistive Switching Devices", ***ACS Nano***. 15, 17214-17231, 2021

12. Yaqing Shen, Wenwen Zheng, **Kaichen Zhu**, Yiping Xiao, Chao Wen, Yingwen Liu, Xu Jing, Mario Lanza*, "Variability and Yield in h-BN-Based Memristive Circuits: The Role of Each Type of Defect", ***Advanced Materials***. 2103656, 2021.

11. Chao Wen, Xuehua Li, Tommaso Zanotti, Francesco Maria Puglisi, Yuanyuan Shi, Fernan Saiz, Aleandro Antidormi, Stephan Roche, Wenwen Zheng, Xianhu Liang, Jiaxin Hu, Steffen Duhm, Juan B. Roldan, Tianru Wu, Victoria Chen, Eric Pop, Blas Garrido, **Kaichen Zhu**, Fei Hui, Mario Lanza*, "Advanced Data Encryption using 2D Materials", ***Advanced Materials***. 2100185, 2021. Selected as front cover.

10. Xuehua Li, Tommaso Zanotti, Tao Wang, **Kaichen Zhu**, Francesco Maria Puglisi, Mario Lanza*, "Random Telegraph Noise in Metal-Oxide Memristors for True Random Number Generators: A Materials Study", ***Advanced Functional Materials***. 2102172, 2021.

9. Thales Becker, Xuehua Li, Pedro Alves, Tao Wang, **Kaichen Zhu**, Yiping Xiao, Gilson Wirth, Mario Lanza*, "An Electrical Model for Trap Coupling Effects on Random Telegraph Noise", ***IEEE Electron Device Letters***. 41, 1596, 2020.

8. **Kaichen Zhu**, Bin Yuan, Xianhu Liang, Xu Jing, Chao Wen, Marco A. Villena, Mario Lanza*, "Graphene-Boron Nitride-Graphene Cross-Point Memristors with Three Stable Resistive States", ***ACS Applied Materials & Interfaces***. 11, 37999, 2019.

7. Tao Wang, Yuanyuan Shi, Francesco Maria Puglisi, Shaochuan Chen, **Kaichen Zhu**,

Ying Zuo, Xuehua Li, Xu Jing, Tingting Han, Biyu Guo, Kristýna Bukvišová, Lukáš Kachtík, Miroslav Kolíbal, Chao Wen, Mario Lanza*, "Electroforming in metal-oxide memristive synapses", ***ACS Applied Materials & Interfaces***. 12, 11806, 2020.

6. Ying Zuo, Huizi Lin, Jingchun Guo, Yue Yuan, Hanglin He, Yutong Li, Yiping Xiao, Xuehua Li, **Kaichen Zhu**, Tao Wang, Xu Jing, Chao Wen, Mario Lanza*, "Effect of the Pressure Exerted by Probe Station Tips in the Electrical Characteristics of Memristors", ***Advanced Electronic Materials***, 1901226, 2020.

5. Christoph Metzke, Werner Frammelsberger, Jonas Weber, Fabian Kühnel, **Kaichen Zhu**, Mario Lanza, Günther Benstetter*, "On the Limits of Scanning Thermal Microscopy of Ultrathin Films", ***Materials***, 13, 518, 2020.

4. Marco A. Villena, Fei Hui, Xianhu Liang, Yuanyuan Shi, Bin Yuan, Xu Jing, **Kaichen Zhu**, Shaochuan Chen, Mario Lanza*, "Variability of metal/h-BN/metal memristors grown via chemical vapor deposition on different materials", ***Microelectronics Reliability***, 102, 113410, 2019.

3. Mario Lanza*, H.-S. Philip Wong, Eric Pop, Daniele Ielmini, Dimitri Strukov, Brian C. Regan, Luca Larcher, Marco A. Villena, J. Joshua Yang, Ludovic Goux, Attilio Belmonte, Yuchao Yang, Francesco M. Puglisi, Jinfeng Kang, Blanka Magyari-Köpe, Eilam Yalon, Anthony Kenyon, Mark Buckwell, Adnan Mehonic, Alexander Shluger, Haitong Li, Tuo-Hung Hou, Boris Hudec, Deji Akinwande, Ruijing Ge, Stefano Ambrogio, Juan B. Roldan, Enrique Miranda, Jordi Suñe, Kin Leong Pey, Xing Wu, Nagarajan Raghavan, Ernest Wu, Wei D. Lu, Gabriele Navarro, Weidong Zhang, Huaqiang Wu, Runwei Li, Alexander Holleitner, Ursula Wurstbauer, Max C. Lemme, Ming Liu, Shibing Long, Qi Liu, Hangbing Lv, Andrea Padovani, Paolo Pavan, Ilia Valov, Xu Jing, Tingting Han, **Kaichen Zhu**, Shaochuan Chen, Fei Hui, Yuanyuan Shi, "Recommended methods to study resistive switching devices", ***Advanced Electronic Materials***, 1800143, 2019. Selected as Excellence in Electronics and most assessed paper.

2. Fei Hui, Marco A. Villena, Wenjing Fang, Ang-Yu Lu, Jing Kong, Yuanyuan Shi, Xu Jing, **Kaichen Zhu**, Mario Lanza*, "Synthesis of large-area multilayer hexagonal boron nitride sheets on iron substrates and its use in resistive switching devices", ***2D Materials***, 5, 031011, 2018.

1. Bingru Wang, Na Xiao, Chengbin Pan, Yuanyuan Shi, Fei Hui, Xu Jing, **Kaichen Zhu**, Biyu Guo, Marco A. Villena, Enrique Miranda, Mario Lanza*, "Experimental observation and mitigation of dielectric screening in hexagonal boron nitride based resistive switching devices", ***Crystal Research and Technology***, 53, 1800006, 2018.

CONFERENCE PAPERS

15. **Kaichen Zhu**, Giovanni Vescio, Sergio Gonzalez-Torres, Julia Lopez-Vidrier, Juan Luis Frieiro, Xu Jing, Mario Lanza*, Albert Cirera, Blas Garrido*, "Fully inkjet printed h-BN memristors", ***2021 Graphene and 2DM Online Conference (GO2021)***, 20th April – 21st April 2021, Virtual conference. – Oral presentation. Agenda: http://www.confstreaming.com/GO2021/program.php

14. **Kaichen Zhu**, Ying Zuo, Huizi Lin, Jingchun Guo, Yue Yuan, Hanglin He, Yutong Li, Yiping Xiao, Xuehua Li, Tao Wang, Xu Jing, Chao Wen, Mario Lanza*, "Effect of the Pressure Exerted by Probe Station Tips in the Electrical Characterization of Memristors", ***2020 International Integrated Reliability Workshop (IIRW)***, 1st October – 1st November 2020, Virtual conference. – Poster presentation.

13. **Kaichen Zhu**, Giovanni Vescio, Sergio González-Torres, Julià López-Vidrier, Mario Lanza*, Albert Cirera, Blas Garrido*, "Two-dimensional material based memristor fabricated by inkjet printing technology", ***2020 European Materials Research Society (E-MRS)***. – Oral presentation. Agenda: www.european-mrs.com/solid-state-ionics-advanced-concepts-and-devices-emrs-1. (Conference cancelled due to COVID-19).

12. **Kaichen Zhu**, Yiping Xiao, Tao Wang, Fei Hui, Chao Wen, Mario Lanza*, "Forming free metal oxide based memristors with initial low resistive state", ***2020 International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)***. – Abstract accepted. (Conference cancelled due to COVID-19).

11. **Kaichen Zhu**, Giovanni Vescio, Sergio González-Torres, Julià López-Vidrier, Mario Lanza*, Albert Cirera, Blas Garrido*, "Two-dimensional material based memristor fabricated by inkjet printing technology", ***2020 International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)***. – Abstract accepted. (Conference cancelled due to COVID-19).

10. **Kaichen Zhu**, Xianhu Liang, Bin Yuan, Marco A. Villena, Chao Wen, Tao Wang, Shaochuan Chen, Mario Lanza*, Fei Hui, Yuanyuan Shi, "Tristate resistive switching in heterogenous van der Waals dielectric structures", ***2019 IEEE International Reliability Physics Symposium (IRPS)***, 11th March – 4th April 2019, Monterey, California (USA). DOI: 10.1109/IRPS.2019.8720485. – Poster presentation.

9. **Kaichen Zhu**, Xianhu Liang, Bin Yuan, Marco A. Villena, Fei Hui, Yuanyuan Shi, Mario Lanza*, "Tri-state cross point memristors based on heterogeneous van der Waals structure", ***China Semiconductor Technology International Conference (CSTIC)***. March 18th – 19th, 2019, Shanghai (China) – Oral presentation. Agenda: http://www.semiconchina.org/en/138

8. E. Miranda, J. Munoz-Gorriz, **Kaichen Zhu**, T. Wang, J. Sune, M. Lanza*, "SPICE Model for the Conduction Characteristics of Hexagonal Boron Nitride-Based Resistive Switching Devices", ***The International Conference Nano-M&D***. June 4th – 8th, 2019 (Italy) – Poster presentation.

7. **Kaichen Zhu**, Bin Yuan, Xianhu Liang, Xu Jing, Chao Wen, Mario Lanza*, "Development of memristors based on G/h-BN/G van der Waals structures", ***2nd International Workshop on Future Computing***. December 17th-18th, 2018, Shenzhen (China) – Poster presentation.

6. Chao Wen, Xianhu Liang, Bin Yuan, **Kaichen Zhu**, Tingting Han, Mario Lanza*, "Non-conductive dielectric breakdown in multilayer h-BN stacks", ***2nd International Workshop on Future Computing***. December 17th-18th, 2018, Shenzhen (China) – Poster presentation.

5. Tao Wang, Biyu Guo, Zi He, **Kaichen Zhu**, Xianhu Liang, K. Bukvišová, M. Kolíbal, Mario Lanza*, "Resistive switching in ALD and sputtered TiO_2 memory device: electroforming observed by constant voltage bias", ***2nd International Workshop on Future Computing***. December 17th-18th, 2018, Shenzhen (China) – Poster presentation.

4. IEMS 2018 – The International Emergent Memory Symposium, August 31st – September 1st 2018, Jiangxi, China

3. IPFA 2018 - 2018 International Symposium on the Physical and Failure Analysis of Integrated Circuits, July 16th -19th 2018, Singapore.

2. MEMRISYS 2018 – International Conference on Memristive Materials, Devices & Systems, July 3rd – 6th 2018, Beijing

1. ChinaRRAM 2017 – The first International Workshop on RRAM, June 12th – 13th, 2017, Suzhou.

AWARDS AND REVIEWER

- ❖ International Corroboration Scholarship for working at Universitat de Barcelona during 12 months, Soochow University (January 2020)
- ❖ Reviewer of Scientific Reports (IF = 4.011 – 2018) (since January 2019)
- ❖ Reviewer of Microelectronic Engineering (IF = 1.654 – 2018) (Since August 2019)
- ❖ Best Poster Presentation Award, 2nd International Workshop on Future Computing (IWOFC 2018) (December 2018)
- ❖ 2019 Postgraduate Academic Scholarship, Soochow University (November 2019)

- 2018 Postgraduate Academic Scholarship, Soochow University (October 2018)
- 2017 Postgraduate Academic Scholarship, Soochow University (October 2017)
- CNST Excellence Award of College of Nano- Science and Technology, Soochow University (October 2013)

INTERNSHIP

- 2015-07 - 2015-08 Volunteer teaching to Gansu Province to support education
- 2013-10 - 2015-06 Secretary of the Minister of Soochow University Student Mental Health Association

Appendix B: Summary in official language

En conclusión, en esta tesis doctoral he realizado un profundo análisis del memristor basado en h-BN, desde un solo dispositivo hasta circuitos integrados.

Los principales objetivos de esta tesis han sido diseñar, fabricar, caracterizar y optimizar memristores basados en h-BN, con el objetivo principal de lograr una integración a nivel de circuito con la tecnología CMOS basada en Si. Comparando los objetivos principales, concluyo mi tesis doctoral de la siguiente manera:

1- El h-BN producido mediante los métodos *Chemical Vapour Deposition* (CVD) y *Liquid Phase Exfoliated* (LPE) (consulte el Apéndice D: *Inkjet-printed h-BN based memristor*) se puede utilizar como un medio de conmutación resistivo confiable para memristores basados en materiales 2D. Ambos métodos son escalables para la integración de áreas grandes.

2- La estructura en capas del h-BN multicapa ha sido confirmada por el TEM transversal, y el espesor es de alrededor de 6 nm (18 capas). La naturaleza bidimensional del h-BN y del Grafeno ha sido confirmada estadísticamente por el espectro Raman.

3- Los defectos locales en la estructura de capas se confirmaron como debilidades desde el punto de vista eléctrico del material que pueden funcionar como la región de conmutación confinada para memristores. Esto fue confirmado por el microscopio de fuerza atómica conductora (CAFM).

4- Con estructura en forma de matriz y fotolitografía se han fabricado memristores basados en h-BN con un tamaño de 25 μm^2. Utilizando una única vía de interconexión entre los niveles de Metal 3 y 4, se logran obtener memristores basados en h-BN tan pequeños como 0,053 μm^2.

5- Los memristores h-BN con inserción de grafeno muestran un funcionamiento tri-estado estable y altamente controlable, al controlar la corriente límite (CC) y el voltaje de reset (V_{RESET}).

6- La estructura de un transistor y un memristor (1T1M) puede mejorar en gran medida el rendimiento del memristor h-BN y proporcionar un control más preciso del nivel de resistencia, usando un electrodo de tungsteno (W) como electrodo inerte inferior (W).

7- La celda 1T1M con electrodo superior Ti puede lograr una operación de baja corriente y conmutación multinivel. La celda 1T1M con electrodo superior Au puede lograr una alta durabilidad de 2,5 millones de ciclos. La celda 1T1M con electrodo superior Ag puede lograr un funcionamiento rápido y un bajo consumo de energía con una energía de reset de 1,41 pJ.

8- La celda 1T1M con electrodo superior Ti puede mostrar una conmutación resistiva tanto volátil como no volátil en función del voltaje de puerta (V_G) utilizado. Si V_G es de 1.5 V, el dispositivo muestra un comportamiento no volátil. Sin embargo, si VG es de 0.5 V, ahora la conmutación resistiva es volátil.

9- La celda 1T1M con electrodo superior Ti puede mostrar un comportamiento sináptico tipo *spike-time-dependent-plasticity* (STDP).

10- Mediante el uso de la estructura de celda 1T1M, las corrientes parásitas generadas en la matriz de dispositivos se pueden suprimir eficientemente.

11- Con dos celdas 1T1M podemos realizar operaciones lógicas como "OR" o "IMP".

12- Según la característica STDP, se modeló una red neuronal de picos (SNN) con el conjunto de datos MNIST obteniendo una precisión promedio del ~90 % reconociendo números manuscritos.

Hemos demostrado que es posible fabricar microchips híbridos CMOS/2D-LMs y que puede proporcionar un rendimiento eléctrico notable. En particular, hemos integrado h-BN multicapa en las interconexiones de final de línea de chips de silicio que contienen transistores CMOS organizados en una configuración de matriz. La matriz resultante de celdas 1T1M muestra un rendimiento muy superior al de los memristores h-BN independientes.

El trabajo futuro debe centrarse en la implementación completa del hardware del SNN y lograr la clasificación in situ; multiplicación de matriz vectorial para redes neuronales artificiales (ANN) y electrónica completamente basada en materiales 2D. Aunque la alta velocidad de operación se puede lograr en un solo memristor h-BN (tiempo de conmutación inferior a 10 ns), el tiempo de conmutación de la celda 1T1M es mucho más largo, de cientos de ns o incluso unos pocos µs. La integración debe ser optimizada. La retención de alta temperatura debe mejorarse con una mejor conexión o integración entre los materiales 2D y la vía CMOS.

Appendix C: Experimental equipment and materials

Experimental types of equipment

Equipment	Model	Manufacturers
Ultrasonic cleaner	KQ-100KDB	Kunshan Chaosheng
Mask aligner	MJB4	SUSS MicroTec
Electron beam evaporator Magnetron sputter	PVD75	Kurt J Lesker
Plasma stripper	IoN 40	PVA TePla
Optical microscope	DM4000M	Leica
Probe station	M150	Cascade Microtech
Semiconductor parameter analyzer	4200	Keithley
Semiconductor parameter analyzer	B1500	Aglient
Scanning electron microscope	Supra 55	Carl Zeiss
Transmission electron microscope	JEM-2100	JEOL
Raman spectrometer	HR800	Jobin Yvon

Experimental materials/reagents

Materials/reagents	Model	Manufacturers
Silicon oxide wafer	300 nm SiO_2	Yancai Micro&Nano
Ethanol	Analytical pure	Yonghua Chem
Acetone	Analytical pure	Qiangsheng Chem
Photoresistive with the developer: positive & negative	AR-P 5350 & AR 300-26	Allresist
Lithographic mask	3 inch, Suda glass	Suzhou Zhiban
Pure Au/Ti	99.999%	Zhongnuo Xincai
PMMA powder	996000	Sigma-Aldrich

Appendix D: Inkjet-printed h-BN based memristor

Multilayer hexagonal boron nitride (h-BN) synthesized by chemical vapour deposition (CVD) has shown interesting resistive switching (RS) behaviour. However, the high temperatures (>900 ºC) employed during the CVD synthesis complicate the integration on silicon wafers containing pre-patterned circuitry, and its use requires a delicate transfer process that introduces contamination and reduces the yield due to the generation of cracks. Here we report the easy and cheap fabrication of h-BN memristors via liquid phase exfoliation (LPE) and inkjet-printing technology at low temperatures below 100 ºC. The Ag/h-BN/Ag devices exhibit random telegraph noise signals with a low energy consumption when biased at 0.1V, as well as stable volatile unipolar RS for over 0.6 million cycles for high resistances up to 10^7 Ω in low resistive state. At lower resistance ranges, the devices exhibit vera multiple stochastic phenomena that are suitable for data encryption applications, including a high device-to-device variability of the initial resistance and forming voltage, non-volatile bipolar resistive switching with a high cycle-to-cycle variability of the state resistances — which is highly reliable because never gets stuck permanently at a given resistive state — and low-energy random telegraph noise. Moreover, the devices can be operated in h-We also observed that the devices exhibit random telegraph noise when applying a constant bias, which produces stochastic switching between different states by consuming very low energy. These behaviours are interesting in the field of data encryption, as they could be exploited to construct physical unclonable functions and true random number generators directly on silicon wafers containing complementary metal-oxide-semiconductor circuits.

1. Introduction

Two-dimensional (2D) layered materials have exhibited outstanding physical, chemical, mechanical, electrical, optical and magnetic properties, which have been used for a wide range of applications. Among them, the fabrication of solid-state microelectronic devices and circuits made of 2D materials has attracted much attention,

as it may be a solution to produce advanced electronic devices and circuits beyond the complementary metal-oxide-semiconductor (CMOS) technology. However, most 2D materials based electronic devices consist on prototypes fabricated using methods that are not compatible with wafer-scale circuital technologies, such as mechanical exfoliation of small crystals. Chemical vapor deposition (CVD) is a scalable method to produce large-area 2D materials; however, CVD method requires the use of high temperatures >800 ºC, which impedes the direct growth of the 2D material on wafers containing complementary metal-oxide-semiconductor (CMOS) circuits. This represents an important complication to the fabrication process, as it makes necessary the use of a delicate process to transfer the 2D material from the wafer used for the growth to the target wafer. During this his process, which is time-consuming, the CVD-grown 2D materials get normally contaminated and cracked.

Liquid phase exfoliation (LPE) is an inexpensive method to produce large amounts of few-nanometer-thick and few-micrometer-long sheets of different 2D layered materials, and it consists on the use of an ultrasonic bath to exfoliate sheets from small crystals —this is possible because the van der Waals forces that make the 2D planes attach to each other in the crystals are much weaker than the in-plane covalent bonging. Moreover, 2D materials synthesized by LPE method can be deposited on any wafer at low temperature (by inkjet printing, screen printing and spray, among others) to pattern wires, electrodes, channels and other complex devices with irregular shapes. This method has been used to fabricate a plethora of solid-state microelectronic devices, including transistors, photodetectors, capacitors, solar cells and light emitting diodes (among others).

Memristors are emerging devices that are being commercialized as non-volatile memory (with a growing market size), and that have exhibited huge potential for data computation, encryption and radio-frequency communication. Memristors made of 2D materials have shown interesting performance beyond flexibility and transparency, such as coexistence of volatile and non-volatile switching, controllable potentiation and depression, ultra-low power consumption, and high thermal stability. However, few studies reported 2D-materials-based memristors made by LPE method. However, in all these prototypes the dielectrics used (e.g. graphene oxide, MoS_2, BP) are not very good electrical insulators, which results in low endurance and short retention times. On the contrary, hexagonal boron nitride (h-BN) has a higher bandgap of ~6 eV and has

exhibited low leakage currents under polarization, which has been used to fabricate high-performance memristors using CVD method.

In this work, we have formulated printable h-BN inks from 2D crystals and afterward produced memristors with Ag/h-BN/Pt and Ag/h-BN/Ag structure (top-to-down) entirely by inkjet-printing. The devices show a high device-to-device variability of the initial resistance and forming voltage, as well as a high cycle-to-cycle variability of the state resistances. These electrical features, combined with a high reliability and simple low-temperature fabrication could enable the use of h-BN memristors for data encryption circuits.

2. Results and discussion

Figure 1a shows the photograph of the h-BN ink, which consists on a carboxy metyl cellulose solution containing the h-BN nanoflakes; it shows a characteristic whitish colour. Figure 1b shows the top-view scanning electron microscopy image of an Ag/h-BN/Pt memristor as fabricated on a 300 nm SiO_2/Si wafer. The h-BN nanoflakes form a continuous film with a thickness of ~250 nm and a surface roughness of 5.6 nm, as demonstrated via cross-sectional scanning electron microscopy (Figure 1d) and top-view atomic force microscopy (Figure 1c), respectively. Cross-sectional transmission electron microscopy (TEM) images of the Ag/h-BN/Pt devices reveal that the crystalline structure of the h-BN nanoflakes is excellent (see Figure 1e-f and Supplementary Figures 1-2), and that their thickness fluctuates between 5 and 15 nm, which corresponds to 15 to 45 layers — one monolayer of h-BN is around 0.33 nm thick). Top-view TEM inspection of the flakes on a microperforated copper grid reveals that the average lateral length is ~210 nm. The good crystal quality of the h-BN flakes was further confirmed via Raman spectroscopy.

We investigate the switching voltages and state currents of the all the devices by applying sequences of ramped voltage stresses (RVS) and plotting the current versus voltage (I-V), and we test the reproducibility and the cycle-to-cycle variability of the switching by applying voltage list stresses (VLS) and pulsed voltage stresses (PVS). Constant voltage stresses (CVS) are also applied to test the retention time at a given voltage at room temperature and 85 ºC — this is the typical temperature at which reliability tests are normally performed. The stresses are always applied on the top electrode of the memristors, keeping bottom electrode grounded.

During the first RVS the Ag/h-BN/Pt memristors are initially in HRS, and they transition to a LRS sharply as the voltage increases. The initial resistance and the dielectric breakdown voltage of the devices exhibits a very high device-to-device variability: from rom ~10^7 to ~10^{11} Ω and from ~0.4 to ~16 V (respectively, see Figure 2a-b). This variability is related to the different structure of the devices, i.e., thickness fluctuations and formation of random junctions between the h-BN flakes, as well as different amounts of solvent present between the Pt and Ag electrodes. While this inherent behaviour is really bad for most memristive applications (memory, computation, communication), it can be exploited to fabricate physical unclonable functions, i.e., circuits in which multiple devices are in an unpredictable state, and that can be used as unclonable fingerprint within a system. The sharp current increase and it high nonlinearity suggest that the switching mechanism is filamentary, most probably due to the drift of Ag^+ ions from the top Ag electrode into the h-BN film when this is positively biased.

When positive and negative RVS are applied sequentially to the devices they exhibit correct non-volatile bipolar RS at high current limitations between 1 and 500 μA (see Figure 2c). When few RS cycles (<50) are measured via RVS the cycle-to-cycle variability of the state resistances are low (Figure 2d). However, when the devices are stressed using PVS the cycle-to-cycle variability of the state resistances remarkably increases, and it is remarkably higher in HRS (Figure 2e). The devices programmed to LRS tend to self-switch back to the HRS (i.e., recover) after some time if the bias is reduced below 0.1V, and the higher the resistance of the LRS the longer the recovery time (see Figure 2f). These observations further indicates that the RS mechanism is filamentary, and it is consistent with the diffusive nature of the Ag^+ ions penetrating into the dielectric (in our case LPE h-BN).

When the current limitation is reduced below 1 μA, the devices exhibit volatile unipolar RS, also known as threshold-type (Figure 3a-d). This switching behaviour is very reproducible in all devices analysed, and in some devices it can reach high switching endurances up to 0.6 million cycles (see Figure 3e) — this is the highest endurance ever reported for any memristor fabricated by LPE method. This behaviour could be exploited to fabricate electronic neurons for artificial neural networks. At this volatile unipolar regime we also find a good source of randomness when applying CVS, which consists on random fluctuations of the current across the Ag/h-BN/Pt memristor between different stable states (see Figure 4). This type of behaviour is known as

random telegraph noise (RTN), and it could be employed to fabricate true random number generators (TRNG) for one time password generation and data encryption. This behaviour is related to the trapping and de-trapping of electrons at the local defects in the dielectric, a behaviour that should take place at the junctions between the h-BN nanoflakes, as these are defect free (see Figure 1e). The advantage of stochastic RTN (Figure 4) compared to stochastic resistive switching (Figure 2e) is the lower energy consumption because the currents never exceed the nanoampere regime (no filament is formed and disrupted in every transition). However, the ratio between the resistance states is much lower (~2), which complicates the hardware necessary to identify the state.

The randomness of the RS behaviour in LPE h-BN memristors could be enhanced by using a bottom Ag electrode.

The fully-printed Ag/h-BN/Ag devices also exhibit stable non-volatile bipolar RS when applying sequences of RVS, but the cycle-to-cycle variability of the resistance in each state is much higher than in the Ag/h-BN/Pt devices (see Figure 5a). One genuine feature of these devices is that they never produce failure by getting stuck in one of the resistive states, but they simply show random fluctuations of the conductance from one cycle to another. This can be better observed in LVS stresses (see Figure 5b-d).

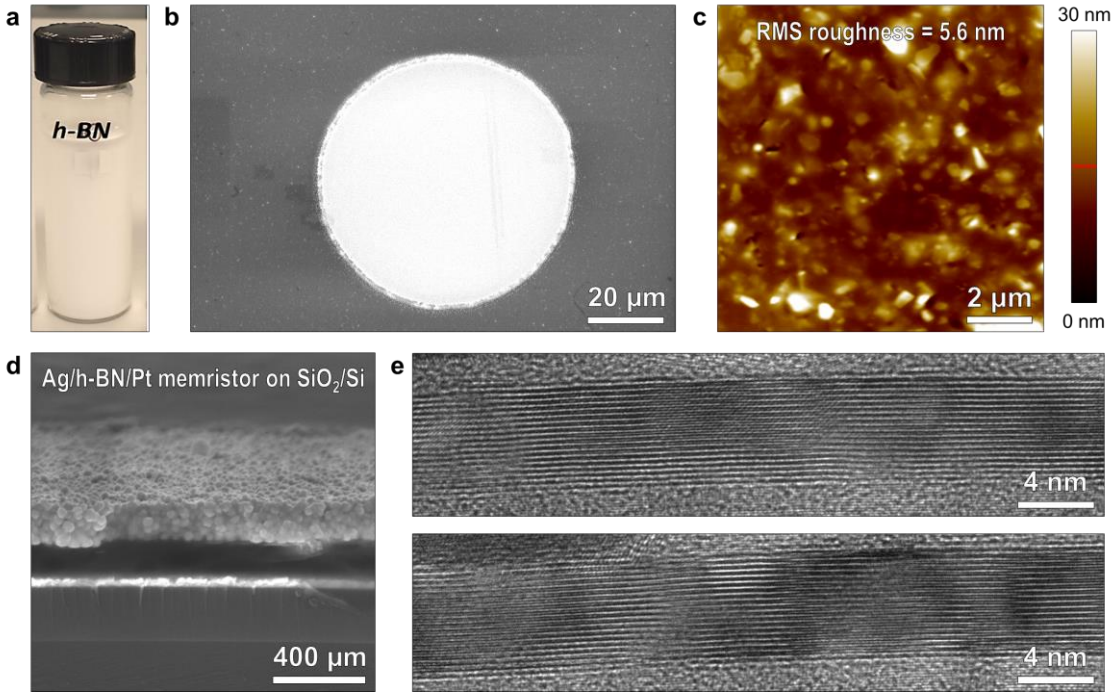


Figure 1. (a) Photo of liquid exfoliated h-BN ink. (b) SEM image of a single h-BN memristor. AFM (c), SEM (f) and (XTEM) images of inkjet printed h-BN thin films.

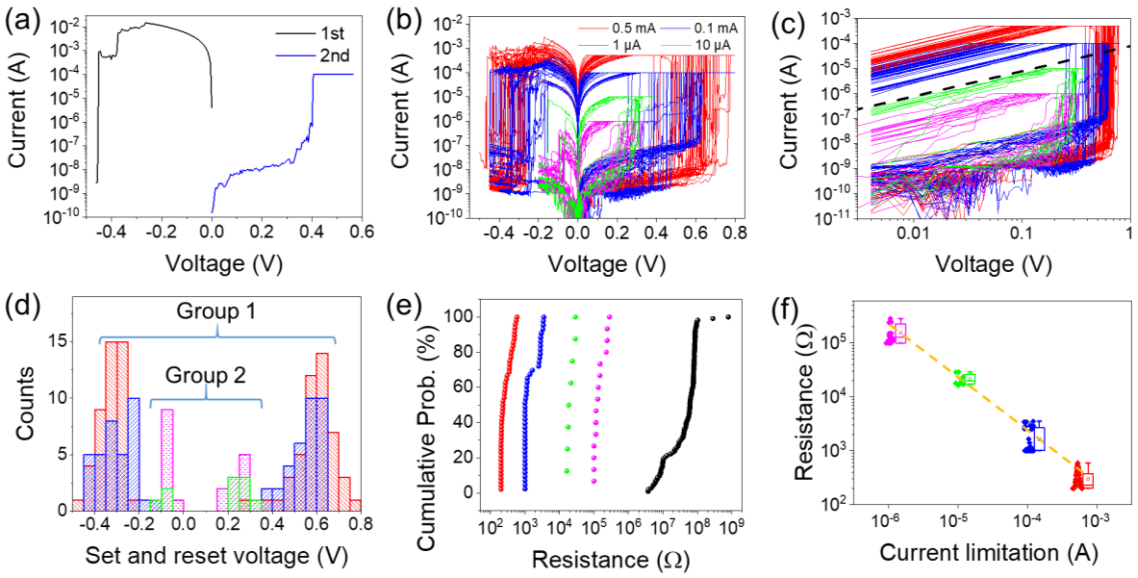


Figure 2. (a) Device starts from initial LRS, which could be Ag penetration into h-BN layer due to high T. Typical curing temperature for Ag ink is 100 ^{o}C. (b) I-V curves displaying 200 cycles bipolar RS stressed by RVS mode with different current limitation on the same device. (c) Replot (b) in log-log scale. Quantum point conduction black dash line is drawn for reference. (d) Statistical analysis of set/reset voltages of different CL in (b). The set/reset voltage can be seperated into two groups, purple & green, and red & blue, which is agreement with the seperation above and below the QPC line in (c). (e) The distribution of the resistance read at 0.1 V using different current limitation. (f) Box plot of resistance vs. current limitation. In this figure, same color shares the same current limitation.

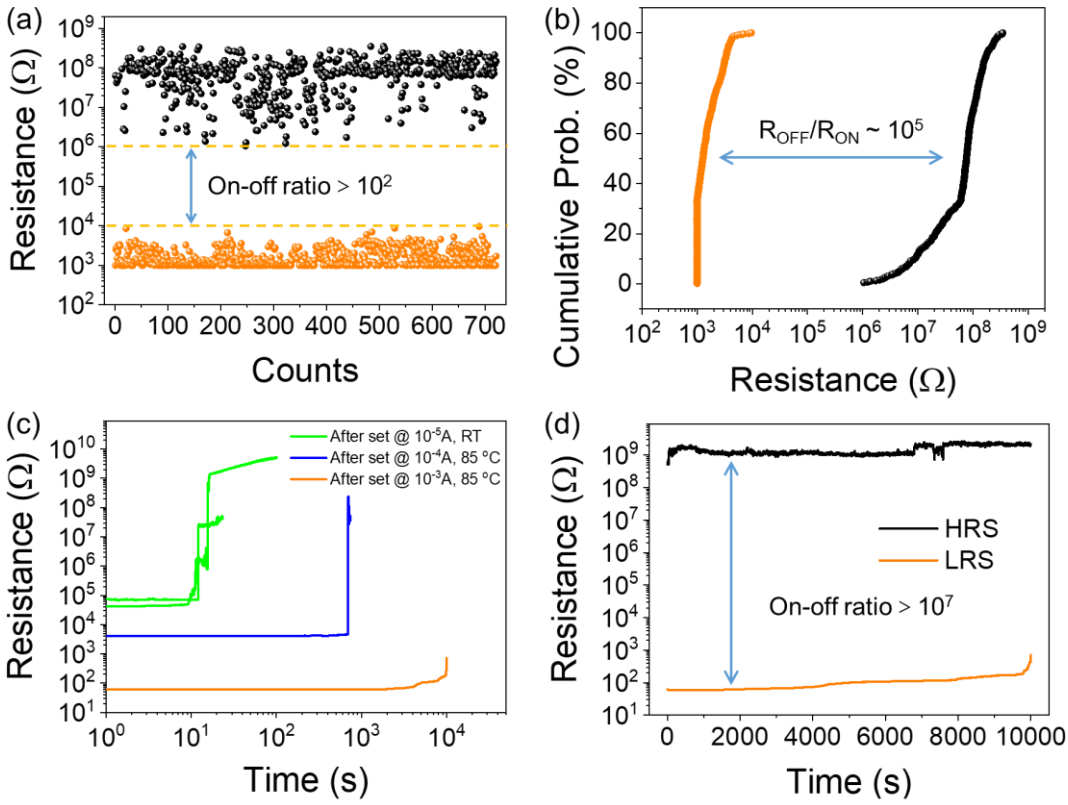


Figure 4. Endurance and retention test. Resistance versus cycle (a) and cumulative probability versus resistance plots (b) obtained under 0.1 V read voltage using VLS mode with delay 1s. 1 mA CL is applied during the set process. Set voltage varies from 0.4 to 0.8 V. Reset voltage varies from -0.4 to -0.5 V. Retention test measured with constant voltage stress 0.1 V. (c) First we set the device to LRS with different current limitation (CL), then we applied constant voltage stress 0.1 V to do the retention. (d) Set with current limitation 1 mA. Retention test read at 0.1 V 85 ^{o}C of LRS after set at 1 mA (orange line) and of HRS after reset in RVS mode (black line).

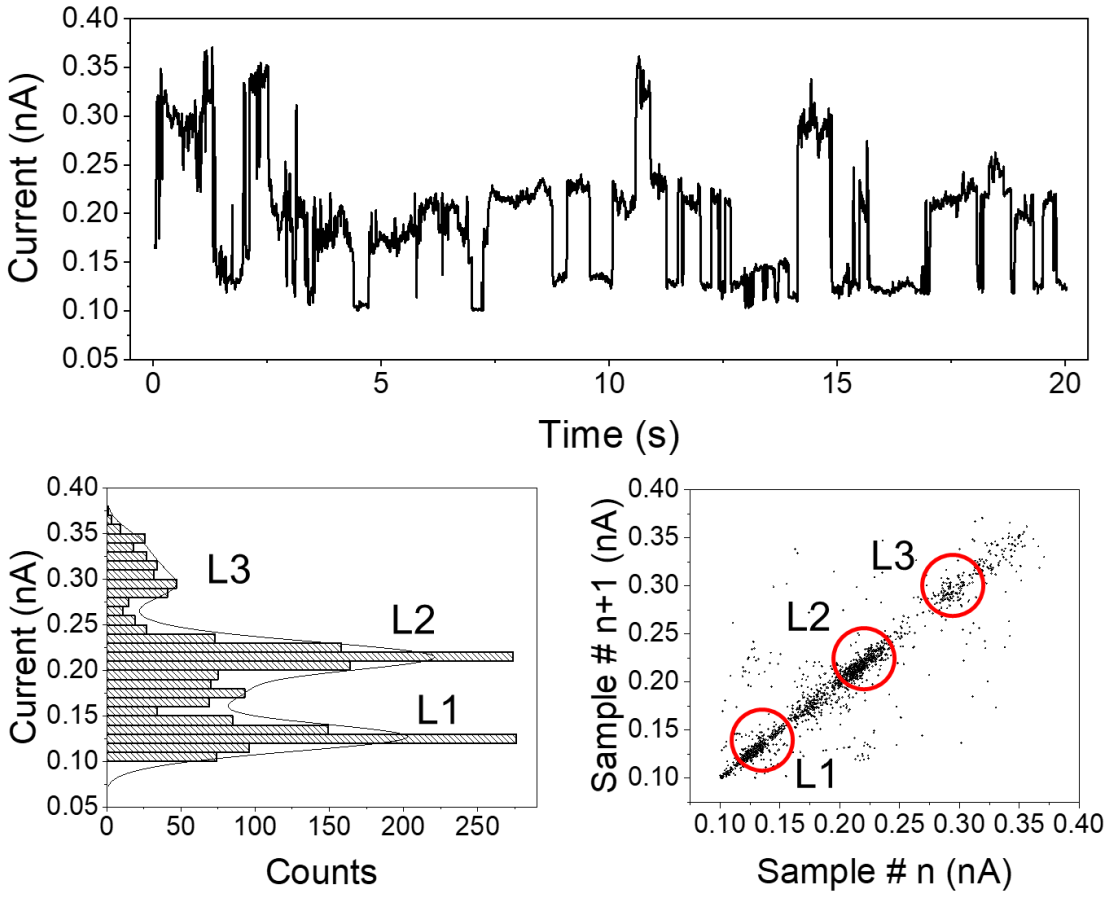


Figure 5. (a) RTN signal read at 0.1 V with time interveral 10 ms generated by two defects, resulting in three descrete levels. (b) Corresponding histrogram representation showing three main states revealed by three peaks with a Kernel Smooth fitting. (c) Corresponding time-lag plot representation.

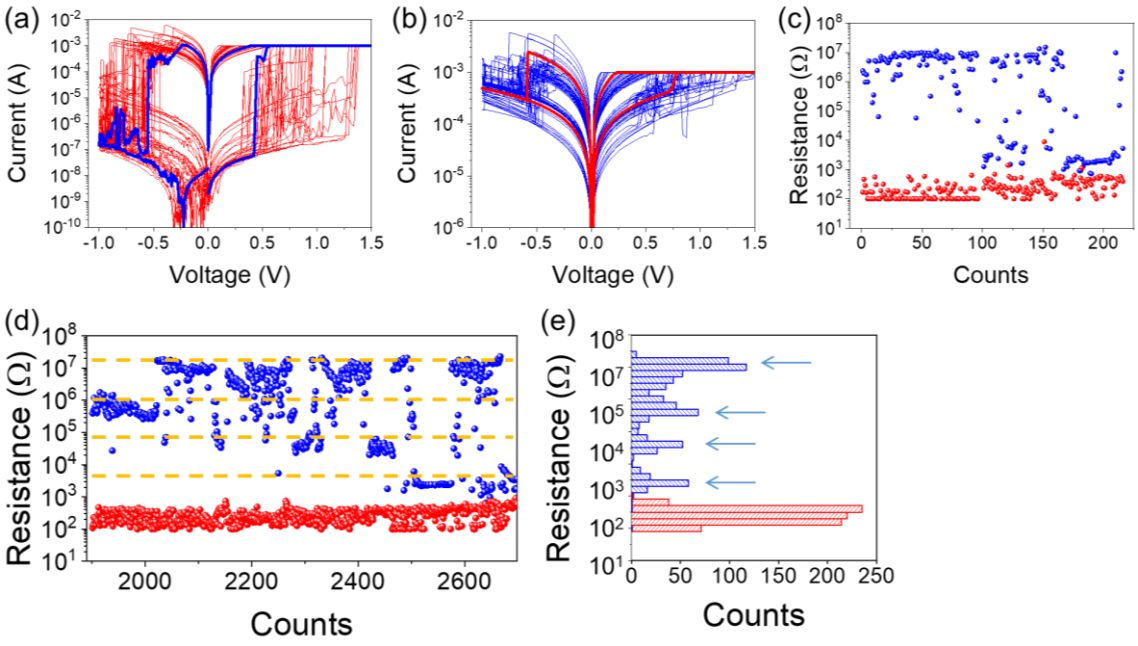


Figure 6. First 100 I-V curves shows bipolar switching between HRS and LRS (a) and the last 60 I-V curves shows bipolar switching between intermidiate HRS and LRS (b). (c) Resistance versus cycle plot showing the trends from (a) to (b). (d) More states are revealed by resistance vs cycles plot by using VLS mode. It is replotted in histogram plot (e), which shares the same y axis. The relatively stable states are the peaks highlighted by the arrows. It is interesting to see each state has a similar separation of one order of magnitude.

Appendix E: Home made 1T1M pretest

After the above work on the three resistance state memristor of graphene hexagonal boron nitride graphene van der Waals structure is completed satisfactorily. In order to solve the current overshoot problem of memristor in the setting process (even if the limiting current is applied) and move towards the development direction of practical industrialization, I continue to try to build 1t1r memory array. The grid voltage can control the current flowing through the source drain and memristor, so as to effectively control the current limiting level from the circuit.

A normal source / measurement unit instrument for semiconductor parameter analysis needs about 100-500 μ S to really reach the level of limiting current [1]. When the current in the high resistance state increases to the low resistance state instantaneously under the set voltage, the instrument needs 100-500 even if the limiting current is applied μ S, during which the actual current may be one order of magnitude higher than the limiting current, that is, the so-called overshoot phenomenon. The research group of teacher Zhou Peng of Fudan University used an oscilloscope to observe in situ and found that μ Under the current limit of a, the overshoot will reach Ma level [2]. The experiment also proves that 1t1r architecture has higher reliability than 1R architecture. In the future integrated circuit level architecture, 1R arrays can be constructed, but a transistor is integrated at the end of each word line and bit line to realize 1t multi r arrays, so as to achieve lower energy consumption and higher work efficiency.

In this experiment, I try to construct a full two-dimensional material based 1t1r structure, in which graphene will be used as the channel material of the transistor, and h-BN will be used as the gate insulation layer of the transistor and the resistive layer of the memristor. However, untreated graphene based transistors show high conductivity and almost negligible switching ratio. In the follow-up work, the more traditional but mature channel semiconductor material indium tin oxide (izgo) [3] is used to ensure the normal operation of the device, and then continue to try the full two-dimensional material-based device.

5.2 design and preliminary physical implementation of 1t1r optical mask

Figure 5.1 is the design diagram of the lithography mask designed by me for patterning 1t1r storage array architecture, in which the light blue frame is the memristor with

intersection structure, and the pink pattern is the top electrode; The device in the red box is the top gate transistor, the blue pattern is the source drain electrode, the gray pattern is the gate electrode, and the green pattern is the channel material of the transistor. There are four patterns in a set of templates for patterning different parts in 1t1r structure. In the process of mask design, pay attention to the linewidth that can be achieved by laboratory instruments and the optical registration between different masks, so as to avoid the problem of short circuit in the actual preparation of devices.

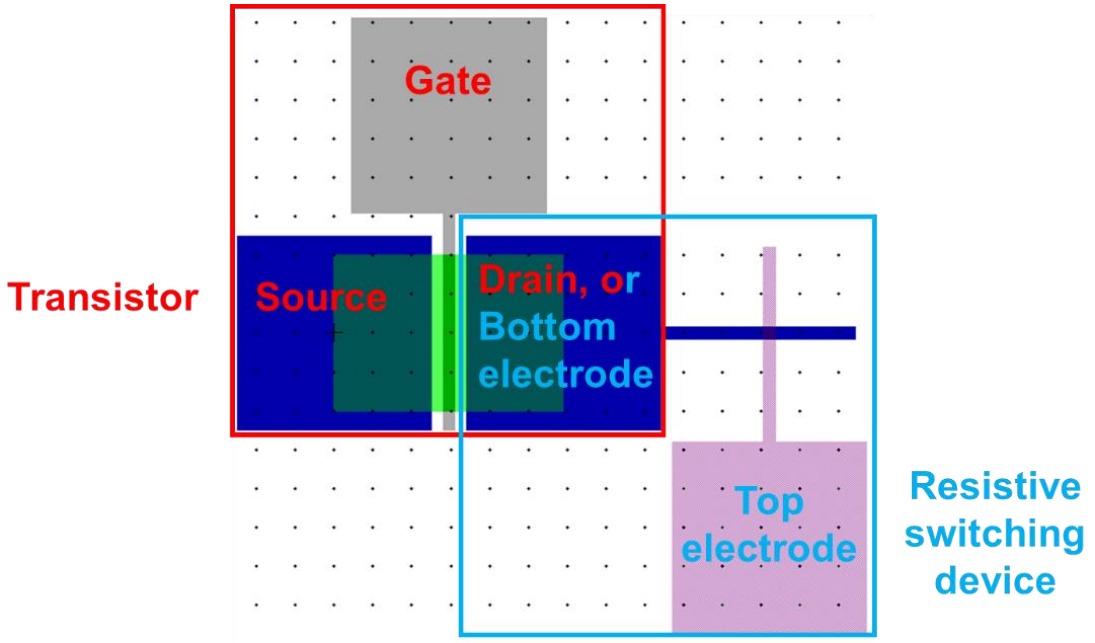


Figure 5.1 lithography mask design of 1t1r storage array architecture. In the red box is the 1t transistor part, and in the blue box is the 1R memristor resistance transformer.

The actual preparation process is shown in Fig. 5.2. Firstly, a 300 nm SiO2 / Si substrate (Fig. 5.2a) is prepared, and then a piece of multilayer h-BN prepared by wet transfer CVD method is used as the substrate of the whole architecture (Fig. 5.2b), so as to obtain better roughness and reduce the electron scattering of graphene channel material. In the third step, a piece of multilayer graphene prepared by CVD method is transferred by wet method as channel material (Fig. 5.2c). In the fourth step, the source drain electrode is synchronously patterned and evaporated by photolithography and electron beam evaporation (Fig. 5.2d). In the fifth, sixth and seventh steps, I patterned the channel material graphene by photoresist negative glue and oxygen plasma etching to quantify the channel width. In step 8, a piece of multi-layer h-BN is transferred by wet transfer as the gate insulation layer of 1t part, and this layer h-BN will also be used as the resistive layer of 1R part. Finally, I photolithographed and evaporated the gate

electrode of 1t part and the top electrode of 1R part; These two electrodes can be obtained step by step, that is, different metal materials can be evaporated to 1t gate electrode and 1R top electrode.

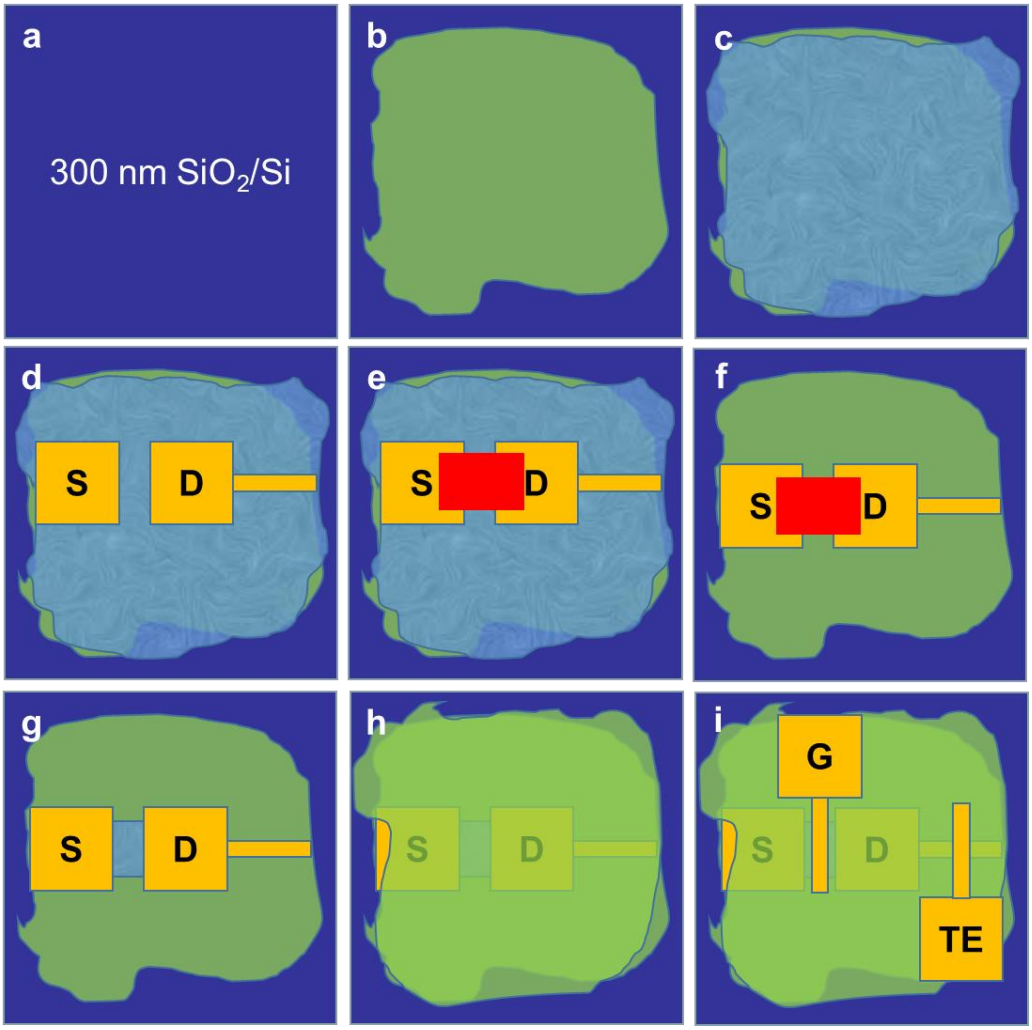


Fig. 5.2 detailed schematic diagram of step-by-step preparation of two-dimensional material based 1t1r framework. Among them, dark blue is SiO2 / Si substrate, green is h-BN, light blue is multilayer graphene, gold is metal electrode, and red is negative photoresist used to pattern fossil graphene channel material.

Figure 5.3a is the physical diagram of the 1t1r storage array architecture prepared by me. In the 1R part of the device, the structure is 50 nm Au / h-BN / 10 nm Ti / 40 nm Au (top); In the 1t part of the device, the channel material I selected is multilayer graphene, the source drain electrode and gate electrode materials are 50 nm Au, and the gate dielectric layer material is also h-BN. In mask design, the minimum linewidth is 3 μ m. In the actual preparation, the linewidth is increased to 4 μ m. This part also needs

further optimization. Figure 5.2b shows the electrical characterization of the transistor part. Firstly, under any gate voltage, the device shows a large current which is linear with the voltage, indicating that the graphene based transistor will not limit the current; In addition, different gate voltages have little effect on the switching of the device, indicating that the graphene based transistor will be able to apply different current limits to achieve the purpose of multi-state resistance.

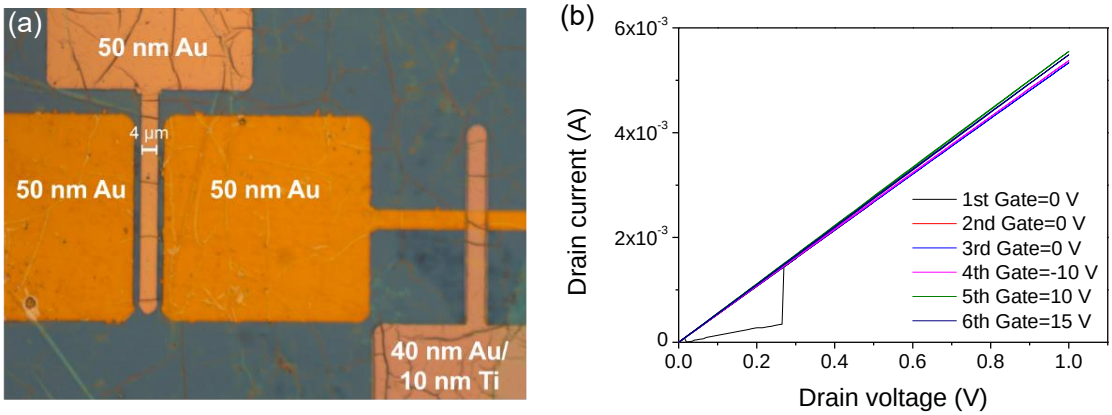


Figure 5.3 physical structure diagram of 1t1r architecture and electrical characterization of 1t part. (a) Optical micrograph of 1t1r device with multilayer graphene as channel material, h-BN as transistor gate dielectric and memristor resistive layer at the same time. The wrinkles indicate that the two-dimensional material has been successfully transferred. (b) Current voltage curve of 1t part between source and drain under different gate voltages.

After the construction of graphene based transistors failed, I used izgo as the channel material and a simpler bottom gate structure. Figure 5.4A shows the structural diagram of 1t part with izgo as channel material. 1t part has current limiting function, and the gate voltage can adjust the current limiting value for multi state storage. Figure 5.4b shows that the switching current of the device can reach three orders of magnitude when the gate electrode changes from 0 V to 10 v. However, the saturation current needs to be further optimized according to the working current range of memristor.

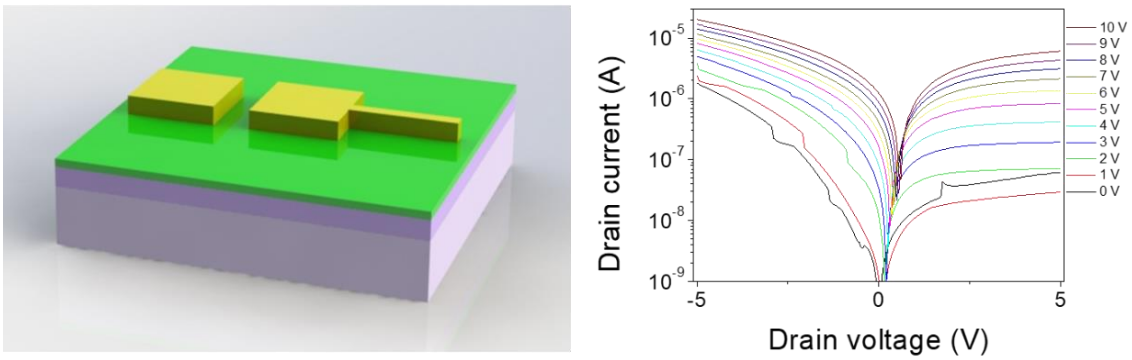


Figure 5.4 (a) bottom gate transistor structure is used for 1t1r. Among them, yellow is 1t part of source drain electrode, green part is channel material igzo, and dark / lavender is SiO2 gate oxide and highly doped silicon gate electrode. (b) Source drain I-V curve under different gate voltages.