University of Sheffield

A Contribution Towards Intelligent Autonomous Sensors Based on Perovskite Solar Cells and Ta_2O_5/ZnO Thin Film Transistors

A thesis submitted for the degree of Doctor of Philosophy by:

Xiaoyao Song

Submitted to the Department of Electronic and Electrical Engineering

The University of Sheffield

DOCTOR OF PHILOSOPHY

June 2023

DECLARATION

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or processional qualification except as specified.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisor, Prof. Maria Merlyne De Souza. Starting from the moment she offered me the opportunity to join her team, her continuous guidance, support, and encouragement, especially during difficult times, have been invaluable. Her constant encouragement and unparalleled expertise have served as the driving forces behind my academic journey, providing me with the motivation and knowledge to overcome numerous challenges.

I would also like to express my sincere thanks to my colleagues, Premlal Pillai, Ashwani Kumar, and Nagaditya Poluri. Their unwavering support and collaborative spirit have significantly enhanced my research experience. Their unique perspectives and thoughtful conversations have substantially contributed to my professional growth, making this journey all the more enjoyable.

I wish to acknowledge the significant roles played by the dedicated staff members Jon Wall, Jonathan D. Milner, Paul Haines, and Stephen Gricmanis. The completion of the experimental work and recovery from the fire accident would not have been possible without their consistent support.

Lastly, and most importantly, I dedicate this work to my parents, Guihua Wang and Youzhong Song. Their unconditional love and support have made my journey possible, and I am forever grateful for their unwavering faith in me. They have taught me the value of hard work and perseverance, lessons that have been fundamental throughout this journey.

To everyone mentioned and those unmentioned who have contributed to this journey, I extend my heartfelt thanks. This work stands as a testament to your invaluable contribution.

ABSTRACT

Many broad applications in the field of robotics, brain-machine interfaces, cognitive computing, image and speech processing and wearables require edge devices with very constrained power and hardware requirements that are challenging to realize. This is because these applications require sub-conscious awareness and require to be always "on", especially when integrated with a sensor node that detects an event in the environment. Present day edge intelligent devices are typically based on hybrid CMOS-memristor arrays that have been so far designed for fast switching, typically in the range of nanoseconds, low energy consumption (typically in nano-Joules), high density and endurance (exceeding 10^{15} cycles). On the other hand, sensory-processing systems that have the same time constants and dynamics as their input signals, are best placed to learn or extract information from them. To meet this requirement, many applications are implemented using external "delay" in the memristor, in a process which enables each synapse to be modeled as a combination of a temporal delay and a spatial weight parameter.

This thesis demonstrates a synaptic thin film transistor capable of inherent logic functions as well as compute-in-memory on similar time scales as biological events. Even beyond a conventional crossbar array architecture, we have relied on new concepts in reservoir computing to demonstrate a delay system reservoir with the highest learning efficiency of 95% reported to date, in comparison to equivalent two terminal memristors, using a single device for the task of image processing. The crux of our findings relied on enhancing our capability to model the unique physics of the device, in the scope of the current thesis, that is not amenable to conventional TCAD simulations. The model provides new insight into the redox characteristics of the gate current and paves way for assessment

of device performance in compute-in-memory applications. The diffusion-based mechanism of the device, effectively enables time constants that have potential in applications such as gesture recognition and detection of cardiac arrythmia.

The thesis also reports a new orientation of a solution processed perovskite solar cell with an efficiency of 14.9% that is easily integrable into an intelligent sensor node. We examine the influence of the growth orientation on film morphology and solar cell efficiency. Collectively, our work aids the development of more energy-efficient, powerful edge-computing sensor systems for upcoming applications of the IOT.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF PUBLICATIONS

Journals:

(1) Gaurav, A.; **Song, X**.; Manhas, S.; Gilra, A.; Vasilaki, E.; Roy, P.; De Souza, M. M. Reservoir Computing for Temporal Data Classification Using a Dynamic Solid Electrolyte ZnO Thin Film Transistor. *Front. Electron.* **2022**, *3*. https://doi.org/10.3389/felec.2022.869013.

(2) Kumar, A.; **Song, X**.; De Souza, M. M. Necessary Conditions for Steep Switching in a Constant Resistor-Capacitor RCFET. *MRS Adv.* **2021**, *6* (21), 540–545. https://doi.org/10.1557/s43580-021-00119-9.

(3) **Song, X**.; Kumar, A.; De Souza, M. M. Off-State Operation of a Three Terminal Ionic FET for Logic-in-Memory. *IEEE J. Electron Devices Soc.* **2019**, *7* (July), 1232–1238. https://doi.org/10.1109/JEDS.2019.2941076. pp 154–156. https://doi.org/10.1109/EDTM.2019.8731277.

(4) Kumar, A.; Balakrishna Pillai, P.; **Song, X**.; De Souza, M. M. Negative Capacitance beyond Ferroelectric Switches. *ACS Appl. Mater. Interfaces* **2018**, *10* (23), 19812–19819. https://doi.org/10.1021/acsami.8b05093.

(5) Balakrishna Pillai, P.; Kumar, A.; **Song, X**.; De Souza, M. M. Diffusion-Controlled Faradaic Charge Storage in High-Performance Solid Electrolyte-Gated Zinc Oxide Thin-Film Transistors. *ACS Appl. Mater. Interfaces* **2018**, *10* (11), 9782–9791. https://doi.org/10.1021/acsami.7b14768.

Conferences:

(1) **Song, X**.; Gaurav, A.; Pillai, P. B.; Kumar, A.; Manhas, S.; Gilra, A.; Vasilaki, E.;

Souza, M. M. De. Reservoir Computing Based on a Solid Electrolyte ZnO TFT : An Attractive Platform for Flexible Edge Computing. *Int. Flex. Electron. Technol. Conf.* **2023**.

(2) Souza, M. M. De; **Song, X**.; Gaurav, A.; Manhas, S. K.; Pillai, P. B.; Sikdar, S.; Kumar, A.; Gilra, A.; Vasilaki, E. A Delay System Reservoir Based on a Nano-Ionic Solid Electrolyte FET *. Nanotechnol. Mater. Devices Conf. 2023.

(3) Gaurav, A.; **Song, X**.; Manhas, S. K.; Roy, P. P.; De Souza, M. M. A Solid Electrolyte ZnO Thin Film Transistor for Classification of Spoken Digits Using Reservoir Computing. In 2023 7th IEEE Electron Devices Technology & Manufacturing Conference (EDTM); IEEE, 2023; pp 1–3. https://doi.org/10.1109/EDTM55494.2023.10103131.

(4) **Song, X**.; Kumar, A.; De Souza, M. M. An Ultra-Low Power 3-Terminal Memory Device with Write Capability in the off-State. In 2019 Electron Devices Technology and Manufacturing Conference (EDTM); IEEE, 2019; pp 154–156. https://doi.org/10.1109/EDTM.2019.8731277.

(5) **Song, X**.; Pillai, P. B.; Batten, T.; Souza, M. M. De. Highly Oriented (202) Mixed Halide Perovskite for Enhanced Solar Cell Performance. In UK semi 2017 conference; 2017.

CHAPTER 1

Introduction

1.1 Motivation

The escalating demand in global energy presents a formidable challenge to our planet, primarily due to its significant contribution to the alarming increase in greenhouse gas emissions. This situation is further complicated by the ambitious objective of achieving net-zero emissions by 2050. By 2021, the worldwide demand for electricity surged to 24,700 TWh, representing a 6% year-on-year increase from 2020, which was the largest annual increase seen since 2010 [1]. Approximately 46% of this demand, which accounts for a significant proportion of the global increase in CO2 emissions, is attributable to increasing demand in electricity and heat production [2].

Among various sectors contributing to this relentless demand for electricity, technology-oriented sectors, particularly those reliant on heavy data processing, have experienced the steepest growth. For instance, cryptocurrency mining alone accounted for an energy consumption of 100-140 TWh in 2021, marking an astonishing increase of 2000-3000% from 2015 [3].

In the modern landscape, a burgeoning technology known as the Internet of Things (IoT), also referred to as the Industrial Internet, is gaining traction. This technology fosters interaction among global networks, paving way for the emergence of intelligent environments capable of autonomous self-regulation. These networks link sensors and communication devices on a global scale. Currently, IoT implementations necessitate data analysis in the cloud, where data is transferred from the point of generation and subsequently processed for decision-making. Furthermore, it falls short of achieving the

envisioned fully integrated smart functionality of IoT due to its high power consumption, estimated to be between 320 and 460 TWh for data centers alone [4]. It is projected that by 2025, the quantity of IoT devices will proliferate, reaching an astounding figure of 4.7 billion [5]. A substantial portion of the data, generated by these devices, is characterized by its unstructured nature and redundancy. This leads to heightened energy consumption and latency, particularly due to the transmission of this data for processing [6].

The cloud-centric model of the IoT, based on conventional Von Neumann computing architecture that distinctly separates memory and computational elements, is unsustainable (Fig 1.1), due to data transfer bottlenecks and energy inefficiency [7].

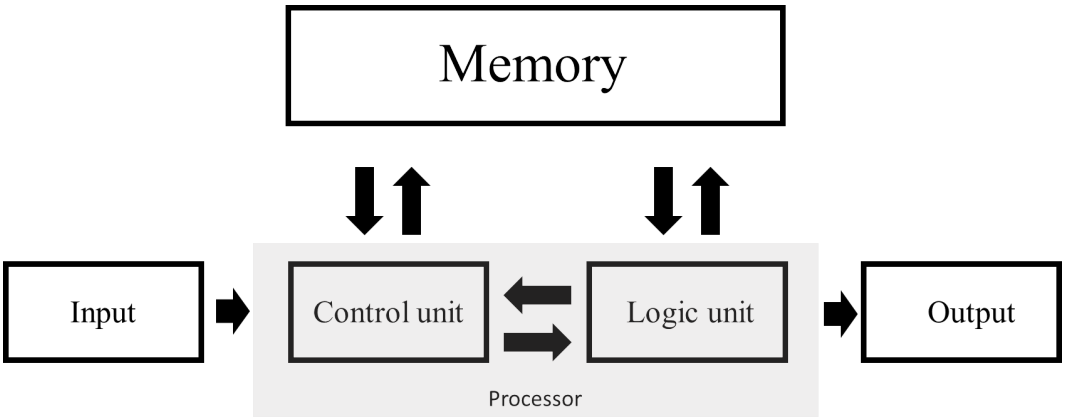


Fig 1.1 Schematic diagram of Von Neumann architectures

Despite the strength of CMOS in handling mathematical and logic operations, these devices fail to match the efficiency of human neurons and synapses in processing analog information. Furthermore, Von Neumann architectures inherently lack features that could support learning and problem-solving. For example, despite not having the raw computational power found in modern machines, the human brain excels at solving sensory-related problems through the use of memory and learning. As a result, software based neural networks require tens of MW of power to mimic only a fraction of human

brain function [8], such as identifying images, that too, after repeated training with substantial data. The human brain accomplishes similar image identification tasks using a mere 20 watts of energy, with less training and data [9]. Therefore, developing hardware akin to brain-inspired computing systems that can more accurately mimic biological neurons and synapses is viewed as a significant step towards achieving energy efficiency [10].

Alternatives to the Von Neumann model such as Compute-in-Memory (CIM) and Neuromorphic Computing are being actively explored, which promise improved computational speed and energy efficiency integrating processing and memory functions into the same unit, thereby avoiding speed and energy losses due to data transmission between separate processing and memory units [11].

The implementation of CIM architectures necessitates hardware that offers high speed, high density, and low energy cost [11][12]. Typically, this involves the integration of scalable memory devices with CMOS devices in a crossbar array for massively parallel operations [12], which are inherently compatible with multiplication and addition operations (Fig 1.2).

Memristors are typically used as synapses in such artificial neural networks (ANNs), where they store conductance values as weights. These weights determine the connections between layers and are a key part of the network's learning functions. The sum of currents in the crossbar are fed into neurons which undertake the task of thresholding. Implementing a neuron in large scale crossbars to achieve the activation function or learnable kernels is challenging and expensive, often relying on field-programmable gate

arrays (FPGA) [13], as well as Analogue to Digital (ADC) and Digital to Analogue Converters (DAC) [14][15] to achieve the desired network function.

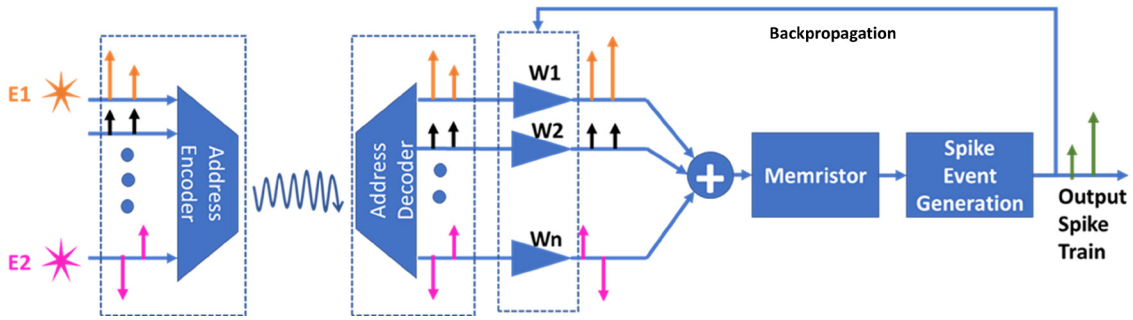


Fig 1.2 Sensor connected IoT application

Deep neural networks (DNNs) based on memristors devices such as CBRAMs [16], PCMs [17][18], and OxRAMs[19][20] are increasingly popular due to their compatibility with CMOS, scalability, bidirectional switching behavior, and reduced power consumption by three orders of magnitude compared to a CPU with off-chip memory. However, the change in conductance in a ReRAM, which is caused by the formation and rupture of a conducting filament, is inherently analog and stochastic. This leads to variability, relaxation, nonlinear conductance behavior, and a reduction in yield. This approach necessitates additional efforts at the device, architecture, or algorithmic levels to address the loss of accuracy and make it unsuitable for high precision computing. The accuracy of memristor-based computing can be improved by off-chip training methods, which are energy-intensive and occur in the cloud. However, discrepancies between the platforms used in training (usually GPUs) and the crossbar implementation of the ReRAM lead to a loss of learning efficiency. In contrast, on-chip training methods can eliminate this issue, especially as secure communication with the cloud is a critical constraint in health applications. Examples of embedded intelligence in alternate technologies (beyond

CMOS) are virtually nonexistent in the commercial world, primarily due to challenges associated with integration, especially when substrate platforms are not silicon-based, as is common in sensor technologies, which may rely on alternate materials that need to meet the constraints of their applications.

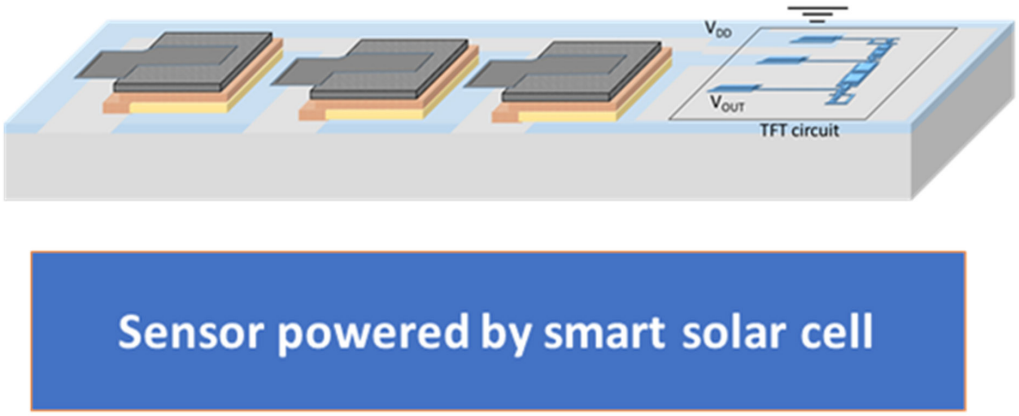


Fig 1.3 Schematic diagram of of a integrated smart solar cell

Edge computing (also known as Processing-in-sensor) allows decisions to be made at the source, thereby reducing the need for extensive data storage and transmission to the cloud. This approach emphasizes the processing of information over the transmission of raw sensor data. This effectively reduces the volume of data required for transmission, thereby decreasing energy usage and latency [21]. The ultimate goal is to develop smart IoT devices capable of operation purely on harvested energy and preserving their computational state even when energy sources are depleted. Furthermore, large-scale implementation of such systems requires commitment to low-cost, environmentally sustainable materials. Moreover, the implementation of such a solution necessitates a meticulous balancing act, harmonizing the demands of space, power supply, computational power, and memory within the sensor framework.

In particular, a major challenge faced by autonomous wireless systems is power supply. When power sources such as batteries or capacitors are combined with renewable energy sources, wireless sensor nodes can operate for longer durations, potentially eliminating the need for battery replacement [22]. This leads to an extended device lifespan and a reduction in weight. There are several examples of energy harvesting integration with CMOS transistors [15], such as photovoltaic energy harvesters integrated with power management ICs in CMOS, which have been demonstrated for at least two decades [23]. These are embedded in watches and other wearable sensors today. For instance, a-Si-H solar cells on top of CMOS transistors [15] and a self-powered solar switch fabricated by integrating a crystalline-Si interdigitated back contact (IBC) PV cell with a-IGZO TFTs in a direct 3D stacked single chip have been demonstrated [24].

More recently, a perovskite photovoltaic cell integrated with a 2D MoTe2 transistor was demonstrated [25], in which double-sided scotch tape was used for bonding to integrate these two technologies at the package level. Since 2009, perovskites have emerged as promising photovoltaic materials due to their potential for high efficiency and low cost. Over the past decade, the efficiency of perovskite solar cells has increased from 3.8% [26] to 25.7% [27], making them an emerging candidate for the next generation of solar cells. Furthermore, the hysteresis observed in perovskite solar cells has led to the development of perovskite-based memristors [28], which has sparked interest in their use in neuromorphic systems [29][30].

This thesis contributes to this field, specifically focusing on neuromorphic devices that can be seamlessly integrated with solar cells for energy harvesting and flexible electronics. We aim to explore materials and processes beyond CMOS that are earth-

abundant, can be processed at low temperatures, and are flexible. These innovative devices hold potential for meeting the rising demands of the healthcare sector, particularly in applications such as assisted living and robotics. Technologies capable of real-time contextualization and filtering of sensitive and safety-critical data at the point of sensing can empower artificial intelligence-based systems to proactively provide support, guidance, and intervention in such applications. Furthermore, these devices could be powered by, or even provide control, to the solar cell, leading to the concept of the 'smart' energy harvesting.

1.2 Outline of the Thesis

This thesis is built upon two distinct types of materials that constitute the components for the autonomous intelligent sensor node. Chapters 2 and 3 serve as foundational background chapters, containing an exploration of the physics and a literature review of devices relevant to this work. Chapters 4 to 7 are the primary contributory chapters. The organization of this thesis is as follows:

Chapter 2 provides an introduction to the fundamentals of perovskite materials and their use in solar cells, detailing the fabrication and characterization methods employed. The chapter also outlines the basics of amorphous oxide semiconductors and the operating principles of thin-film transistors based on these materials.

Chapter 3 delves into the principles of brain-inspired computing, tracing the evolution of artificial neural networks from feed-forward neural networks (FNNs) to recurrent neural networks (RNNs). This progression includes an overview of the fundamental structures and training processes associated with these networks. The chapter reviews the progression from RNNs to reservoir computing, offering a critique of recent

advancements in physical reservoir computing. It also discusses the key features of the physical synaptic devices and the emerging memory devices used in neuromorphic systems.

Chapter 4 provides a thorough background on a novel non-filamentary thin-film transistor, known as the Solid Electrolyte Field Effect Transistor (SE-FET). This chapter highlights the device performance in synaptic applications, including its Excitatory Post Synaptic Current (EPSC). Moreover, we present a refined model of the SE-FET that facilitates the simulation of the gate current, a crucial aspect for circuit simulation which in Chapter 5. Furthermore, we delve into the physical interpretation of the gate current and discuss the limitations inherent in previous models [31].

Chapter 5 revolves around the deployment of Solid Electrolyte field-effect transistors (SE-FETs) in logic operations and supervised learning within crossbar architectures. We illustrate how the three-terminal nature of these devices facilitates write operations at a remarkably low power consumption level.

Chapter 6 revolves around the basics of physical reservoirs computing and explores the potential of SE-FETs as physical reservoirs in recurrent neural networks (RNNs).

Chapter 7, we demonstrate a new method for controlled perovskite growth, followed by an evaluation of the physical characteristics and performance of perovskite solar cells produced through our method and conventional techniques.

Finally, Chapter 8 presents the concluding remarks, summarizing the findings and contributions of this thesis. We also discuss potential future directions to further extend the scope of this work.

1.3 References

[1] International Energy Agency, “International Energy Agency (IEA) World Energy Outlook 2022,” *Https://Www.Iea.Org/Reports/World-Energy-Outlook-2022/Executive-Summary*, p. 524, 2022.

[2] I. E. Agency, “Global Energy Review : CO2 Emissions in 2021 Global emissions rebound sharply to highest ever level INTERNATIONAL ENERGY,” 2021.

[3] IEA, “Data Centres and Data Transmission Networks,” 2022. .

[4] E. Covi *et al.*, “Adaptive Extreme Edge Computing for Wearable Devices,” *Front. Neurosci.*, vol. 15, no. May, pp. 1–27, May 2021.

[5] L. S. Vailshery, “IoT connected devices globally in 2019 & 2030.” [Online]. Available: https://www.statista.com/markets/418/technology-telecommunications/

[6] F. Zhou and Y. Chai, “Near-sensor and in-sensor computing,” *Nat. Electron.*, vol. 3, no. 11, pp. 664–671, Nov. 2020.

[7] D. V Christensen *et al.*, “2022 roadmap on neuromorphic computing and engineering,” *Neuromorphic Comput. Eng.*, vol. 2, no. 2, p. 022501, Jun. 2022.

[8] S. Dai *et al.*, “Recent Advances in Transistor-Based Artificial Synapses,” *Adv. Funct. Mater.*, vol. 29, no. 42, p. 1903700, Oct. 2019.

[9] R. Ananthanarayanan *et al.*, “The cat is out of the bag: cortical simulations with 109 neurons, 1013 synapses,” *High Perform. Comput. Networking, Storage Anal. Proc. Conf.*, no. c, pp. 1–12, 2009.

[10] Y. Zhang *et al.*, “Brain-inspired computing with memristors: Challenges in devices, circuits, and systems,” *Appl. Phys. Rev.*, vol. 7, no. 1, p. 011308, Mar. 2020.

[11] S. Yu *et al.*, “Compute-in-Memory Chips for Deep Learning: Recent Trends and

Prospects," *IEEE Circuits Syst. Mag.*, vol. 21, no. 3, pp. 31–56, 2021.

[12] D. Ielmini and H. S. P. Wong, "In-memory computing with resistive switching devices," *Nat. Electron.*, vol. 1, no. 6, pp. 333–343, 2018.

[13] R. Fernandez Molanes *et al.*, "Deep learning and reconfigurable platforms in the internet of things: Challenges and opportunities in algorithms and hardware," *IEEE Ind. Electron. Mag.*, vol. 12, no. 2, pp. 36–49, 2018.

[14] S. Mittal, "A Survey of ReRAM-Based Architectures for Processing-In-Memory and Neural Networks," *Mach. Learn. Knowl. Extr.*, vol. 1, no. 1, pp. 75–114, 2018.

[15] T. Zhang *et al.*, "Memristive Devices and Networks for Brain-Inspired Computing," *Phys. status solidi – Rapid Res. Lett.*, vol. 13, no. 8, p. 1900029, Aug. 2019.

[16] J.-H. Cha *et al.*, "Conductive-bridging random-access memories for emerging neuromorphic computing," *Nanoscale*, vol. 12, no. 27, pp. 14339–14368, 2020.

[17] V. Joshi *et al.*, "Accurate deep neural network inference using computational phase-change memory," *Nat. Commun.*, vol. 11, no. 1, p. 2473, May 2020.

[18] G. Pedretti *et al.*, "A Spiking Recurrent Neural Network With Phase-Change Memory Neurons and Synapses for the Accelerated Solution of Constraint Satisfaction Problems," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 6, no. 1, pp. 89–97, Jun. 2020.

[19] S. Kim *et al.*, "Metal-oxide based, CMOS-compatible ECRAM for Deep Learning Accelerator," in *2019 IEEE International Electron Devices Meeting (IEDM)*, 2019, pp. 35.7.1-35.7.4.

[20] Y. Zhang *et al.*, "Oxide-based filamentary RRAM for deep learning," *J. Phys. D. Appl. Phys.*, vol. 54, no. 8, p. 083002, Feb. 2021.

[21] F. De Roose *et al.*, “12.3 Memory Solutions for Flexible Thin-Film Logic: up to 8kb, >105.9kb/s LPROM and SRAM with Integrated Timing Generation Meeting the ISO NFC Standard,” in *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, 2019, vol. 31, pp. 206–208.

[22] D. Prakash Guragain *et al.*, “Programmable timer triggered energy harvesting wireless sensor-node using long range radio access technology,” *Int. J. Electr. Comput. Eng.*, vol. 12, no. 4, p. 3869, Aug. 2022.

[23] J. Lu *et al.*, “Integration of Solar Cells on Top of CMOS Chips Part I: a-Si Solar Cells,” *IEEE Trans. Electron Devices*, vol. 58, no. 7, pp. 2014–2021, Jul. 2011.

[24] S. Ali *et al.*, “Solar Powered Smart Irrigation System,” *Pakistan J. Eng. Technol.*, vol. 5, no. 1, pp. 49–55, 2022.

[25] Y.-M. Juan *et al.*, “Integration of a-IGZO Thin-Film Transistor and Crystalline-Si Interdigitated Back Contact Photovoltaic Cell With 3D Stacking Structure as Self-Powered Solar Switch,” *IEEE Electron Device Lett.*, vol. 35, no. 10, pp. 1040–1042, Oct. 2014.

[26] A. Kojima *et al.*, “Organometal halide perovskites as visible-light sensitizers for photovoltaic cells,” *J. Am. Chem. Soc.*, vol. 131, no. 17, pp. 6050–6051, 2009.

[27] M. A. Green *et al.*, “Solar cell efficiency tables (Version 61),” *Prog. Photovoltaics Res. Appl.*, vol. 31, no. 1, pp. 3–16, Jan. 2023.

[28] C. Gu and J.-S. Lee, “Flexible Hybrid Organic–Inorganic Perovskite Memory,” *ACS Nano*, vol. 10, no. 5, pp. 5413–5418, May 2016.

[29] J. Choi *et al.*, “Organolead Halide Perovskites for Low Operating Voltage Multilevel Resistive Switching,” *Adv. Mater.*, vol. 28, no. 31, pp. 6562–6567, Aug.

2016.

[30] B. Hwang *et al.*, “Effect of halide-mixing on the switching behaviors of organic-inorganic hybrid perovskite memory,” *Sci. Rep.*, vol. 7, no. 1, p. 43794, Mar. 2017.

[31] A. Kumar *et al.*, “Negative Capacitance beyond Ferroelectric Switches,” *ACS Appl. Mater. Interfaces*, vol. 10, no. 23, pp. 19812–19819, Jun. 2018.

CHAPTER 2

Basics of Perovskite solar cells and Amorphous Oxide Thin Film Transistors

2.1 Introduction

Perovskite materials and amorphous oxide semiconductors (AOS) have been subjects of increasing attention, motivated primarily for their potential low cost and broad applications in energy harvesting and flexible electronics. Perovskites, with their distinct properties, are viewed as a promising catalyst for significant strides in solar energy conversion, while AOS, due to their exceptional electronic characteristics, hold promise in diverse areas, including display technology, neuromorphic computing, and renewable energy (as carrier transporting layer).

This chapter, therefore, offers an exhaustive background to these vital materials, starting in part 2.2 with a brief history of perovskite solar cells, followed by a comprehensive examination of their fundamental material properties, detailing methodologies employed in their fabrication and characterization and principle of operation of solar cells. Section 2.3 is devoted to Amorphous oxide semiconductors. Subsection 2.3.1 highlights applications of AOSs in the areas particularly of displays and memory. In particular, recent development of using AOS, especially ZnO, as electron transporting layer to address cost and stability in perovskite solar cells is highlighted. This is followed by the structure and properties, with a special emphasis on distinguishing the material differences between AOS and their amorphous silicon counterparts. Subsection 2.3.3 highlights methods of deposition and is followed by the principle of operation of thin film transistors in AOS.

2.2 Perovskite solar cells

2.2.1 A brief history of perovskite solar cells

The history of perovskite-based solar cells is relatively brief, spanning just over a decade. The inaugural perovskite-based solar cell was unveiled by A. Kojima and colleagues in 2006. They employed $CH_3NH_3PbBr_3$ as the light-absorption material in a dye-sensitized solar cell (DSSC), achieving a power conversion efficiency (PCE) of 2.2% [1]. The first report specifically centered on perovskite-based solar cells was released by the same team in 2009, where $CH_3NH_3PbBr_3$ was substituted with the now widely used $CH_3NH_3PbI_3$, leading to an improved PCE of 3.83% [2].

A landmark development for perovskite solar cells transpired in 2012, with the enhancement of an all-solid-state structure premised on TiO2 [3]. This pioneering work propelled the perovskite solar cell away from the electrolyte-based DSSC, resulting in the achievement of a PCE exceeding 10% for the first time [4].

In 2013, a notable enhancement was reported that involved doping the film with Cl. It was conjectured that the creation of $CH_3NH_3PbI_{3-x}Cl_x$ led to improved carrier diffusion lengths (exceeding 1 μm for both electrons and holes) [5]. In the same year, N. Pellet and associates put forth an improved two-step solution technique that elevated the PCE of perovskite solar cells to 15.0% [6]. This group also introduced the usage of $HN{=}CHNH_3^+$ (FA^+) as a substitute cation for $CH_3NH_3^+$(MA^+). The combined perovskite using both FA^+ and MA^+ exhibited a promising enhancement in the device's short-circuit photocurrent and elevated the PCE to 14.9% [7]. A significant efficiency record of 19.3% was achieved in 2014 by H. Zhou and colleagues by optimizing the interfaces between the perovskite, the electron transport material (ETM), and the hole transport layer (HTL), and

aligning the band structure [8][9]. In 2016, a novel vacuum flash-assisted solution processing method was developed for large-area perovskite solar cells, significantly boosting their certified power conversion efficiency of perovskite solar cells over 1 square centimeter from 15.6% to 19.6% [10].

In 2017, a new technique known as defect passivation was introduced, aimed at enhancing the performance and addressing the stability issues of perovskite solar cells. This method specifically targets and mitigates the imperfections in the perovskite material, leading to improved efficiency and durability of these solar cells [11]. Defect passivation targets the amelioration of surface defect sites at the perovskite layer interface by incorporating a passivation layer between the perovskite and the ETM or hole transport material (HTM), or in some instances, both. Notably, in 2019, Q. Jiang et al. utilized phenethyl ammonium iodide as a passivation layer between the perovskite and the HTM, which led to the achievement of a certified efficiency of 23.32%. Remarkably, the device retained this efficiency for over 500 hours [12]. In 2023, a new record of 25.7% was set for thin film perovskite solar cell [13]. Moreover, using a mixed of different perovskite materials, the Multijunction perovskite/perovskite and perovskite/Si is on the rise. In 2022, Haowen Luo et al. reported a perovskite/perovskite (two hybrid perovskite layer with different band gap) based Multijunction solar with recorded efficiency of 28% [13], and the device is able to maintain 75% of its efficiency after 500 hours of operation [14].

2.2.2 Basic structure and property of perovskite material

Perovskites represent a versatile family of materials unified by their ABX_3 structure, as illustrated in Fig 2.1. In this configuration, 'A' denotes a large positively charged compound, 'B' signifies a smaller positively charged ion, and 'X' is a negatively

charged ion, typically arranged in a cubic crystal. B ions inhabit the body center of the cubic cell, while A ions reside at each vertex, and X ions are located at each face-center of the unit cell. This ideal cubic structure can accommodate a broad range of metal oxides, such as $CaTiO_3$, $LaAlO_3$, or $MgSiO_3$. Here, A, B, and X ions possess charge states of A^{2+}, B^{4+}, and X^{2-} or A^{3+}, B^{3+}, and X^{-2}. Intriguingly, organic compounds can also form structures with metal trihalides, where A, B, and X ions are A^{+}, B^{2+}, and X^{-}, respectively [15].

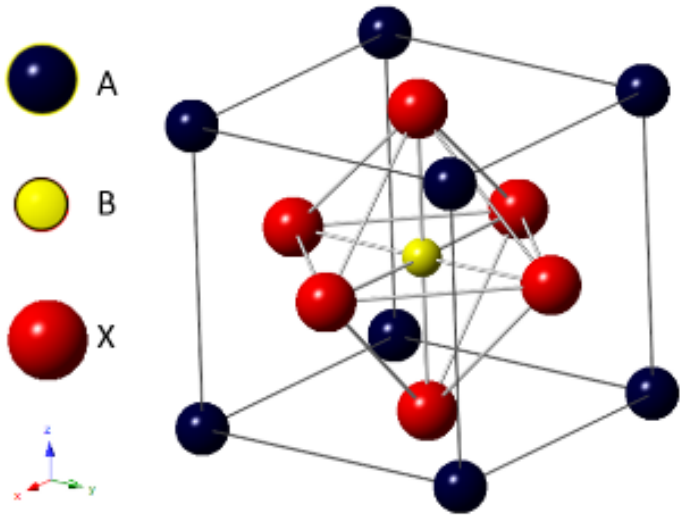


Fig 2.1 Ideal perovskite crystal structure (cubic).

Despite the underlying structural consistency of perovskites, distortions frequently occur depending on factors such as formation temperature and the ionic radii of A, B, and X [16]. Depending on the formation temperature and the ionic radii of the respective ions, the same perovskite group's structure can transition from cubic to tetragonal or even orthorhombic. Such structural shifts are often quantified using a tolerance factor (t) defined as:

$$t = \frac{r_A + r_X}{\sqrt{2}(r_B + r_X)} \qquad \text{(Eq 2.1)}$$

Here r_A , r_B, and r_X are the ionic radii of A, B, and X, respectively. The tolerance factor 't' provides an indication of hexagonal perovskite instability [17]. A tolerance factor exceeding 1 suggests challenges in maintaining the perovskite structure, whereas a tolerance factor within the range of $0.9 \leq t \leq 1$ is indicative of an optimal cubic perovskite structure. Should the tolerance factor fall between $0.7 \leq t \leq 0.9$, it is usually a sign that 'A' is undersized for the cubic structure, consequently yielding orthorhombic, rhombohedral, or tetragonal structures [17]. A tolerance factor of $t < 0.7$ implies that 'A' is equivalent to or smaller than 'B', a scenario rarely encountered in perovskite structures.

In the realm of perovskite-based solar cells, the term 'perovskite' generally refers to the organic–inorganic metal trihalide, with $CH_3NH_3PbI_3$ being the most extensively utilized perovskite-based material. Here, 'A' is the large organic ion $CH_3NH_3^+$(commonly referred to as MA^+), 'B' is typically a Pb^{2+} ion, and 'X' is I. At room temperature, the ionic radii of MA+, Pb2+, and I- are 0.18, 0.132, and 0.206 nm, respectively, yielding a tolerance factor of 0.81 [18]. This value suggests a tetragonal structure for $MAPbI_3$, corroborating experimental findings [19].

The tetragonal structure of the MAPbI3 crystal, delineated in Fig 2.2, has lattice constants a = b = 8.85 Å, c = 12.69 Å, and angles $\alpha = \beta = \gamma = 90°$ [20][21]. The Pb-I framework constitutes the fundamental structure of the $MAPbI_3$ unit cell, while $CH_3NH_3^+$ molecules are situated within the Pb-I framework. Due to these molecules, the Pb-I framework often exhibits distortion, leading to a less symmetrical structure with each unit cell containing four units of $CH_3NH_3PbI_3$ [21].

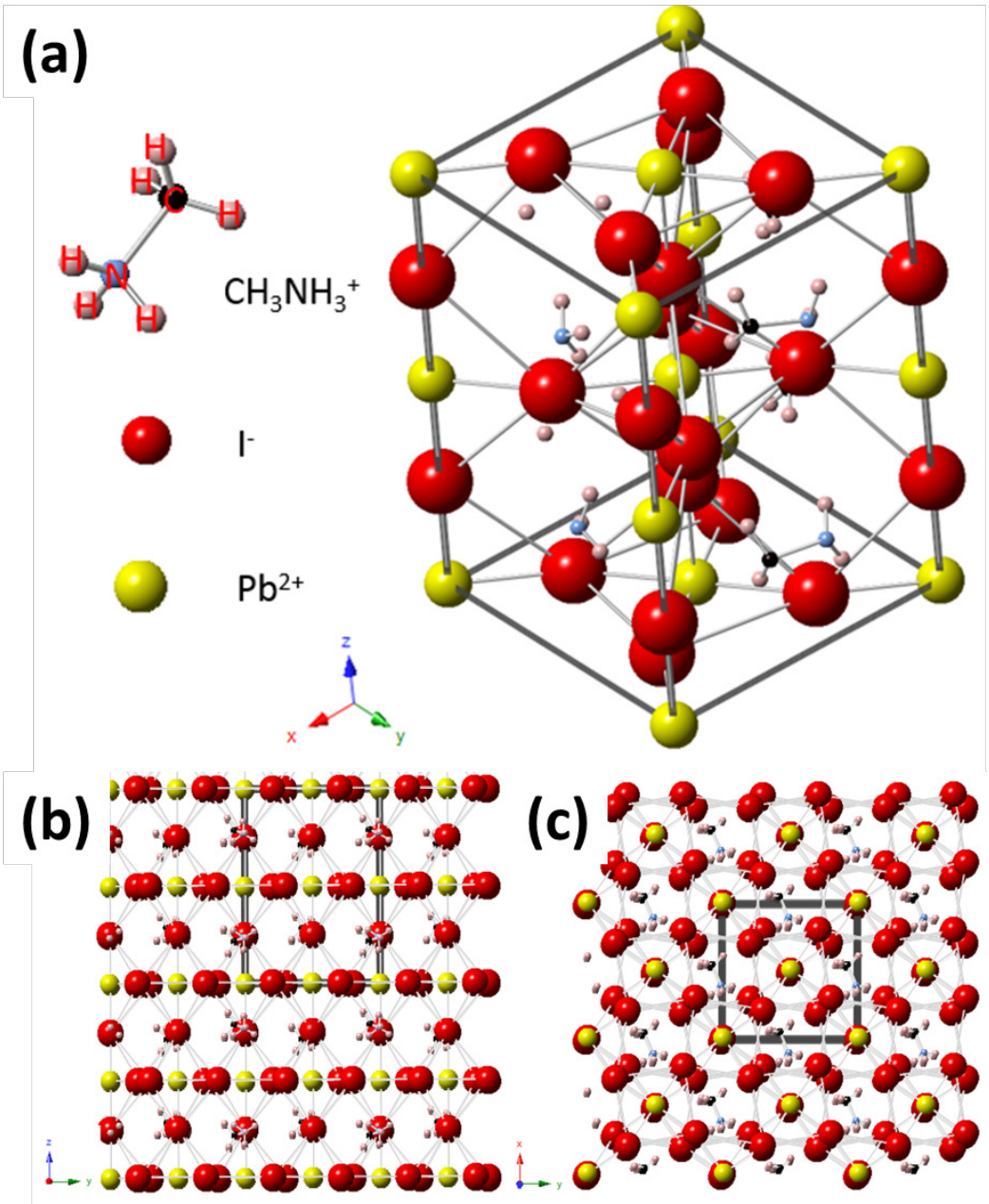


Fig 2.2 (a) Unit cell of the tetragonal structure of $CH_3NH_3PbI_3$. Here, A is $CH_3NH_3^+$, B is Pb^{2+}, and X is I^-. (b) Side view and (c) top view of the tetragonal $CH_3NH_3PbI_3$ structure.

$CH_3NH_3PbI_3$ demonstrates a direct band gap of 1.55~1.6 eV [22][23], in close proximity to the ideal optical band gap for a single-junction solar cell of approximately

1.3~1.4 eV [24][25]. This halide perovskite's band gap is adjustable through mixed-halide superlattices or by altering the halide composition, typically by substituting I with Cl or Br (1.95eV for $CH_3NH_3PbBr_3$ and 2.46 eV for $CH_3NH_3PbCl_3$) [26]. Moreover, polycrystalline $CH_3NH_3PbBr_3$ exhibits an impressively sharp absorption coefficient, exceeding that of Si and aligning closely with the band gaps of GaAs and CdTe [27]. This property permits high absorption even for thin perovskite films, conferring an advantage in thin-film devices over traditional Si materials. The majority of perovskite solar cells are less than 1 μm thick and achieve efficiencies nearing 20%, significantly surpassing the peak reported efficiency for thin-film Si solar cells (less than 2 μm thick) at 10.5% [28].

Single-crystal $CH_3NH_3PbI_3$ has demonstrated a carrier mobility of 30 cm^2/Vs, with the trap state density reaching as low as 3.6×10^{10} cm^3, comparable to that of Si [29]–[32]. Despite these exceptional properties, the factors contributing to the superior photovoltaic performance and their relation to the perovskite structure are yet to be fully understood [33]. Based on the ferroelectric properties of $CH_3NH_3PbI_3$ perovskite [34] and its slow photocurrent response [33], the $CH_3NH_3PbI_3$ structure is thought to constitute a well-defined Pb-I framework with a rotational dynamic cation, $CH_3NH_3^+$. Positioned in the Pb-I framework, as shown in Fig 2.2, the $CH_3NH_3^+$ molecule is comprised of a neutral NH3 group, with the charge mainly distributed to the CH_3^+ group [35]. This uneven charge distribution within the Pb-I framework influences numerous properties relevant to solar cell performance, and is considered to be one of the causes for the ferroelectric properties and hysteresis observed in these cells [33][36]. The $CH_3NH_3^+$ orientation is postulated to be influenced by illumination [33], applied bias [37], and environmental temperature [38],

but the precise rotation mechanism under operational conditions remains a topic of active investigation.

2.2.3 Deposition method for $CH_3NH_3PbI_3$

Since the introduction of perovskite solar cells in 2009, a multitude of deposition methods have been established for growing $CH_3NH_3PbI_3$ perovskite. Despite the diverse array of methods, all are centered around a simple reaction:

$$CH_3NH_3I + PbI_2 = CH_3NH_3PbI_3$$

This reaction typically takes place at temperatures exceeding 60°C [39] and can be prompted by a reaction between solutions of CH_3NH_3I and PbI_2 [40], annealing a dried film of CH_3NH_3I and PbI_2 [41], or through a vapor deposition method [19].

The solution-based process is generally the more popular technique for $CH_3NH_3PbI_3$-based perovskite deposition. Here, the manufacturing process involves coating solutions of CH_3NH_3I and PbI_2, layer by layer, in the shape of the desired device structure. As this technique doesn't necessitate a vacuum or high-temperature conditions, fabrication costs are kept to a minimum.

The one-step and two-step solution-based methods are illustrated in Fig 2.3. In the one-step method, a blend of CH_3NH_3I and PbI_2 in dimethylformamide (DMF) or dimethyl sulfoxide (DMSO) solution is spin-coated onto the chosen substrate. An annealing process usually follows the spin-coating process to dry the remaining solvent and facilitate the formation of the perovskite layer. Due to the high reaction rate, the one-step method often resulted in a non-uniform film with randomly oriented perovskite in earlier works [42][43]. However, the introduction of solvent engineering [44] techniques and other additives, such

as Triazolium [45] and ammonium benzenesulfonate [46], has significantly enhanced the quality and uniformity of perovskite films.

The two-step method splits the perovskite formation process into two phases. The first step involves depositing a layer of PbI_2 via spin coating. This PbI_2 film is then converted into perovskite through various methods, such as spin coating of an MAI/isopropyl alcohol (IPA) solution [42][40], immersion in a solution of MAI/IPA [47][48], or annealing with a solid MAI film [49]. On the whole, the two-step method is viewed as superior, as it typically yields a better average efficiency and morphology [43].

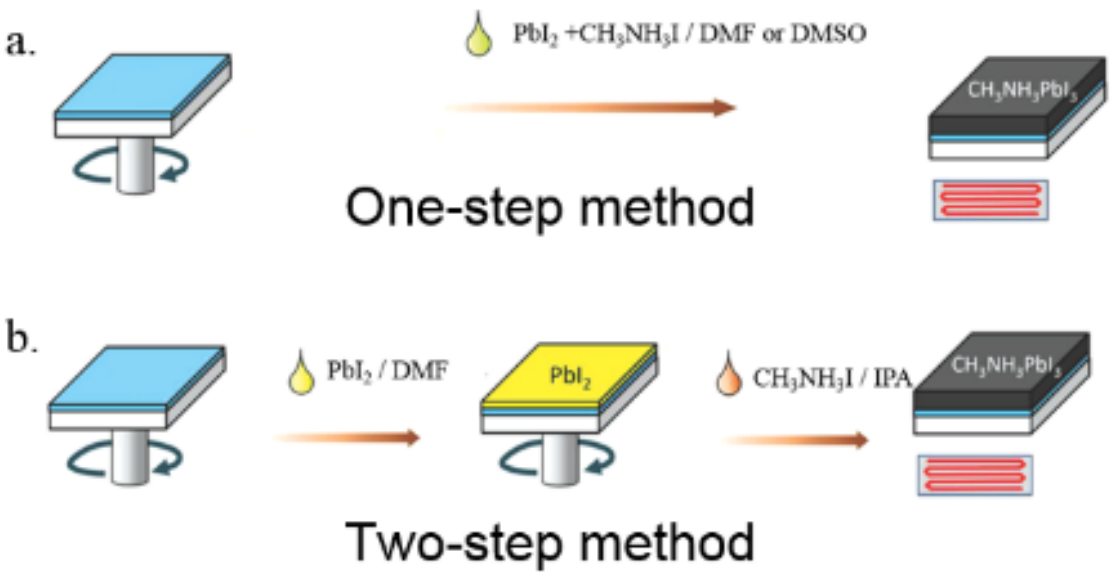


Fig 2.3 One-step and two-step solution-based methods for $CH_3NH_3PbI_3$ deposition.

In 2013, Stranks et al. [5] suggested doping $CH_3NH_3PbI_3$ with Cl, a step that has since been widely implemented for both one-step and two-step methods. In the solution method, doping is accomplished by dissolving a small amount of CH_3NH_3Cl in a CH_3NH_3I solution (or $PbCl_2$ in a PbI_2 solution). It was assumed that Cl would diffuse into the

$CH_3NH_3PbI_3$ and form a mixed $CH_3NH_3PbI_3$-xClx perovskite [5][50]. $CH_3NH_3PbI_3$-xClx is thought to have a superior carrier lifetime compared to the CH3NH3PbI3 perovskite, with reports showing a device efficiency improvement from 4.2% to 12.2% [5].

Despite the widespread acceptance of the Cl doping method for performance enhancement, the role of Cl doping was ambiguous, as Cl loss is commonly observed during perovskite growth [51]. Utilizing XRD and X-ray photoelectron spectroscopy (XPS) techniques, H. Yu et al. detected no Cl in a sample prepared with 10% Cl doping in an MAI+MACl solution [52]. These results confirmed that while Cl doping is an effective technique, the formation of $CH_3NH_3PbI_3$-xClx is clearly not the cause of the improvement. Given the observed Cl loss and the final reaction product, it is currently believed that Cl enhances perovskite quality by forming an intermediate state during the reaction, rather than by forming the CH3NH3PbI3-xClx perovskite. Yu et al. [48] and Williams et al. [53] proposed a possible reaction of the Cl doping technique as follows:

$$PbCl_2 + 3CH_3NH_3I \xrightarrow{\Delta} PbI_2 + CH_3NH_3I + 2CH_3NH_3Cl$$

$$PbI_2 + xCH_3NH_3I + yCH_3NH_3Cl \xrightarrow{\Delta} (CH_3NH_3)_{x+y}PbI_{2+x}Cl_y$$

$$(CH_3NH_3)_{x+y}PbI_{2+x}Cl_y \xrightarrow{\Delta} CH_3NH_3PbI_3 + CH_3NH_3Cl \uparrow$$

2.2.4 Perovskite solar cells: device structure and working principle

Early perovskite solar cells followed a structure akin to Dye-Sensitized Solar Cells (DSSCs); the cell consisted of a photoelectrode with a perovskite layer deposited on TiO2 and Pt as a counter electrode. The two electrodes were separated by an organic electrolyte solution [2]. However, it was discovered that the liquid electrolyte could dissolve the

perovskite layer, resulting in limited performance and stability of perovskite-based DSSCs [54].

Since 2012, all-solid-state structures have replaced liquid-electrolyte perovskite solar cells. There are generally three types of solid-state structures: the conventional p-i-n structure, the mesoscopic p-i-n structure, and the inverted n-i-p structure. Both the conventional and mesoscopic p-i-n structures (Fig 2.4 (a) and (b)) are usually bottom-up structures based on a transparent bottom contact such as fluorine-doped tin oxide (FTO). The electron transport layer (ETL) is typically a compact layer composed of a transparent n-type semiconductor, such as TiO_2 [55] or SiO_2 [56]. In the conventional p-i-n structure, the perovskite layer is deposited on top of the ETL, while in the mesoscopic structure, the perovskite is partly formed within the mesoscopic ETL. Both conventional and mesoscopic p-i-n structures usually require a hole transport layer (HTL) to form the heterojunction. Spiro-MeOTAD is the most commonly used material for the HTL, and the top contact is typically gold (Au) or silver (Ag).

For an inverted n-i-p perovskite solar cell (Fig 2.4 (c)), the transparent bottom contact is usually based on Indium Tin Oxide (ITO). The bottom HTL is usually a transparent p-type semiconductor, commonly PEDOT:PSS [42][57] or a p-type oxide such as NiO_x [58], and the ETL on top of the perovskite is usually an n-type semiconductor such as PCBM [57], C_{60} [59], or ZnO [60].

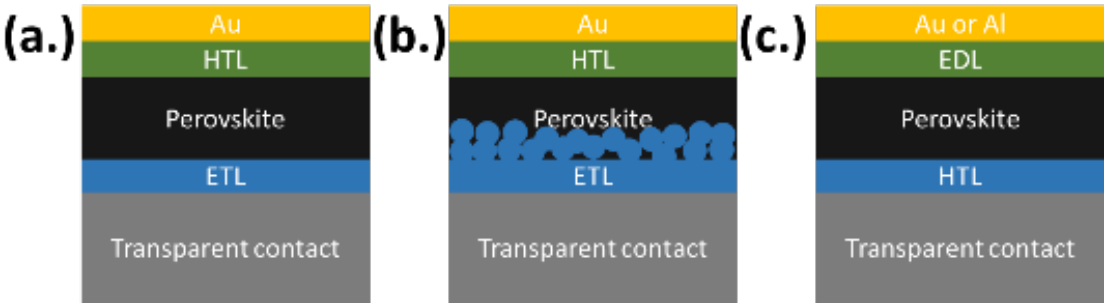


Fig 2.4 (a) Conventional p-i-n, (b) mesoscopic p-i-n, and (c) inverted n-i-p perovskite solar cell structures.

It is worth noting that researchers have demonstrated ETL-free or HTL-free perovskite solar cells based on a simple structure with a perovskite layer and an ETL or HTL sandwiched between two contacts [61][62]. However, these devices usually suffer from relatively poor performance due to a lower fill factor and open circuit voltage.

2.2.5 Working principle of perovskite solar cells

Perovskite solar cells, despite their diverse materials and device structures, fundamentally operate like conventional p-n or p-i-n photodiodes, with the utilization of perovskite materials [63].

For a conventional p-n diode, due to the carrier concentration gradient at the interface between the p and n side, the electrons tend to diffuse toward the p-type region and leave a space of positively charged donor ions in the n-type material near the interface. The holes migrate from the p-type region toward the n-type region, leaving a region of negatively charged dopants near the interface in the p-type semiconductor. This migration of carriers results in a built-in electric field from n to p, which generates a drift current opposite to the direction of diffusion. This drift and diffusion eventually reach a dynamic equilibrium, creating a region where carriers are depleted.

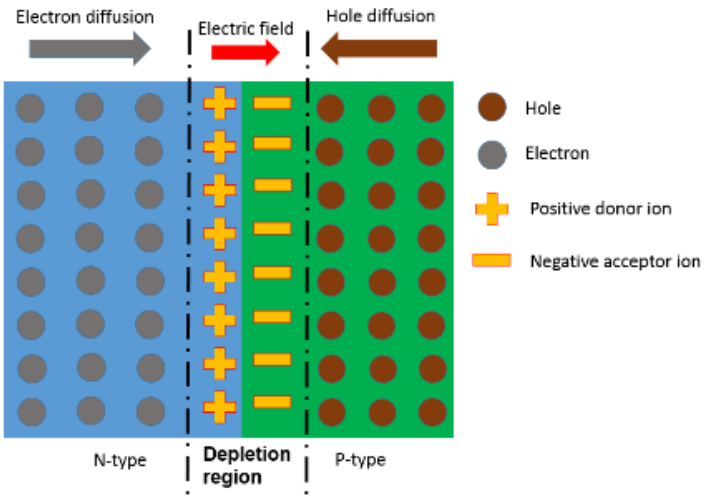


Fig 2.5 Structure of a p-n diode.

In the case of p-i-n diodes like perovskite solar cells, the p-type material is typically PEDOT:PSS or spiro-MeOTAD, the n-type material is usually PCBM or TiO2, and the intrinsic semiconductor is the perovskite. Both the p-i and i-n interfaces experience diffusion and drift, with the built-in electric field from n to p driven by the potential difference at the p-i and i-n interfaces. The intrinsic perovskite semiconductor acts as an extended depletion region between p- and n-type semiconductors, with its band tilted due to the potential difference between the p-i and i-n interfaces [63].

Upon exposure to light, if a photon with energy higher than the band gap of the perovskite is absorbed in this region, the electron and hole generated in the perovskite layer become separated due to the built-in electric field in this region, as illustrated by the band diagrams in Fig 2.6 (a). Because perovskite has a diffusion length of more than 1 μm [5], the electrons generated in this region have a high likelihood of passing through the intrinsic semiconductor, which is usually a few hundred nanometers thick, and recombining at the

n-type semiconductor. Similarly, holes have a higher probability of crossing the intrinsic layer to the n-type region. This mechanism produces a photocurrent in the device.

The corresponding circuit model for the p-i-n structure is presented in Fig 2.6 (b), where the current source represents the photocurrent I_L and the p-i-n structure is modeled as a diode. The circuit model also includes series and shunt resistances, R_s and R_{sh}, which account for the resistance introduced by the body of the p-i-n diode and the resistance of the leakage path between the p-type and n-type materials. Following this circuit model, the current and voltage are expressed as follows:

$$I = I_L - I_0[\exp\left(\frac{q(V-IR_S)}{nKT}\right) - 1] - \frac{V+IR_S}{R_{sh}} \quad \text{(Eq 2.2)}$$

where q is the element charge, V is the voltage applied across the diode terminals, I is the net current flowing through the diode, I_0 is dark saturation current, n is the ideality factor indicating how closely the diode behavior matches the ideal diode behavior, K is Boltzmann's constant, and T is the absolute temperature.

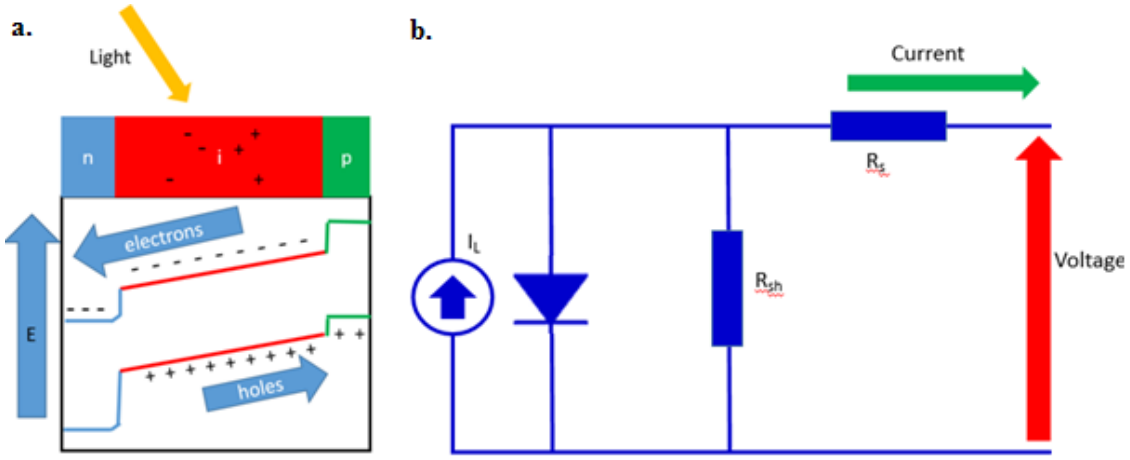


Fig 2.6 (a) Schematic band diagram of a p-i-n diode and (b) circuit model of a solar cell.

The operation environment of a solar cell varies greatly depending on the location and time of the day, but in a laboratory environment, solar cell measurements are usually performed under a standardized spectral irradiance defined as AM1.5 with a power equivalent to 1000 W/m^2. The AM1.5 standard represents the average sunlight at all wavelengths reaching the Earth's surface at an incident angle of 48.2° at mid-latitudes. Measurements of perovskite solar cells are performed under this illumination at an applied bias that is usually within the limits of the open circuit voltage of the cell, as an applied bias higher than the set potential tends to damage the perovskite solar cell [63].

The working conditions of any solar cell can be divided into three types as depicted in the current density-voltage (J-V) and power-voltage (P-V) curves in Fig 2.7 for ambient AM1.5 sunlights: open circuit, short circuit, and an intermediate condition. In the open circuit condition, when V_{oc} is measured, all of the generated carriers must recombine, as there is no current. The theoretical limit of V_{oc} is the band gap, which is approximately 1.55 eV at room temperature for perovskite solar cells. However, in actual operation, the limit is heavily affected by non-radiative recombination and the band gap of the ETM and HTM [64][65]. In contrast, in the short circuit condition, recombination in the cell is minimized, and the current density is dominated by the amount of carriers produced and the series resistance [65].

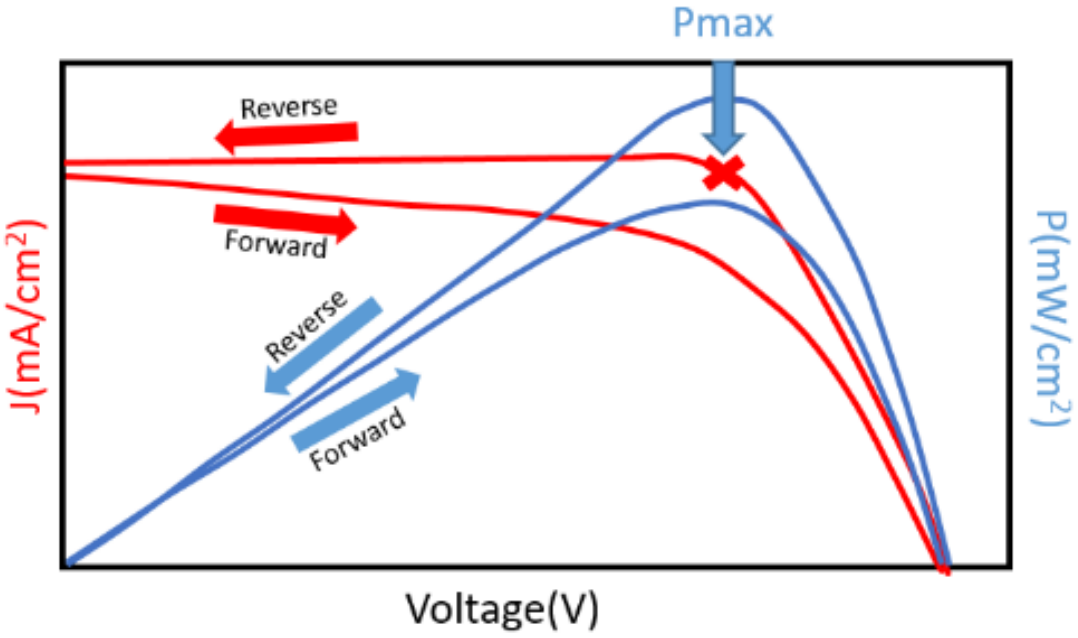


Fig 2.7 Schematic J-V and P-V curves of a perovskite solar cell measured under simulated solar light at AM1.5.

Indeed, the generation of a photocurrent in perovskite solar cells primarily occurs with the absorption of light at a wavelength below 800 nm. The series resistance within the cell is influenced by the materials and thickness of each layer and is dominated mainly by the carrier mobility within the intrinsic perovskite layer. This carrier mobility is influenced by grain size. There have been reports indicating that the carrier mobility is significantly reduced when the grain size is smaller than 100 nm, but when the grain size exceeds approximately 300 nm, the carrier mobility no longer increases because it reaches the single-crystal limit of 24.8 ± 4.1 $cm^2V^{-1}s^{-1}$ [29][31][66].

The conditions of open circuit and short circuit do not produce any output power. The maximum power output is often influenced by factors such as environmental temperature and internal factors like shunt and series resistances, and usually occurs at a

specific point on the J–V curve. At this point, the current density and voltage (marked as Vmax and Jmax) are often smaller than Voc and J_{SC}. The fill factor (FF) is used alongside Voc and J_{SC} to define the maximum power output of a solar cell. In general, the fill factor represents the 'squareness' of the J-V curve and is defined as:

$$FF = \frac{V_{max}J_{max}}{V_{oc}J_{SC}} = \frac{P_{MAX}}{V_{oc}J_{SC}} \quad \text{(Eq 2.3)}$$

Indeed, hysteresis is often observed in perovskite solar cells [67][68], as indicated in Fig 2.7. This behavior, where the forward scan typically generates a current lower than the reverse current, isn't desired for photovoltaic applications due to its unpredictable nature and the complications it introduces in understanding device performance. Interestingly, while hysteresis is a challenge for photovoltaic technology, it may provide an opportunity for other electronic applications, such as memory devices.

The origins of hysteresis in perovskite solar cells are not entirely understood currently [67][68]. Several contributing factors include charge accumulation at the interface [69], electron trapping at defects [37], crystal size and defects within the perovskite [70], structural instability, and the rotational MA compound leading to different electrical characteristics [33].

Although this hysteresis generally does not affect the operation of the solar cell (since the system stabilizes after a few seconds), it does affect the estimation of shunt and series resistances due to the varying forward and reverse scan results. Determining the ideality factor also poses a challenge in perovskite solar cells. While the factor remains at a constant value between 1 and 2 for a typical solar cell, in perovskite solar cells, it heavily depends on the HTM, where an ideality factor greater than 2 is possible.

2.2.6 Stability of perovskite solar cells

The perovskite solar cell often suffers performance losses when exposed to air or operating in high humidity, caused by the well-documented instability in perovskite-based photovoltaic devices. This instability of the perovskite solar cell is a complex phenomenon influenced by multiple factors, including the selection of the Electron Transport Material (ETM) and Hole Transport Material (HTM) [71][72], the deposition method [59], and the morphology and composition of the perovskite [73][74].

The environmental instability of $CH_3NH_3PbI_3$ perovskite is among the most significant factors contributing to its overall instability. Moisture-induced degradation is a prevalent issue for this perovskite material [75]. The hydrophilic nature of the methylammonium ion and the inherent instability of the perovskite material make the perovskite prone to water absorption and degradation into PbI_2 [76]. This degradation process begins at the grain boundaries and edges of the perovskite film [77]. Though moisture-induced degradation initially involves a reversible process via the formation of $(CH_3NH_3)_4PbI_6 \cdot 2H_2O$, this compound can break down further into PbI_2, resulting in permanent damage to the perovskite film.

The perovskite is also susceptible to oxygen-induced degradation upon light absorption, another major degradation pathway [78][79]. This process usually happens during light absorption, and in the presence of poor ETM coverage or insufficient carrier transport, photo-excited electrons may accumulate on the perovskite surface. Assisted by acidic protons in $CH_3NH_3^-$, these accumulated electrons can convert O_2 into superoxide O^{2-} [80]. This superoxide then catalyzes the degradation of $CH_3NH_3PbI_3$ following the reaction below:

$$4H_2O + 4CH_3NH_3PbI_3 \leftrightarrow 4[CH_3NH_3PbI_3 \cdot H_2O]$$

$$4[CH_3NH_3PbI_3 \cdot H_2O] \leftrightarrow (CH_3NH_3)_4PbI_6 \cdot 2H_2O + 3PbI_2 + 2H_2O$$

$$(CH_3NH_3)_4PbI_6 \cdot 2H_2O \xrightarrow{\text{moisture or light}} 4CH_3NH_3I + PbI_2 + 2H_2O$$

$$O_2 + 4CH_3NH_3PbI_3 \xRightarrow{\text{accumulated electrons}} 4PbI_2 + 4CH_3NH_2 + 2I_2 + 2H_2O$$

Superoxide-induced degradation not only breaks down $CH_3NH_3PbI_3$ into PbI2 but also generates H_2O, which can further induce moisture-related degradation. While superoxide-induced degradation can rapidly degrade an exposed perovskite film, it often isn't the primary cause of a perovskite solar cell's instability. With efficient ETM coverage and, in some cases, surface passivation, the likelihood of superoxide formation is significantly reduced compared to a perovskite film directly exposed to air. Consequently, by selecting an efficient electron transport layer and employing techniques such as defect passivation, oxygen-induced degradation can be minimized [80][81].

2.3 Amorphous Oxide Semiconductors

2.3.1 Applications of AOS semiconductors.

Amorphous Oxide Semiconductors (AOS), with their substantial bandgap of approximately 3.3eV, present an exciting prospect for transparent electronics. Their compatibility with low-temperature processing augments their applicability in flexible electronic applications compatible with plastic substrates [82]. Notably, their potential was realized in 2012 when Sharp Corporation pioneered the use of Indium Gallium Zinc Oxide-Thin Film Transistors (IGZO-TFTs) in manufacturing Liquid Crystal Display (LCD) panels. These panels have since been widely adopted in smartphones, tablets, and 32" LCDs [82]. Despite these advancements, challenges remain. High reactivity of oxygen in

AOS leads to accelerated aging and reduced voltage sensitivity over time. This issue becomes particularly pronounced when compared to displays made of amorphous Silicon (a-Si:H) and Low-Temperature Polycrystalline Silicon (LTPS). Another challenge involves the intricate process of production, which necessitates precise control over quarternary material ratios and utilization of rare earth elements, including Indium and Gallium [83].

Previous applications of polysilicon TFT LCDs were faced with the challenge of integrating a capacitor for refreshment due to inadequate capacitance of a TFT in a Dynamic Random Access Memory (DRAM) architecture. Mitsubishi's pioneering solution in 2001 was to integrate DRAM at the pixel level [84]. In 2008, a capacitorless polysilicon TFT DRAM was introduced, featuring an embedded trench for charge storage. Despite requiring a dual gate architecture and having an unreported retention time, it offered a memory window twice as large as a conventional TFT [85].

Recent advancements have seen the introduction of a novel capacitor-less DRAM cell architecture by imec in AOS materials fully compatible with displays. Implemented on a 45 nm gate length on 300 mm wafers, this architecture combines two IGZO-TFTs and offers a remarkable retention time exceeding 400 seconds due to an exceptionally low off-current of 3×10^{-19}A/μm [86].

Another innovation was the demonstration of a multilayer stackable 3D NAND Flash memory in polysilicon TFTs by Macronix in 2006 [87]. This showcased the feasibility of 3D NAND in TFTs. Although NOR flash memory in polysilicon TFTs has been applied in synaptic applications by Kim et al. [88], no instances of such applications in AOS semiconductors have been reported to date.

The most recent development has been the proposal of IGZO-based DRAM cells to mitigate IR voltage drops on both the summation and activation lines of a crossbar array. These cells, with their extremely low off-currents and compatibility with low-temperature BEOL processing, were proposed as a potential solution for large-scale cross-bar arrays aiming for 10000 Tera Operations Per Second per Watt (TOPs/W), offering an advantage over STT MRAM and projected PCM [89].

On the photovoltaics front, amorphous oxides, specifically ZnO, has been drawing attention in the development of air-stable perovskite solar cells. They are being explored for their potential to enhance stability [90] and cut costs compared to other electron transport materials like PCBM and TiO_2. Demonstrations since 2015 have reported successful integration of ZnO in both standard [91] and inverted perovskite solar cell structures [90].

The process of positioning ZnO varies depending on whether it is placed below or on top of the perovskite layer. For instance, the solution-based method can be employed both above [92] and below the perovskite layer [93]. On the other hand, Radio frequency (RF) sputtering is recommended if the ZnO layer is deposited prior to the perovskite layer, with such cells yielding the highest reported efficiency of 20.2% among ZnO-based perovskite solar cells [91].

Given the ease with which ZnO—a binary compound—can be fabricated across large areas and at high rates using techniques like RF sputtering, Pulse Laser Deposition, Chemical Vapour Deposition, and solution-based processing, its economic value has skyrocketed, exceeding 500M Euros. This value stems from diverse applications such as UV filters, surge protection varistors, light scattering layers in silicon solar cells [94], and

gas sensors, which function on conductivity alterations induced by gas adsorption and desorption.

Moreover, ZnO's capability for monolithic integration into not just solar cells but also silicon-based electronic circuitry, coupled with its amenability to chemical processing, has led to its employment in Surface Acoustic Wave (SAW) multiplexing/demultiplexing bandpass filters. These filters are mainly used in telecommunications applications, particularly in cellular phones and base stations [94].

2.3.2 Basic structure and property of amorphous oxide semiconductors

Amorphous oxide semiconductors (AOS) have emerged as formidable competitors in flexible, large-scale applications, given their capability to deliver superior performance while requiring relatively low processing temperatures of below 300 ◦C [95]. Notably, room temperature processed mobilities of up to 10 cm^2/Vs have been realized, representing an advancement that is an order of magnitude higher than that of a:Si:H [96]. This diminution in mobility within a:Si:H (less than 1 cm^2/Vs) can be attributed to the material's bonding [97], which is contingent on the overlapping of disordered sp^3 hybridized orbitals among shared atoms. Conversely, in 1996, Hosono et al. proposed a mechanism to explain the enhanced mobility in AOSs [98]. This discovery paved the way for the invention of unique materials and devices based on amorphous oxide semiconductors, thereby expanding the horizons for large-scale flexible electronic applications.

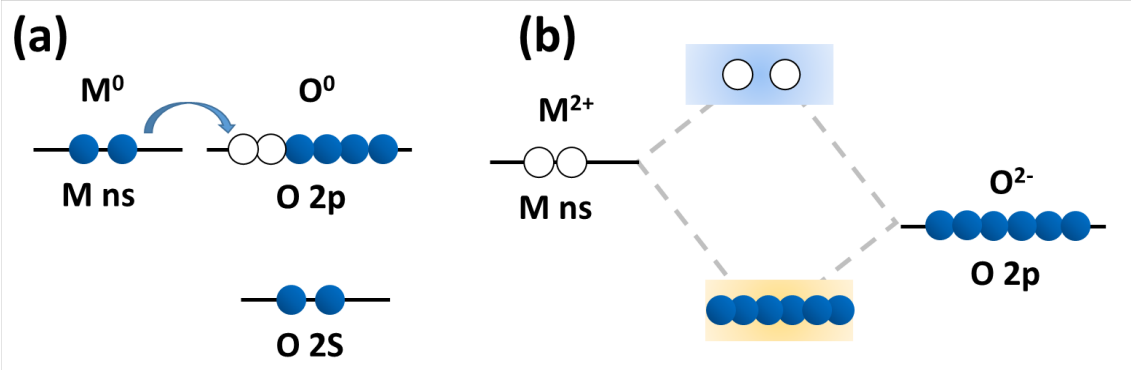


Fig 2.8 Schematic electronic structures of ionic oxide semiconductors, (a) the charge transfer from metal to oxygen atoms, (b) the unoccupied s orbitals forms the CBM and the fully occupied O 2p orbitals forms the VBM [99].

In amorphous oxide semiconductors, the marked iconicity of metal atoms within an AOS triggers a charge transfer from the metal to the electronegative oxygen atoms (Fig 2.8 (a)). This electronic structure gains stability by raising cation levels and diminishing anion levels. The Conduction Band Maximum (CBM) consists of unoccupied s orbitals, whereas the Valence Band Maximum (VBM) is comprised of completely occupied O 2p orbitals of cations [82]. This configuration results in a minimal electron effective mass, thereby contributing to high electron mobility and reduced power consumption. Furthermore, the expansive radii provided by the vacant s orbitals permit robust overlap with each other, constructing a potent conduction pathway. This differs starkly from a:Si:H, which relies on the overlap of sp^3 hybridized, highly directional bonds for current conduction.

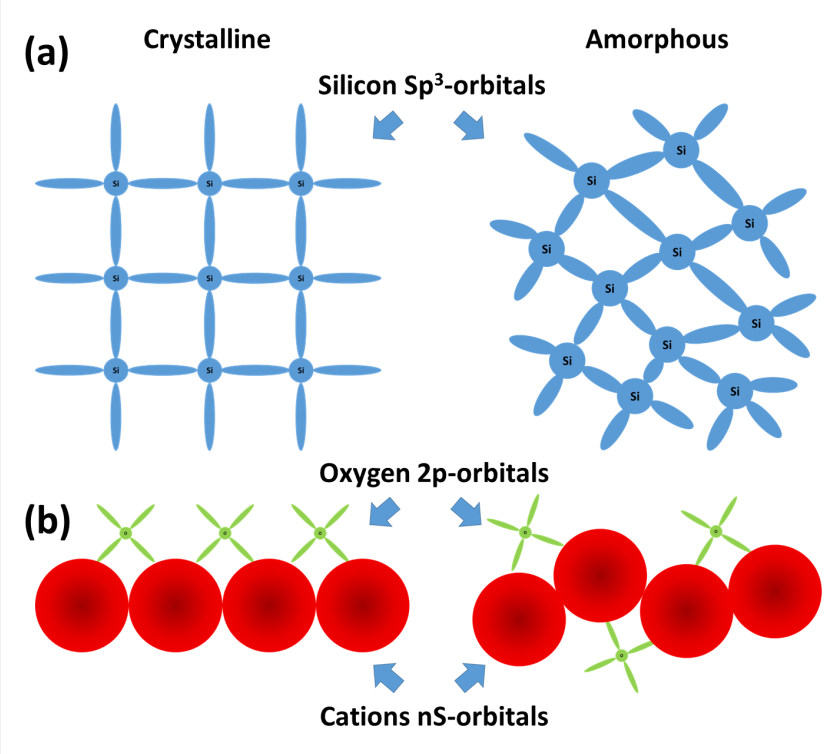


Fig 2.9 Schematic drawing of the CBM atomic orbital and carrier transport paths of crystalline and amorphous Si (a) and (b) ionic oxide semiconductors [99].

The band structure of amorphous semiconductors also diverges from their crystalline counterparts. Current theories regarding the electronic configuration of amorphous semiconductors encompass extended state bands, band tails, and defect states (Fig 2.10). The manifestation of extended state bands can be credited to the short-range order present in a-Si:H as compared to traditional single-crystal silicon. Within this context, states within the valence band can be viewed as bonding states, while those in the conduction band can be considered antibonding states. Situated below the conduction band or above the valence band are localized electronic states known as Band tail states [100]. These tail states originate from the lack of long-range order or the existence of charged defects that locally disrupt the energy levels at the band edges. The valence band tail can

be perceived as strained or weak bonding states, while the conduction band tail represents the corresponding antibonding states. In pristine materials like a-Si, defect states are postulated to consist of three-fold coordinated silicon dangling bonds, or non-bonding states. In ionic oxide semiconductors, it's established that specific structural defects, such as vacancies, substitutions, anti-site configurations, and interstitial atoms, can generate carrier traps when they introduce additional electron states within the bandgap, thereby giving rise to electronic defects [101].

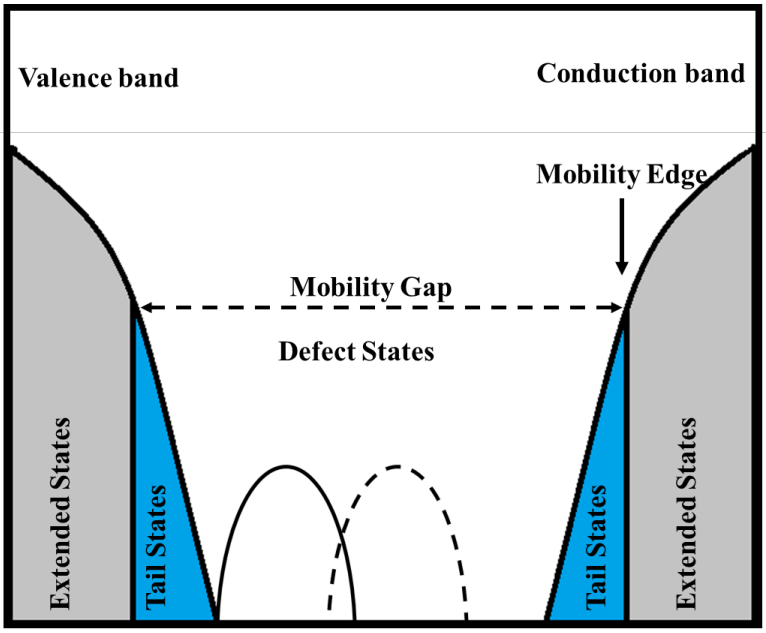


Fig 2.10 Schematic drawing of amorphous semiconductors Density of States

Contrary to other kinds of semiconductor materials, where carrier transport is frequently impeded by bulk defects, amorphous oxide semiconductors showcase degenerate band conduction that is unhindered by band tails. Instead, their conduction depends on channel conduction overseen by metal cations, resulting in films with high electron mobilities [102]. Carrier transport in amorphous oxide semiconductors is typically

regulated by two primary mechanisms: trap-limited conduction (TLC) [100][103] and the percolation mechanism [104][105].

Trap-limited conduction (TLC) is usually governed by low conductance state under low voltage conditions in AOSs. In this case, carrier transport occurs through a succession of trapping and de-trapping events under the applied electric field [100]. Carriers encounter localized trap states within the material's bandgap. These carriers are temporarily captured in the traps before being released under the applied electric field. This cycle of trapping and de-trapping events forms the overall conduction process. However, it leads to a reduced carrier mobility compared to higher voltage regimes or crystalline semiconductors with fewer defects.

The percolation mechanism, on the other hand, gains prominence under high voltage conditions. In this regime, carriers can find a continuous path through the material by hopping between adjacent conducting regions, referred to as a percolation path [104]. The formation of these paths is influenced by the density and distribution of conducting regions and insulating barriers within the material. As the applied voltage or electric field increases, the probability of percolation occurring also rises, leading to enhanced conductivity and improved carrier mobility in amorphous oxide semiconductors.

2.3.3 Deposition method of Amorphous Oxide Semiconductors

Sputtering is a physical vapour deposition method routinely utilized for the deposition of metal contacts, magnetic materials, and amorphous oxide semiconductors. The basic configuration of a sputtering system is depicted in Fig 2.12. Sputtering is a PVD technique where deposition is executed via ionized plasma, which transfers the target material onto substrates. The specific type of sputtering can be determined by the voltage

supplied to the system, with two primary types being Direct Current (DC) sputtering and Radio Frequency (RF) sputtering.

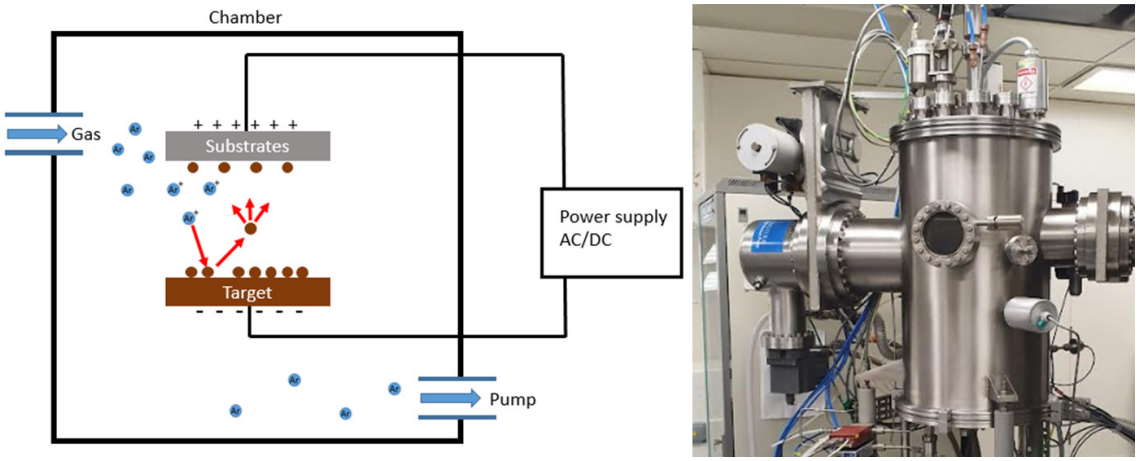


Fig 2.11 Schematic diagram and photo of the Lesker sputter system used in this work.

Direct Current (DC) sputtering is a fundamental version of sputtering systems, commonly used for the sputtering of conductive materials like metal contacts. During the deposition process, an inert gas, generally argon (Ar), is introduced into the system. The gas flow is managed by a mass flow controller to maintain the necessary deposition pressure. A DC voltage is applied between the substrate and the target, with the substrate acquiring a positive voltage and the target receiving a negative one. Under this high electric field, an Ar atom loses an electron to the anode, initiating the creation of Ar plasma. Propelled by the high electric field between the anode and cathode, these ions collide with the target material. The high-energy Ar^+ ions impact and transfer their kinetic energy to the target material, which subsequently reflects and deposits back onto the substrates.

However, when the target material is an insulator, the insulating target can lead to charge accumulation during plasma generation. This accumulation slows down the deposition process and can eventually bring it to a halt. To circumvent this issue, Radio

Frequency (RF) sputtering is employed for deposition of insulating materials. In this scenario, an AC voltage is supplied to the system, and the deposition process transforms into a two-stage cycle. In the first stage, a negative voltage is applied to the target, and the deposition process proceeds similarly to DC sputtering. In the second stage, a positive voltage is applied to the target. This halts the deposition process and removes the charge accumulated during the first stage. The system rapidly cycles between these two stages at a frequency of 13.56 MHz to facilitate the deposition process.

2.3.4 Thin film Transistor architecture

Thin Film Transistors (TFTs) are three-terminal field-effect devices that are commonly constructed using amorphous semiconductors. These devices typically consist of a gate, source, and drain electrode, with the semiconductor material sandwiched between them. Depending on the order and manner in which the gate, oxide semiconductor, and drain/source electrodes are constructed, there are four commonly investigated device configurations, namely Top-gate coplanar, Top-gate staggered, Bottom-gate bottom coplanar and Bottom-gate staggered are visually represented in Fig 2.12.

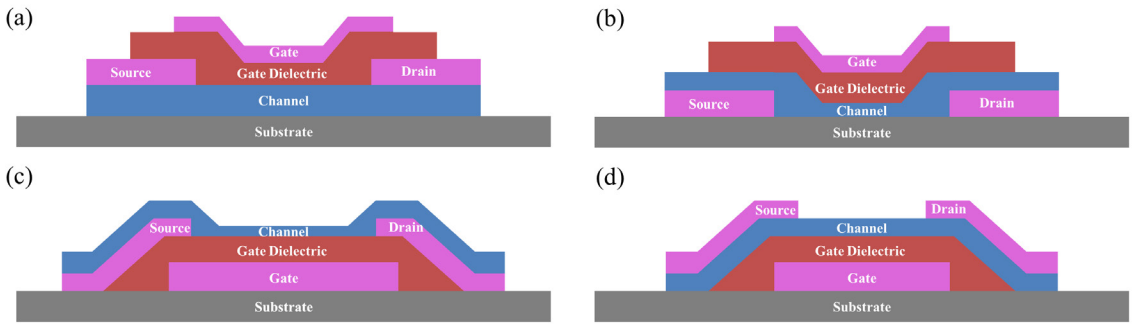


Fig 2.12 Schematic diagram of basic TFT structures. (a) Top-gate coplanar (b) Top-gate staggered (c) Bottom-gate bottom coplanar and (d) Bottom-gate staggered.

2.3.5 Principle of operation and physics of AOS based TFTs

Most of the thin-film transistors (TFTs) are "normally off" devices, a state achieved via their usage of amorphous semiconductors renowned for their high resistivity in the non-conducting state. The operational procedures and device characterization typically mandate the grounding of the drain terminal. Current flow modulation from the source to the drain is governed by the applied gate voltage, an interaction that occurs via the Metal-Insulator-Semiconductor (MIS) interface. Noteworthy is the TFT's operational preference for the accumulation region within the MIS structure.

The distinctive current-voltage ($I_{DS} – V_{DS}$) characteristics of TFTs are divisible into three significant operational regions (Fig 2.13). Of primary importance is the threshold voltage, V_{TH}, which signifies the minimal gate voltage required to form a conducting channel bridging the drain and source terminals. Under conditions where $V_G < V_{TH}$, the device resides in the cut-off region, thereby inhibiting the formation of a conductive path between the drain and source and effectively reducing the drain-source current to zero.

A contrasting operational state is observed when $V_G > V_{TH}$ and $V_D << V_G - V_{TH}$. Here, the device transitions into the linear region, characterized by the channel resistance adhering to Ohm's law, which consequently establishes a linear relationship between the drain-source current and drain voltage. The drain current within this linear region can be encapsulated in Equation (2.4):

$$I_D = \left(\mu C_0 \frac{W}{L}\right)[(V_G - V_{TH})V_D] \quad \text{(Eq 2.4)}$$

In this equation, W represents the channel width, L is the channel length, C_0 denotes the gate oxide capacitance, μ symbolizes the carrier mobility, V_G is the gate voltage, V_D is the drain-source voltage, and V_{TH} is the threshold voltage.

Upon reaching a state where $V_G > V_{TH}$ and $V_D = V_G - V_{TH}$, the device enters the saturation region. In this regime, the escalating drain potential causes the potential difference between the gate and the drain to decrease, and at times, turn negative. As a consequence, the device becomes incapable of maintaining a conducting channel between the drain and source, thereby causing the region near the source to be pinched off. Despite this occurrence, the drain-source remains highly conductive owing to the significant drain-source voltage, which allow carriers to tunnelling through the pinched-off region. Additional increases in V_D expand the pinched-off region towards the source. However, this expansion is offset by the enhanced tunnelling effect that accompanies an increased drain-source voltage. The device then enters a saturation region where I_D remains constant despite any subsequent increases in the drain-source voltage. The drain current within this saturation region is defined by Equation (2.5):

$$I_D = \left(\mu C_0 \frac{W}{2L}\right)(V_G - V_{TH})^2 \qquad \text{(Eq 2.5)}$$

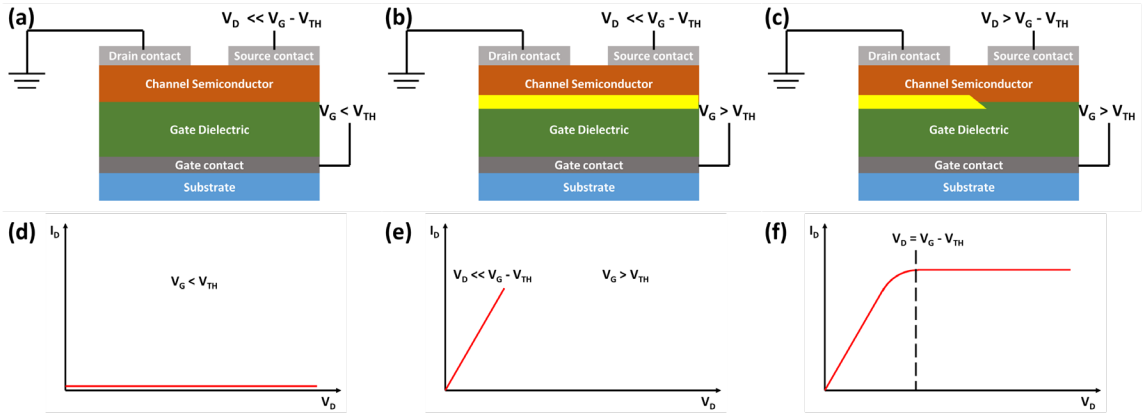


Fig 2.13 Schematic diagram (a-c) and corresponding I_D –V_D curve (d-f) of a TFT in (a) cut-off region, (b) linear region and (c) saturation region.

In the realm of amorphous oxide-based thin-film transistors (TFTs), a significant gate bias dependency of field-effect mobility is often observed, attributed to the interplay of percolation and trap-limited conduction mechanisms [106]. This interaction frequently results in non-linearity within the device's transfer characteristics. These mechanisms can be quantitatively generalized through a modified field-effect mobility equation, denoted as Equation (2.6):

$$\mu_{EF} = \mu_0 (V_G - V_{th})^{\gamma} \quad \text{(Eq 2.6)}$$

In this equation, μ_0 and γ are fitting parameters associated with the field-effect mobility. Given this modified definition of field-effect mobility, Equations (Eq 2.7) and (Eq 2.8) evolve into [106][107]:

$$I_D = \left(C_0 \frac{W}{L}\right) [\mu_0 (V_G - V_{th})^{\gamma} V_D] \quad \text{(Eq 2.7)}$$

$$I_D = \left(W \frac{C_0}{L(1+\gamma)}\right) \mu_0 (V_G - V_{th})^{1+\gamma} V_D \quad \text{(Eq 2.8)}$$

Performance evaluation of TFTs typically entails the consideration of several key parameters that function as benchmarks in the comparative analysis of different devices. Such parameters prominently include the on/off ratio and the subthreshold swing (SS).

The on/off ratio provides a comparative measure of the maximum and minimum drain current I_D, where the minimum I_D usually corresponds to the leakage current in the off state, and the maximum I_D is influenced by various factors including the materials constituting the channel and gate dielectric, the geometry and area of the device, and the applied drain and gate bias. In most applications, a larger on/off ratio is sought after.

The subthreshold swing (SS) is defined as the inverse derivative of the logarithm of the drain current with respect to the gate voltage. It quantifies the gate voltage increment required to alter I_D by an order of magnitude in the subthreshold region, i.e., when the gate voltage (V_G) is lower than the threshold voltage (V_{th}). This relationship is articulated as:

$$SS = \left[\frac{d \log (I_D)}{d V_G}\right]^{-1} \quad \text{(Eq 2.9)}$$

The Eq 2.9 SS could be rewriten as:

$$SS = \left[\frac{d \log (I_D)}{d V_G}\right]^{-1} = \left[\frac{d \Psi_s}{d V_G}\frac{d \log (I_D)}{d \Psi_s}\right]^{-1} = m \times n \quad \text{(Eq 2.10)}$$

In this context, *n* symbolizes the transport factor a constant where $n = 2.3\, k_B T/q$, also referred to as the Boltzmann limit, which is approximately 60 mV/dec at room temperature. The variable *m*, referred to as the body factor, can be simplified as:

$$m = 1 + \frac{C_s}{C_i} \quad \text{(Eq 2.11)}$$

Here, C_i and C_s denote the gate insulator capacitance and a differential capacitance, respectively, the latter representing the depletion region capacitor within the channel semiconductor. Since in typical dielectric materials and semiconductors, C_i and C_s are

invariably positive, the body factor *m* consistently exceeds 1, leading to a subthreshold swing limit of 60mV/dec at room temperature.

2.4 Summary

Perovskites, with their unique optoelectronic properties, offer promising prospects in energy conversion, particularly in enhancing the efficiency and cost-effectiveness of solar cells. In parallel, AOS, with their high carrier mobility and exceptional transparency, are forging new paths in the electronics industry, particularly in supporting the creation of advanced memory technologies, TFT application in display as well as carrier transporting materials in perovskite solar cell. The importance of perovskite materials and amorphous oxide semiconductors (AOS) in the landscape of energy and memory applications motivates this work.

In this context, this chapter provide an exploration of the fundamental attributes of these materials. This exploration incorporates a comparative review of their respective structures and material properties, illuminating the attributes that define their functionality and utility in various applications.

Moreover, the discourse extends to a clear exposition of the fundamental device structure and the associated operational physics underlying perovskite-based solar cells. This dissection of the working principles of these cells provides insights into their efficiency and potential challenge when it comes to stability.

In a similar vein, the chapter presents a detailed analysis of amorphous oxide-based Thin-Film Transistors (TFTs). The operating principles of these transistors, their design, and the influence of the unique properties of amorphous oxide semiconductors on their performance are brought to the fore. This comprehensive exploration provides a deeper

understanding of the role and potential of these materials in the broader context of technological advancement in energy and electronics.

2.5 References

[1] A. Kojima *et al.*, "Novel Photoelectrochemical Cell with Mesoscopic Electrodes Sensitized by Lead-halide Compounds (11)," *ECS Meet. Abstr.*, vol. MA2008-02, no. 1, pp. 27–27, Aug. 2008.

[2] A. Kojima *et al.*, "Organometal halide perovskites as visible-light sensitizers for photovoltaic cells," *J. Am. Chem. Soc.*, vol. 131, no. 17, pp. 6050–6051, 2009.

[3] L. Etgar *et al.*, "Mesoscopic CH 3 NH 3 PbI 3 /TiO 2 Heterojunction Solar Cells," *J. Am. Chem. Soc.*, vol. 134, no. 42, pp. 17396–17399, Oct. 2012.

[4] M. M. Lee *et al.*, "Efficient Hybrid Solar Cells Based on Meso-Superstructured Organometal Halide Perovskites," *Science (80-.).*, vol. 338, no. 6107, pp. 643–647, Nov. 2012.

[5] S. D. Stranks *et al.*, "Electron-Hole Diffusion Lengths Exceeding 1 Micrometer in an Organometal Trihalide Perovskite Absorber," *Science (80-.).*, vol. 342, no. 6156, pp. 341–344, Oct. 2013.

[6] J. Burschka *et al.*, "Sequential deposition as a route to high-performance perovskite-sensitized solar cells," *Nature*, vol. 499, no. 7458, pp. 316–319, Jul. 2013.

[7] N. Pellet *et al.*, "Mixed-Organic-Cation Perovskite Photovoltaics for Enhanced Solar-Light Harvesting," *Angew. Chemie*, vol. 126, no. 12, pp. 3215–3221, Mar. 2014.

[8] P. Schulz *et al.*, "Interface energetics in organo-metal halide perovskite-based photovoltaic cells," *Energy Environ. Sci.*, vol. 7, no. 4, p. 1377, 2014.

[9] H. Zhou *et al.*, “Interface engineering of highly efficient perovskite solar cells,” *Science (80-.).*, vol. 345, no. 6196, pp. 542–546, Aug. 2014.

[10] X. Li *et al.*, “A vacuum flash–assisted solution process for high-efficiency large-area perovskite solar cells,” *Science (80-.).*, vol. 353, no. 6294, pp. 58–62, Jul. 2016.

[11] X. Zheng *et al.*, “Defect passivation in hybrid perovskite solar cells using quaternary ammonium halide anions and cations,” *Nat. Energy*, vol. 2, no. 7, pp. 1–9, 2017.

[12] Q. Jiang *et al.*, “Surface passivation of perovskite film for efficient solar cells,” *Nat. Photonics*, vol. 13, no. 7, pp. 460–466, Jul. 2019.

[13] M. A. Green *et al.*, “Solar cell efficiency tables (Version 61),” *Prog. Photovoltaics Res. Appl.*, vol. 31, no. 1, pp. 3–16, Jan. 2023.

[14] K. Xiao *et al.*, “Scalable processing for realizing 21.7%-efficient all-perovskite tandem solar modules,” *Science (80-.).*, vol. 376, no. 6594, pp. 762–767, May 2022.

[15] H. S. Kim *et al.*, “Organolead Halide Perovskite : New Horizons in Solar Cell Research Organolead Halide Perovskite : New Horizons in Solar Cell,” *J. Phys. Chem. C*, vol. 118, no. 11, pp. 5615–5625, 2014.

[16] V. M. Goldschmidt, “Die Gesetze der Krystallochemie,” *Naturwissenschaften*, vol. 14, no. 21, pp. 477–485, 1926.

[17] T. Jesper Jacobsson *et al.*, “Exploration of the compositional space for mixed lead halogen perovskites for high efficiency solar cells,” *Energy Environ. Sci.*, vol. 9, no. 5, pp. 1706–1724, 2016.

[18] M. A. Green *et al.*, “The emergence of perovskite solar cells,” *Nat. Photonics*, vol. 8, no. 7, p. 506, 2014.

[19] J. Lian *et al.*, “Organic solvent vapor sensitive methylammonium lead trihalide film

formation for efficient hybrid perovskite solar cells," *J. Mater. Chem. A*, vol. 3, no. 17, pp. 9146–9151, 2015.

[20] Q. Wang *et al.*, "Transition from the Tetragonal to Cubic Phase of Organohalide Perovskite: The Role of Chlorine in Crystal Formation of CH3NH3PbI3 on TiO2 Substrates," *J. Phys. Chem. Lett.*, vol. 6, no. 21, pp. 4379–4384, 2015.

[21] T. Oku *et al.*, "Microstructures and photovoltaic properties of perovskite-type CH3 NH3PbI3 compounds," *Appl. Phys. Express*, vol. 7, no. 12, p. 121601, Dec. 2014.

[22] F. Hao *et al.*, "Anomalous band gap behavior in mixed Sn and Pb perovskites enables broadening of absorption spectrum in solar cells," *J. Am. Chem. Soc.*, vol. 136, no. 22, pp. 8094–8099, 2014.

[23] P. Wang *et al.*, "Advances in the structure and materials of perovskite solar cells," *Res. Chem. Intermed.*, pp. 625–639, 2015.

[24] W. Shockley and H. J. Queisser, "Detailed balance limit of efficiency of p-n junction solar cells," *J. Appl. Phys.*, vol. 32, no. 3, pp. 510–519, 1961.

[25] M. R. Filip *et al.*, "Steric engineering of metal-halide perovskites with tunable optical band gaps," 2014.

[26] K. T. Butler *et al.*, "Band alignment of the hybrid halide perovskites CH 3 NH 3 PbCl 3 , CH 3 NH 3 PbBr 3 and CH 3 NH 3 PbI 3," *Mater. Horizons*, vol. 2, no. 2, pp. 228–231, 2015.

[27] S. de Wolf *et al.*, "Organometallic Halide Perovskites: Sharp Optical Absorption Edge and," *J. Phys. Chem. C*, vol. 5, pp. 1035–139, 2014.

[28] M. A. Green *et al.*, "Solar cell efficiency tables (version 50)," *Prog. Photovoltaics Res. Appl.*, vol. 25, no. 7, pp. 668–676, 2017.

[29] O. G. Reid *et al.*, “Grain-Size-Limited Mobility in Methylammonium Lead Iodide Perovskite Thin Films,” *ACS Energy Lett.*, vol. 1, no. 3, pp. 561–565, Sep. 2016.

[30] Q. Han *et al.*, “Single Crystal Formamidinium Lead Iodide (FAPbI3): Insight into the Structural, Optical, and Electrical Properties,” *Adv. Mater.*, vol. 28, no. 11, pp. 2253–2258, 2016.

[31] Q. Dong *et al.*, “Electron-hole diffusion lengths > 175 μm in solution-grown CH 3 NH 3 PbI 3 single crystals,” *Science (80-.).*, vol. 347, no. 6225, pp. 967–970, Feb. 2015.

[32] G. Maculan *et al.*, “CH3NH3PbCl3 Single Crystals: Inverse Temperature Crystallization and Visible-Blind UV-Photodetector,” *J. Phys. Chem. Lett.*, vol. 6, no. 19, pp. 3781–3786, 2015.

[33] R. Gottesman *et al.*, “Extremely slow photoconductivity response of CH3NH 3PbI3 perovskites suggesting structural changes under working conditions,” *J. Phys. Chem. Lett.*, vol. 5, no. 15, pp. 2662–2669, 2014.

[34] T. Chen *et al.*, “Rotational dynamics of organic cations in the CH3NH3PbI3 perovskite,” *Phys. Chem. Chem. Phys.*, vol. 17, no. 46, pp. 31278–31286, 2015.

[35] Y. Wang *et al.*, “Density functional theory analysis of structural and electronic properties of orthorhombic perovskite CH3NH3PbI3,” *Phys. Chem. Chem. Phys.*, vol. 16, no. 4, pp. 1424–1429, 2014.

[36] Y. Rakita *et al.*, “Tetragonal CH 3 NH 3 PbI 3 Is Ferroelectric,” *Proc. Natl. Acad. Sci.*, vol. 114, no. 28, pp. 7191–7193, 2017.

[37] W. Tress *et al.*, “Understanding the rate-dependent J–V hysteresis, slow time component, and aging in CH 3 NH 3 PbI 3 perovskite solar cells: the role of a

compensated electric field," *Energy Environ. Sci.*, vol. 8, no. 3, pp. 995–1004, 2015.

[38] G. R. Berdiyorov *et al.*, "Electronic transport in organometallic perovskite CH3NH3PbI3: The role of organic cation orientations," *Appl. Phys. Lett.*, vol. 108, no. 5, pp. 3–7, 2016.

[39] A. Dualeh *et al.*, "Effect of Annealing Temperature on Film Morphology of Organic-Inorganic Hybrid Pervoskite Solid-State Solar Cells," *Adv. Funct. Mater.*, vol. 24, no. 21, pp. 3250–3258, Jun. 2014.

[40] J.-H. Im *et al.*, "Growth of CH3NH3PbI3 cuboids with controlled size for high-efficiency perovskite solar cells.," *Nat. Nanotechnol.*, vol. 9, no. 11, pp. 927–932, 2014.

[41] H. S. Kim *et al.*, "Lead iodide perovskite sensitized all-solid-state submicron thin film mesoscopic solar cell with efficiency exceeding 9%," *Sci. Rep.*, vol. 2, pp. 1–7, 2012.

[42] C.-H. Chiang *et al.*, "Planar heterojunction perovskite/PC 71 BM solar cells with enhanced open-circuit voltage via a (2/1)-step spin-coating process," *J. Mater. Chem. A*, vol. 2, no. 38, pp. 15897–15903, 2014.

[43] J. H. Im *et al.*, "Morphology-photovoltaic property correlation in perovskite solar cells: One-step versus two-step deposition of CH3NH3PbI 3," *APL Mater.*, vol. 2, no. 8, 2014.

[44] N. J. Jeon *et al.*, "Solvent engineering for high-performance inorganic–organic hybrid perovskite solar cells," *Nat. Mater.*, vol. 13, no. 9, pp. 897–903, 2014.

[45] S. Wang *et al.*, "Water-Soluble Triazolium Ionic-Liquid-Induced Surface Self-Assembly to Enhance the Stability and Efficiency of Perovskite Solar Cells," *Adv.*

Funct. Mater., vol. 29, no. 15, p. 1900417, Apr. 2019.

[46] Y. Yang *et al.*, "Bi-functional additive engineering for high-performance perovskite solar cells with reduced trap density," *J. Mater. Chem. A*, vol. 7, no. 11, pp. 6450–6458, 2019.

[47] Y. Zhou *et al.*, "Growth control of compact CH 3 NH 3 PbI 3 thin films via enhanced solid-state precursor reaction for efficient planar perovskite solar cells," *J. Mater. Chem. A*, vol. 3, no. 17, pp. 9249–9256, 2015.

[48] K. Liang *et al.*, "Synthesis and Characterization of Organic−Inorganic Perovskite Thin Films Prepared Using a Versatile Two-Step Dipping Technique," *Chem. Mater.*, vol. 10, no. 1, pp. 403–411, 1998.

[49] L. Chen *et al.*, "Facile preparation of organometallic perovskite films and high-efficiency solar cells using solid-state chemistry," *Nano Res.*, vol. 8, no. 1, pp. 263–270, Jan. 2015.

[50] Q. Chen *et al.*, "The optoelectronic role of chlorine in CH3NH3PbI3(Cl)-based perovskite solar cells," *Nat. Commun.*, vol. 6, no. 1, p. 7269, Nov. 2015.

[51] M. I. Dar *et al.*, "Investigation Regarding the Role of Chloride in Organic–Inorganic Halide Perovskites Obtained from Chloride Containing Precursors," *Nano Lett.*, vol. 14, no. 12, pp. 6991–6996, Dec. 2014.

[52] H. Yu *et al.*, "The Role of Chlorine in the Formation Process of 'CH 3 NH 3 PbI 3-x Cl x ' Perovskite," *Adv. Funct. Mater.*, vol. 24, no. 45, p. n/a-n/a, Sep. 2014.

[53] S. T. Williams *et al.*, "Role of Chloride in the Morphological Evolution of Organo-Lead Halide Perovskite Thin Films," *ACS Nano*, vol. 8, no. 10, pp. 10640–10654, Oct. 2014.

[54] E. L. Ratcliff *et al.*, "Energy level alignment in PCDTBT:PC70BM solar cells: Solution processed NiOx for improved hole collection and efficiency," *Org. Electron.*, vol. 13, no. 5, pp. 744–749, May 2012.

[55] M. M. Tavakoli *et al.*, "Surface Engineering of TiO 2 ETL for Highly Efficient and Hysteresis-Less Planar Perovskite Solar Cell (21.4%) with Enhanced Open-Circuit Voltage and Stability," *Adv. Energy Mater.*, vol. 8, no. 23, p. 1800794, Aug. 2018.

[56] N. Cheng *et al.*, "Application of mesoporous SiO2 layer as an insulating layer in high performance hole transport material free CH3NH3PbI3 perovskite solar cells," *J. Power Sources*, vol. 321, pp. 71–75, Jul. 2016.

[57] L.-C. Chen *et al.*, "Fabrication and Properties of High-Efficiency Perovskite/PCBM Organic Solar Cells," *Nanoscale Res. Lett.*, vol. 10, no. 1, p. 312, Dec. 2015.

[58] F. Jiang *et al.*, "Post-treatment-free solution-processed Non-stoichiometric NiOx nanoparticles for efficient hole-transport layers of organic optoelectronic devices," *Adv. Mater.*, vol. 27, no. 18, pp. 2930–2937, 2015.

[59] J. H. Kim *et al.*, "Enhanced Environmental Stability of Planar Heterojunction Perovskite Solar Cells Based on Blade-Coating," *Adv. Energy Mater.*, vol. 5, no. 4, pp. 2–7, 2015.

[60] J. You *et al.*, "Improved air stability of perovskite solar cells via solution-processed metal oxide transport layers," *Nat. Nanotechnol.*, no. October, pp. 1–8, 2015.

[61] L. Huang *et al.*, "Toward Revealing the Critical Role of Perovskite Coverage in Highly Efficient Electron-Transport Layer-Free Perovskite Solar Cells: An Energy Band and Equivalent Circuit Model Perspective," *ACS Appl. Mater. Interfaces*, vol. 8, no. 15, pp. 9811–9820, Apr. 2016.

[62] L. Huang and Z. Ge, “Simple, Robust, and Going More Efficient: Recent Advance on Electron Transport Layer-Free Perovskite Solar Cells,” *Adv. Energy Mater.*, vol. 9, no. 24, p. 1900248, Jun. 2019.

[63] E. Edri *et al.*, “Elucidating the charge carrier separation and working mechanism of CH3NH3PbI3-xClx perovskite solar cells,” *Nat. Commun.*, vol. 5, pp. 1–8, 2014.

[64] K. Tvingstedt *et al.*, “Radiative efficiency of lead iodide based perovskite solar cells,” *Sci. Rep.*, vol. 4, no. 1, p. 6071, 2015.

[65] G. J. A. H. Wetzelaer *et al.*, “Trap-Assisted Non-Radiative Recombination in Organic-Inorganic Perovskite Solar Cells,” *Adv. Mater.*, vol. 27, no. 11, pp. 1837–1841, 2015.

[66] W. Nie *et al.*, “High-efficiency solution-processed perovskite solar cells with millimeter-scale grains,” *Science (80-.).*, vol. 347, no. 6221, pp. 522–525, 2015.

[67] K. Domanski *et al.*, “Migration of cations induces reversible performance losses over day/night cycling in perovskite solar cells,” *Energy Environ. Sci.*, vol. 10, no. 2, pp. 604–613, 2017.

[68] J.-P. Correa-Baena *et al.*, “The rapid evolution of highly efficient perovskite solar cells,” *Energy Environ. Sci.*, vol. 10, pp. 710–727, 2017.

[69] B. Wu *et al.*, “Charge Accumulation and Hysteresis in Perovskite-Based Solar Cells: An Electro-Optical Analysis,” *Adv. Energy Mater.*, vol. 5, no. 19, pp. 1–8, 2015.

[70] H. S. Kim and N. G. Park, “Parameters affecting I-V hysteresis of CH3NH3PbI3 perovskite solar cells: Effects of perovskite crystal size and mesoporous TiO2 layer,” *J. Phys. Chem. Lett.*, vol. 5, no. 17, pp. 2927–2934, 2014.

[71] X. Yin *et al.*, “Highly Efficient Flexible Perovskite Solar Cells Using Solution-

Derived NiOx Hole Contacts," *ACS Nano*, vol. 10, no. 3, pp. 3630–3636, 2016.

[72] Y. Wang *et al.*, "Air-stable, hole-conductor-free high photocurrent perovskite solar cells with CH3NH3PbI3-NiO nanoparticles composite," *Nano Energy*, vol. 27, pp. 535–544, 2016.

[73] G. Grancini *et al.*, "One-Year stable perovskite solar cells by 2D / 3D interface engineering," *Nat. Commun.* |, no. 14, pp. 1–8, 2017.

[74] C. C. Stoumpos *et al.*, "Ruddlesden-Popper Hybrid Lead Iodide Perovskite 2D Homologous Semiconductors," *Chem. Mater.*, vol. 28, no. 8, pp. 2852–2867, 2016.

[75] A. M. a. Leguy *et al.*, "Reversible Hydration of CH3NH3PbI3 in Films, Single Crystals, and Solar Cells," *Chem. Mater.*, vol. 27, no. 9, pp. 3397–3407, 2015.

[76] A. Dualeh *et al.*, "Thermal behavior of methylammonium lead-trihalide perovskite photovoltaic light harvesters," *Chem. Mater.*, vol. 26, no. 21, pp. 6160–6164, 2014.

[77] Q. Wang *et al.*, "Scaling behavior of moisture-induced grain degradation in polycrystalline hybrid perovskite thin films," *Energy Environ. Sci.*, vol. 10, no. 2, pp. 516–522, 2017.

[78] A. J. Pearson *et al.*, "Oxygen Degradation in Mesoporous Al2O3/CH3NH3PbI3-xClx Perovskite Solar Cells: Kinetics and Mechanisms," *Adv. Energy Mater.*, pp. 1–10, 2016.

[79] D. Bryant *et al.*, "Light and oxygen induced degradation limits the operational stability of methylammonium lead triiodide perovskite solar cells," *Energy Environ. Sci.*, vol. 9, p. DOI: 10.1039/c6ee00409a, 2016.

[80] N. Aristidou *et al.*, "The Role of Oxygen in the Degradation of Methylammonium Lead Trihalide Perovskite Photoactive Layers," *Angew. Chemie - Int. Ed.*, vol. 54,

no. 28, pp. 8208–8212, 2015.

[81] B. Li *et al.*, “Advancements in stability of perovskite solar cells: degradation mechanisms and improvement approaches,” *RSC Adv.*, vol. 6, pp. 38079–38091, 2016.

[82] T. Kamiya *et al.*, “Origins of High Mobility and Low Operation Voltage of Amorphous Oxide TFTs: Electronic Structure, Electron Transport, Defects and Doping*,” *J. Disp. Technol.*, vol. 5, no. 12, pp. 468–483, Dec. 2009.

[83] J.-Y. Huh *et al.*, “Light and bias stability of a-IGZO TFT fabricated by r.f. magnetron sputtering,” *Curr. Appl. Phys.*, vol. 11, no. 5, pp. S49–S53, Sep. 2011.

[84] ITworld staff, “Mitsubishi unveils DRAM in thin film transistor LCD.”

[85] Jyi-Tsong Lin *et al.*, “Performances of a Capacitorless 1T-DRAM Using Polycrystalline Silicon Thin-Film Transistors With Trenched Body,” *IEEE Electron Device Lett.*, vol. 29, no. 11, pp. 1222–1225, Nov. 2008.

[86] C. J. Wan *et al.*, “Short-Term Synaptic Plasticity Regulation in Solution-Gated Indium–Gallium–Zinc-Oxide Electric-Double-Layer Transistors,” *ACS Appl. Mater. Interfaces*, vol. 8, no. 15, pp. 9762–9768, Apr. 2016.

[87] E. Lai *et al.*, “A Multi-Layer Stackable Thin-Film Transistor (TFT) NAND-Type Flash Memory,” in *2006 International Electron Devices Meeting*, 2006, no. c, pp. 1–4.

[88] J. Kim *et al.*, “Initial synaptic weight distribution for fast learning speed and high recognition rate in STDP-based spiking neural network,” *Solid. State. Electron.*, vol. 165, no. December 2019, p. 107742, Mar. 2020.

[89] S. Cosemans *et al.*, “Towards 10000TOPS/W DNN Inference with Analog in-

Memory Computing – A Circuit Blueprint, Device Options and Requirements," in *2019 IEEE International Electron Devices Meeting (IEDM)*, 2019, vol. 2019-Decem, pp. 22.2.1-22.2.4.

[90] H. Zhang *et al.*, "Pinhole-Free and Surface-Nanostructured NiOx Film by Room-Temperature Solution Process for High-Performance Flexible Perovskite Solar Cells with Good Stability and Reproducibility," *ACS Nano*, p. acsnano.5b07043, 2015.

[91] R. Zhang *et al.*, "Continuous Size Tuning of Monodispersed ZnO Nanoparticles and Its Size Effect on the Performance of Perovskite Solar Cells," *ACS Appl. Mater. Interfaces*, vol. 9, no. 11, pp. 9785–9794, Mar. 2017.

[92] S. Park *et al.*, "Alignment of Cascaded Band-Gap via PCBM/ZnO Hybrid Interlayers for Efficient Perovskite Photovoltaic Cells," *Macromol. Res.*, vol. 26, no. 5, pp. 472–476, May 2018.

[93] J. Duan *et al.*, "ZnO nanostructures for efficient perovskite solar cells," *J. Mater. Sci. Mater. Electron.*, vol. 28, no. 1, pp. 60–66, Jan. 2017.

[94] A. N. Corpus-Mendoza *et al.*, "Design of Schottky Contacts for Optimum Performance of Thin-Film Silicon Solar Cells," *IEEE J. Photovoltaics*, vol. 5, no. 1, pp. 22–27, Jan. 2015.

[95] T. Kamiya *et al.*, "Present status of amorphous In–Ga–Zn–O thin-film transistors," *Sci. Technol. Adv. Mater.*, vol. 11, no. 4, p. 044305, Feb. 2010.

[96] S.-H. K. Park *et al.*, "Light Response of Top Gate InGaZnO Thin Film Transistor," *Jpn. J. Appl. Phys.*, vol. 50, no. 3, p. 03CB08, Mar. 2011.

[97] A. Pandey *et al.*, "Density functional theory model of amorphous zinc oxide (a-ZnO) and a-X0.375Z0.625O (X= Al, Ga and In)," *J. Non. Cryst. Solids*, vol. 455, pp. 98–

101, Jan. 2017.

[98] H. Hosono *et al.*, “Novel oxide amorphous semiconductors: transparent conducting amorphous oxides,” *J. Non. Cryst. Solids*, vol. 203, pp. 334–344, Aug. 1996.

[99] K. Nomura *et al.*, “Room-temperature fabrication of transparent flexible thin-film transistors using amorphous oxide semiconductors,” *Nature*, vol. 432, no. 7016, pp. 488–492, Nov. 2004.

[100] S. Lee *et al.*, “Trap-limited and percolation conduction mechanisms in amorphous oxide semiconductor thin film transistors,” *Appl. Phys. Lett.*, vol. 98, no. 20, pp. 2011–2014, 2011.

[101] K. Ide *et al.*, “Electronic Defects in Amorphous Oxide Semiconductors : A Review,” vol. 1800372, pp. 1–28, 2019.

[102] R. Martins *et al.*, “Transport in high mobility amorphous wide band gap indium zinc oxide films,” *Phys. status solidi*, vol. 202, no. 9, pp. R95–R97, Jul. 2005.

[103] R. H. Bube, “Trap Density Determination by Space-Charge-Limited Currents,” *J. Appl. Phys.*, vol. 33, no. 5, pp. 1733–1737, 1962.

[104] S. Kirkpatrick, “Percolation and Conduction,” *Rev. Mod. Phys.*, vol. 45, no. 4, pp. 574–588, 1973.

[105] T. Kamiya *et al.*, “Origin of definite Hall voltage and positive slope in mobility-donor density relation in disordered oxide semiconductors,” *Appl. Phys. Lett.*, vol. 96, no. 12, pp. 1–4, 2010.

[106] S. Lee *et al.*, “Trap-limited and percolation conduction mechanisms in amorphous oxide semiconductor thin film transistors,” *Appl. Phys. Lett.*, vol. 98, no. 20, p. 203508, May 2011.

[107] M. C. Hamilton *et al.*, "Field-Effect Mobility of Organic Polymer Thin-Film Transistors," *Chem. Mater.*, vol. 16, no. 23, pp. 4699–4704, Nov. 2004.

CHAPTER 3

Artificial Neural Networks and Neuromorphic Computing

3.1 Brain-inspired computing

The human brain is an intricate information-processing system comprising an expansive network of neurons, approximately 10^{11}, and a significantly larger number of synapses, approximately 10^{15}. These neurons and synapses function via an analog mechanism, processing information through current spikes mediated by changes in ionic concentrations and neurotransmitters [1].

Biological neurons, as the principal components of the nervous system, are tasked with the reception, processing, and transmission of information via electrical and chemical signals. A neuron is typically constituted of a cell body to which dendrites and an axon are attached. The dendrites, characterized by their tree-like extensions emanating from the cell body, facilitate signal reception from other neurons, functioning as the neuron's input system. Upon signal reception at a receptor, the neuron integrates these incoming signals, transitioning into either excitatory or inhibitory states that respectively increase or decrease the probability of the neuron generating a spike signal. When the aggregate sum of excitatory signals surpasses a specified threshold, an action potential (or spike) is produced and transmitted to the axon—a lengthy projection that enables signal transfer between neurons via synapses.

As depicted in Fig 3.1, synapses represent microscale junctions connecting one neuron's axon and the subsequent neuron's dendrites, thereby enabling information transference from one neuron to another. This transmission process at a synapse is chemically mediated; the presynaptic neuron, triggered by the action potential, releases

neurotransmitters that bind to receptors on the postsynaptic neuron, modulating its electrical potential.

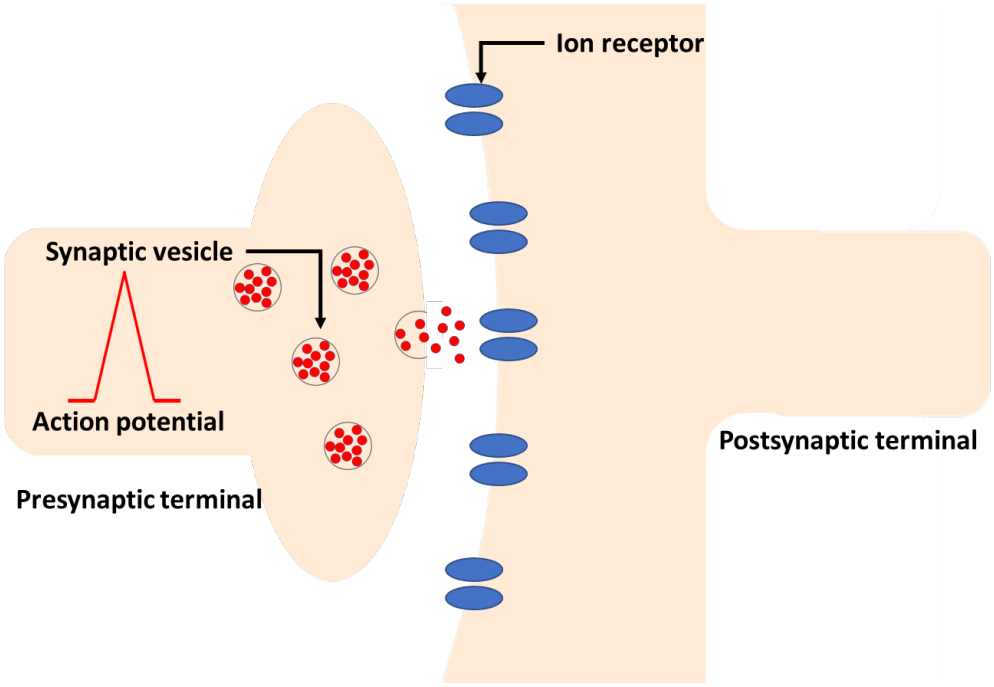


Fig 3.1 Schematic diagram of a chemical synapse transmission [2].

Synapses serve as the foundational pillars to human cognition, including learning and memory functions, attributed to their synaptic plasticity - the capacity to strengthen or weaken connections in response to repetitive neuronal activity [3]. When a presynaptic neuron communicates with a postsynaptic receptor, it triggers the release of ions, which in turn, generates a brief electrical current pulse, the Excitatory Post-Synaptic Current (EPSC), directed towards the cell body (Fig 3.2 (a)). This enhances the likelihood of the neuron to fire. Through repeated activity, synapses can either be strengthened or weakened, a mechanism known as synaptic plasticity, which is believed to be pivotal for the learning and memory processes of the brain [4].

Synaptic plasticity is broadly categorized into two types: short-term plasticity, spanning from milliseconds to minutes, and long-term plasticity, enduring from hours to even decades [5][6]. In the realm of neuromorphic computing and artificial synapses, the term EPSC is employed to denote an analogous concept. Rather than biological neurotransmitters and ion channels, these systems utilize electronic signals and components to simulate the behavior of biological synapses. Thus, in this context, an EPSC symbolizes the response of an artificial synapse to an "excitatory" input signal, and synaptic plasticity is reflected by the ability of the device to alter its conductivity in response to repeated input.

Another important synaptic learning rule is spike timing-dependent plasticity (STDP), where synaptic plasticity could enter a state of potentiation (strengthening) or depression (weakening) depending on the temporal order and the time difference between pre-synaptic and post-synaptic spikes (Fig 3.2 (b)) [7]. This forms an essential part of the adaptability and learning capacity of biological neural networks. In neuromorphic computing, this is often demonstrated via the device's ability to change its conductance state by adjusting the time difference between two input.

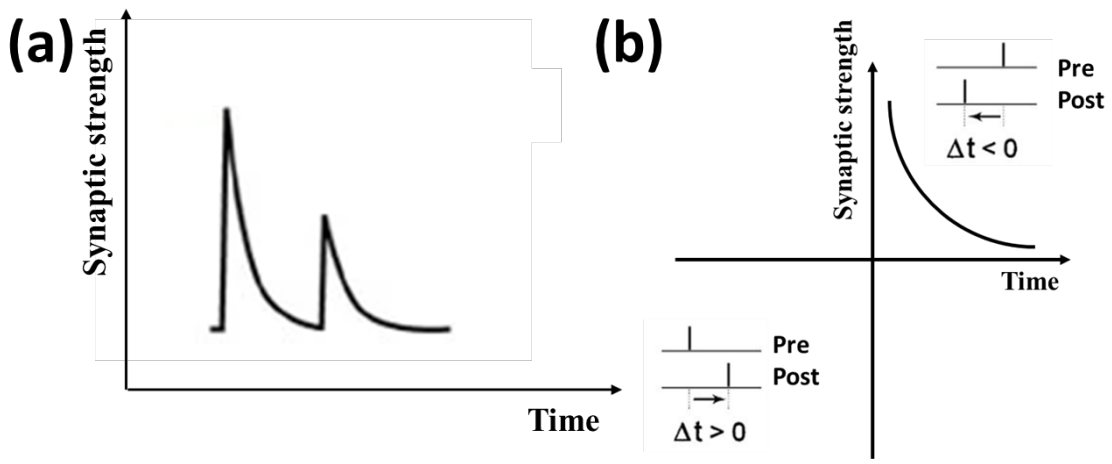


Fig 3.2 Schematic diagram of (a) excitatory post synaptic current (EPSC) and (b) spike timing-dependent plasticity (STDP).

3.2 Artificial Neural Networks

Artificial Neural Networks (ANNs) are computational structures derived from interconnected nodes that process information similarly to neurons in a biological nervous system, albeit in a simplified manner.

Learning or training within artificial neural networks entails adjusting the weights of connections between the neurons in a given network. The weights, (in physical synaptic devices usually represented by the device conductance), serve to strengthen and reduce connections between each layer of neurons, and play a crucial role in realizing the network's intended function. These weights signify the extent of influence one neuron exerts over another, comparable to the synaptic connection strength in a biological neural network. During the learning phase, these weights are modified to enable the ANN to achieve the desired output or function.

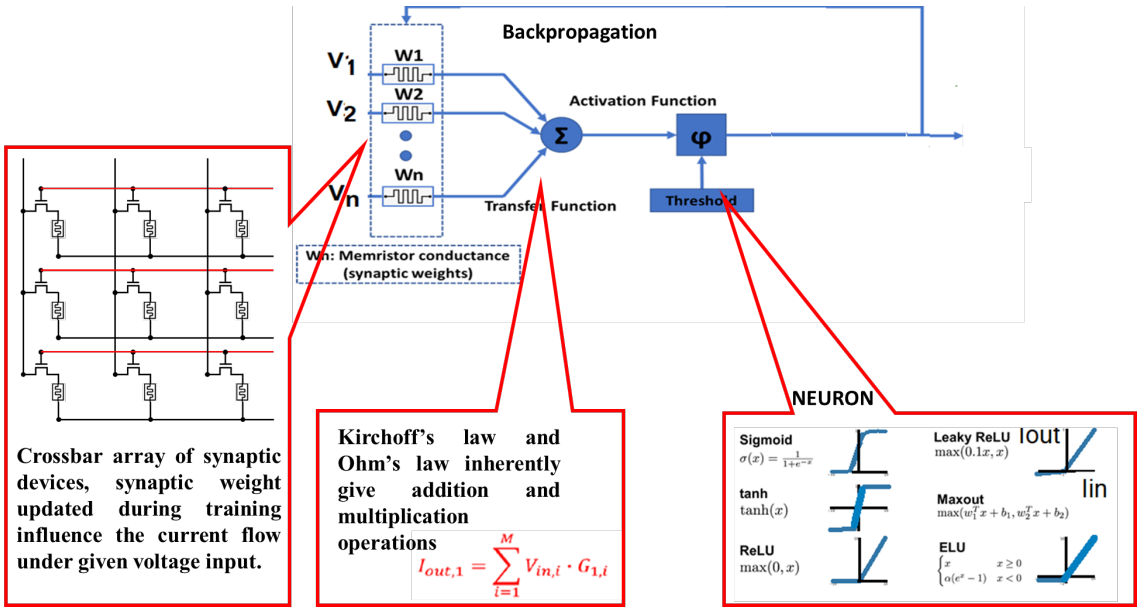


Fig 3.3 Schematic diagram of hardware artificial neural network

Typically, there are two primary learning methodologies within Artificial Neural Networks: supervised learning and unsupervised learning. Supervised learning refers to scenarios where the network is trained on a labelled dataset. The network leverages this training dataset to learn a function that accurately maps inputs to their corresponding outputs. Conversely, unsupervised learning refers to scenarios where the learning operates without the assistance of labelled data.

3.2.1 Feed-forward Neural Network

The Feed-forward Neural Network (FNN) is the most fundamental form of an artificial neural network. This network comprises layers of nodes, specifically input, output, and an arbitrary number of hidden layers, arranged in sequential order. The term "feedforward" signifies that the information flow within these networks is unidirectional—

inputs are introduced through the input layers, processed sequentially through the hidden layers, and the results are produced at the output layers.

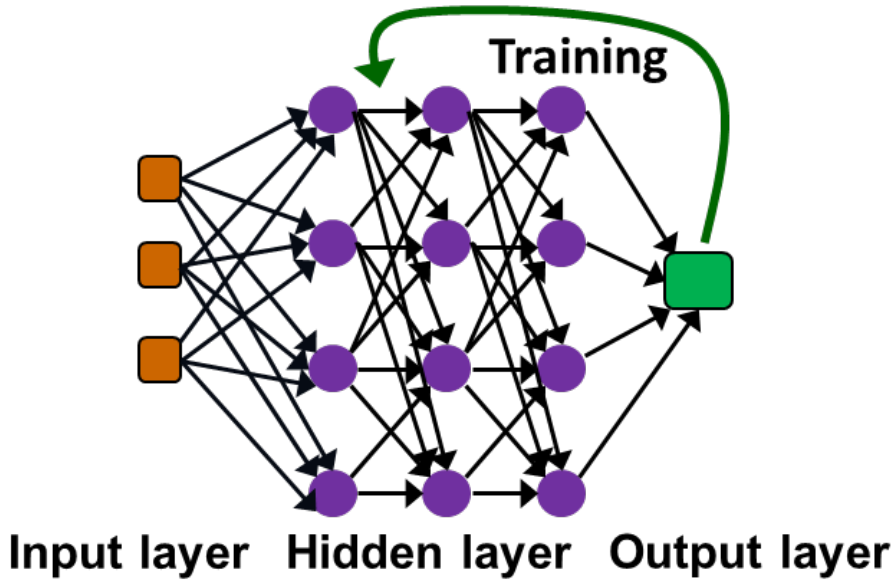


Fig 3.4 Schematic diagram of a Feedforward Neural Network

In an FNN, each neuron is connected to and receives inputs or from all neurons in the previous layer. The weights associated with each connection determine the impact of these inputs. The output of each neuron is then calculated based on the sum of all inputs from the interconnected neurons in the previous layer and their corresponding weights. Mathematically, this is represented as $y_i = f_a\left(\sum_j w_i x_j + b\right)$, where, x_j and w_i denote the inputs and the associated weights, respectively. Here, f_a signifies an activation function, typically a non-linear function such as Sigmoid [8] , Hyperbolic Tangent [8] or Rectified Linear Unit (ReLU) function [9]. The term b, referred to as the bias, is an adjustable parameter employed during network training.

3.2.2 Recurrent Neural Networks

Recurrent Neural Networks (RNNs) represent an advanced class of artificial neural networks designed to recognize and process sequential or time related data. Unlike FNN, which each process inputs independently, RNNs are capable of handling temporal dependencies by retaining information from previous inputs [10][11]. This is made possible by the network's unique architecture, which includes loops or cycles, enabling information to flow from one step in the sequence to the next.

Similar to FNN, the RNN also contains three layers: an input layer, a hidden layer, and an output layer. But unlike FNN, the hidden layer in the RNN is called "recurrent" layer. Different from the hidden layers in FNN, the neurons in the recurrent layer are not strictly separate into clearly defined hidden layers. Instead, the recurrent layer poses cyclic connection, allowing information to persist and be passed to next neurons. This structure of the network creates temporal dimension that allows sequential order or timestep of the input to be recorded and processed within the network.

This characteristic makes RNNs particularly effective for tasks where the temporal dynamics of the input data plays a crucial role, such as in language processing [12][13], speech recognition [14][15], and time series forecasting [16][17]. However, despite their powerful capabilities, training RNNs can be challenging due to issues like vanishing or exploding gradients that can lead to unstable learning behavior [18][19].

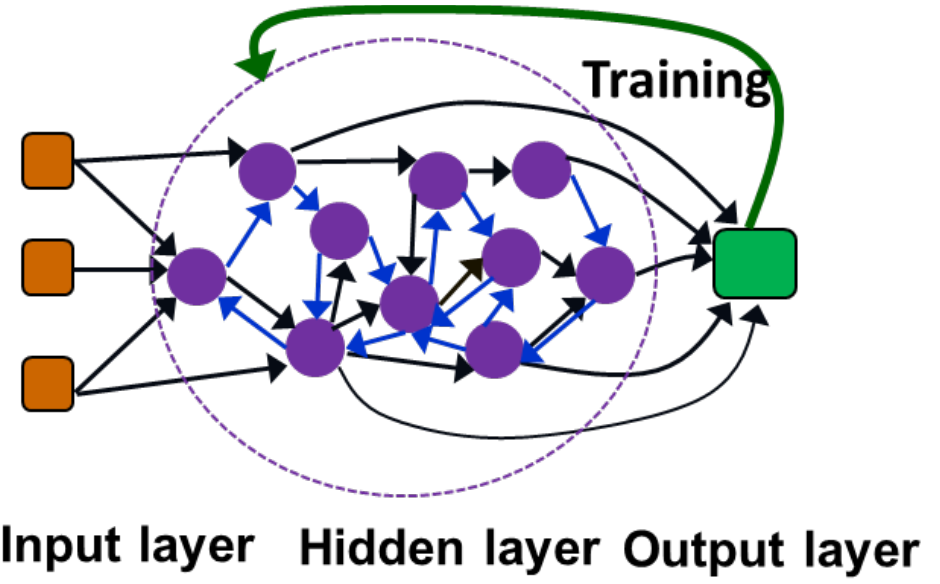


Fig 3.5 Schematic diagram of recurrent neural networks

3.2.3 Backpropagation

Backpropagation stands as the most fundamental algorithm employed in the training of neural networks, particularly in Feed-forward Neural Networks (FNNs) [20]. The training process predominantly involves two stages: a forward pass and a backward pass.

During the forward pass (Fig 3.6 (a)), an input from a labeled dataset is introduced into the network, prompting the network to generate a prediction output marked as o. This prediction is subsequently compared with the actual output from the labeled dataset y to compute the mean square of the error following: $Error(E) = (y - o)^2$. Following this, the system enters the backward pass stage. Here, the calculated error is propagated backwards through the network (Fig 3.6 (b)), starting from the output layer and moving towards the input layer.

In this process, the contribution of each weight to the overall error is determined as the gradient of error is calculated with respect to weight in a particular layer. Based on this information, the network updates the weights and biases in a manner that minimizes the error. This is typically achieved using the gradient descent algorithm or its variants, which iteratively adjust the parameters in the direction of steepest decrease of the error function [21]. In essence, backpropagation employs the principles of differential calculus to optimize the performance of the neural network by minimizing the discrepancy between the network's predicted outputs and the actual outputs.

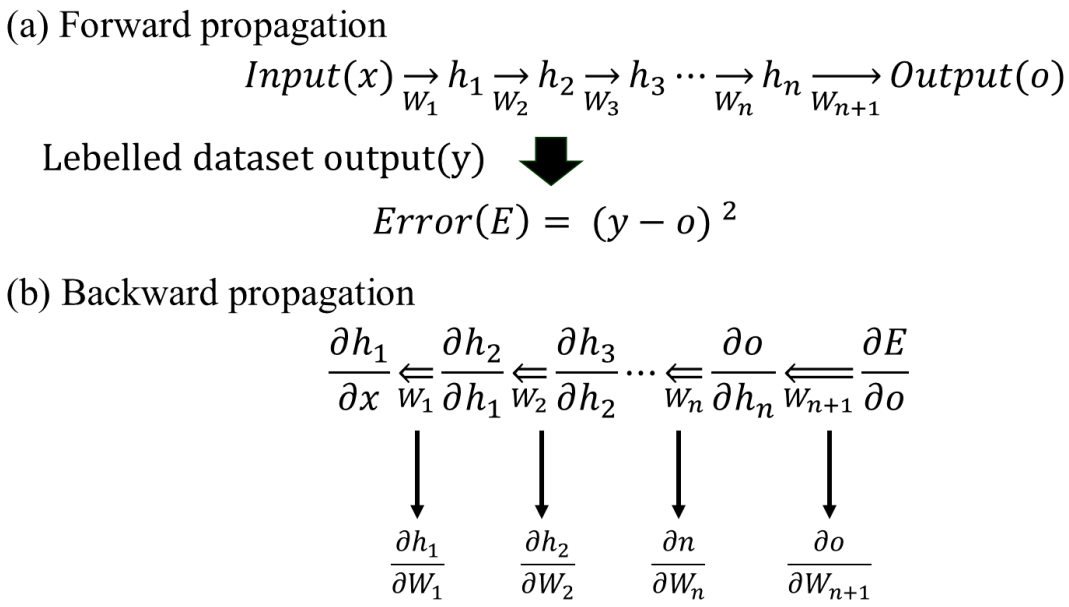


Fig 3.6 Schematic diagram of the forward propagation (a) of the input and backward propagation (b) of the error.

In the context of RNNs, a derivative of backpropagation, termed as backpropagation-through-time (BPTT), is frequently employed [21]. BPTT broadly follows the schematic of standard backpropagation but with a few additional steps. During

RNN training, the intermediate states generated by the recurrent network must be recorded. BPTT entails unfolding the network to trace back to each timestep and its corresponding intermediate states within the RNN, treating each timestep akin to a separate layer within the FNN. Thereafter, the training process aligns with standard backpropagation, albeit with a considerably larger number of layers stemming from intermediate states at each timestep. This training necessitates modifications to all components of the network, including input-to-RNN, RNN-internal, and RNN-to-output weights, to achieve the network's desired functionality [22].

However, this training process is computationally expensive and requires substantial memory for storing the intermediate states at each timestep during the forward pass. Given the structure of the network and the potential for numerous network layers to be involved in the process, optimizing weight updates can become a challenging task. This difficulty often results in vanishing or exploding gradients, contributing to a slow and difficult training process [23]. Another significant challenge posed by conventional RNNs is their struggle with long-term dependencies. Here, information from earlier steps in the sequence is crucial for understanding the context at a later point. Yet, these networks often falter in effectively capturing and leveraging this temporal information [24]. Consequently, designing and implementing effective RNNs requires careful consideration and advanced techniques to handle these challenges.

3.2.4 Reservoir computing

To address these issues, a new framework called Reservoir computing (RC) was proposed in the early 2000s [25]. The concept of Reservoir computing can be traced back to two independently proposed concept called Echo State Networks (ESNs) [26] and

Liquid State Machines (LSMs) [27]. The term "Reservoir Computing" came into existence later to encapsulate these similar concepts under a common umbrella. The RC framework is a type of RNNs which replaces the recurrent neural networks with an internal coupled nonlinear transfer function called reservoir (also known as known as the "echo state property" or "liquid state property"). And the reservoir is a high-dimensional dynamical system which memorizes and processes the temporal dynamics of the input data.

The RC framework is a special type of RNN with a distinctive training method. Unlike the standard RNNs, in an RC framework, the training happens exclusively in the readout function [28]. This significantly reduces the complexity of the training process as the large numbers of recurrent networks no longer need to be adjusted. When a set of input is fed into the reservoir, the reservoir passively excited by the input signal, transforms the input into a high-dimensional representation, capturing the temporal dynamics of the input. The states of the reservoir nodes, at each time step, are collected as the reservoir's response to the input data and fed into the readout function. The training of the networks focuses on adjusting the RNN-to-output layer to achieve the desired output.

Aiming to further develop a more simplified and energy-efficient RC network, a new framework called physical reservoir has been proposed based on the concept of the reservoir computing [29], where instead of using a software-based reservoir network, other nonlinear dynamical systems such as oscillators [30], photonic devices [31][32] or even physical phenomenon such as Quantum Dynamics [33] can be used as reservoir. In recent years, neuromorphic hardware such as memristors [34] has emerged as a preferred choice for implementing physical reservoirs. This development aims to exploit the rich dynamics

of physical systems for computational tasks, potentially leading to more energy-efficient and powerful computing architectures.

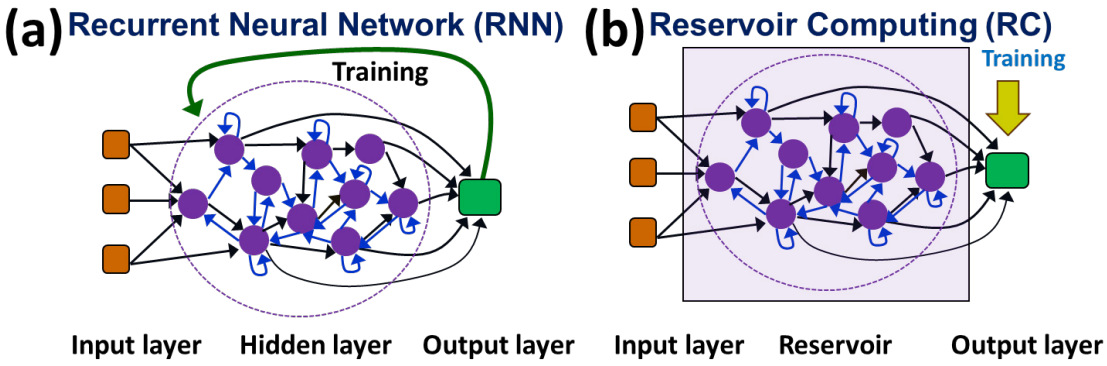


Fig 3.7 (a) Traditional RNN training methods update all connection weights (red) during the training process. (b) In Reservoir Computing, only the activation function (Reservoir-to-output weights) are updated.

The concept of employing memristors in Reservoir Computing (RC) frameworks was first proposed in 2012 [35]. These early memristor-based RC frameworks focused on harnessing the inherent nonlinearity and memory capabilities of the memristors. These reservoirs often utilized standard non-volatile memristors and required multiple memristors to achieve the desired time-dependency and high variability in states.

The structure of these memristor-based reservoirs often mirrored that of Recurrent Neural Networks (RNNs), with randomly interconnected memristors or memristor crossbar arrays forming the hidden layers [34][36]. In some cases, memristor-based circuits, namely inhibitory and excitatory memristive synapse circuits, were employed to fulfill the reservoir function [37][38]. The accuracy of these types of networks reportedly ranged between 61.7% and 85% [36][37], varying based on the tasks undertaken and the specific testing conditions. While these results were promising, the complexity of these memristor-

based RC frameworks prompted further research and development efforts to refine and optimize the architecture.

One strategy to streamline physical reservoirs involves the use of volatile memristive devices [39]. Unlike their non-volatile counterparts, volatile memristive devices can achieve the required nonlinearity and time-dependency via inherent memory decay or fading memory [39]. In 2014, Bürger *et al.* proposed a volatile memristor-based network for image classification. This network structure was akin to reservoir computing, though multi-layer perceptron (MLP) and back-propagation were still necessary. The network was built upon a 28x28 non-volatile memristor crossbar to process input from 784-pixel images. A maximum classification rate of 89.2% was achieved using the memristor crossbars alongside a 100 × 50 × 10 MLP [39].

In 2014, Chao *et al.* reported a hardware reservoir computing network for image classification using WOx-based volatile memristors [34]. The network was based on a 32x32 memristor crossbar array, with each memristor acting as a reservoir operating in parallel with the others. The input was converted into sequential pulses and fed into the reservoir. The memristor based reservoir take a reading after few inputs, effectively down-sampling the input and take a nonlinear transfer at same time [34]. By feeding the input both horizontally and vertically, the network achieved a maximum classification accuracy of 92.1%. In 2019, Midya *et al.* significantly advanced the field of hardware-based Reservoir Computing networks by introducing a volatile memristor-based network in which the readout layer function was also implemented using a memristor crossbar array [40]. This development marked a departure from traditional RC systems where the readout layer was usually implemented in software, thereby demonstrating the potential for fully

hardware-based Reservoir Computing systems. The system developed by Midya et al. achieved a maximum accuracy of 83%, highlighting the promising performance of hardware-implemented Reservoir Computing architectures.

3.2.5 Spiking Neural Networks

Spiking Neural Networks (SNNs) are a new generation of neural network models that more closely emulate biological neural systems. The structure of SNNs largely resembles previous generations of neural networks. However, the primary difference between traditional feed-forward neural networks (FNNs) and SNNs lies in the neuron model.

In FNNs, neurons are continuous mathematical activation functions that generate outputs from inputs continuously. In contrast, SNNs use neuron models that more closely mimic biological neurons, such as the Leaky Integrate-and-Fire (LIF) model [41]. This model can capture the fundamental dynamics of the action potential of biological neurons. In this case, the neuron integrates the input, adjusting its internal state. The membrane voltage of the LIF model is decreased according to the following equation [42]:

$$\tau_m \frac{dv(t)}{dt} = -[v(t) - V_0] + R_m I_s(t) \quad \text{(Eq 3.1)}$$

Here, $v(t)$ is the neuron's membrane potential. τ_m is the passive membrane time constant, which is related to the membrane capacitance C_m and the leak resistance R_m. V_0 is the resting potential. Once the membrane potential $v(t)$ reaches or exceeds a set threshold, the neuron fires the spike to communicate with other neurons and the membrane potential is reset to V_0.

Another key feature of SNNs is the introduction of delay within the network. In biological neurons, the synapses connect to the neuron through long tubular structures called axons. Due to the time required to pass signals through the synapse and the axons, the delay in biological neurons is usually in the tens to hundreds of milliseconds [43]. This introduced delay is crucial for the network to process the temporal information of the input and also plays a role in implementing learning rules such as STDP.

However, implementing this delay in hardware can be extremely challenging. One of the key challenges is the timescale of the delay. Modern CMOS and neuromorphic devices often operate in the nanosecond range. But the delay in SNNs not only needs to achieve a timescale in the tens to hundreds of milliseconds range [44], but it also requires controlled variability as not all neurons share the same delay [45]. To achieve the desired delay function for SNNs, external circuits are often used to introduce delay. For example, more recently, successful demonstrations of ReRAM-based Spiking Neural Networks (SNNs) have been made, wherein the ReRAM is coupled with a capacitor to achieve a desirable delay in the tens to hundreds of milliseconds range [44]. This adds a layer of complexity to the network.

3.3 Emerging memory devices for neuromorphic application.

Motivated by the ambitious objective of constructing computational systems that parallel the energy efficiency of the human brain, research has been intensively directed towards neuromorphic hardware in the recent years. This branch of study focuses on developing devices that can accurately mimic the behavior and functionality of neurons and synapses, reflecting the intricacies of biological information processing systems. This fundamental difference in operation principles results in the need for a disproportionately

large number of CMOS devices and computational power to simulate even the most basic level of neuronal and synaptic interactions [46]. Another major challenge is the high power consumption of modern CMOS technology. While the human brain operates on roughly 20W of power, supercomputers require tens of MW of power to mimic only a fraction of human brain function [47][48]. To overcome these limitations, there has been a growing interest in the development of hardware that can assist in the implementation of brain-inspired computing, specifically focusing on artificial synaptic devices [3].

Given that synapses are the neuronal connections responsible for memory and learning, the design of physical synaptic devices is often based on memory devices that can emulate biological characteristics such as synaptic plasticity [2][49]. In electronic synaptic devices, characteristics such as the devices' on/off ratio, the number of memory states, analog behavior, and the ability to perform an excitatory post-synaptic current (EPSC) upon receiving excitation pulses are typically examined.

Non-volatile memory devices are usually preferred for better implementation of long-term plasticity due to their ability to retain information even when power is disconnected [50][51]. However, volatile memories, which lose their stored data when power is turned off, can be used to process temporal input [34] and implement short-term plasticity [40][52].

The ability to perform STDP is considered a desirable feature for physical synaptic devices [49]. However, its actual role and functional consequences in neuromorphic systems are yet to be fully understood [7].

Apart from the synaptic characteristics, other factors such as power consumption, switching speed, cost, and CMOS compatibility should also be taken into consideration

when benchmarking different devices for use in neuromorphic systems. These criteria help ensure that the selected devices will not only function effectively as artificial synapses but also be feasible for integration into larger systems and practical applications.

Among the assortment of emerging neuromorphic devices, memristors are viewed as one of the most promising prospects for next-generation computing devices. They have recently garnered renewed interest due to advancements in the area of artificial synapses. When contrasted with other potential candidates for physical synapses, memristors exhibit several compelling attributes such as scalability and high-density integration [53], cost-effective fabrication [54], and compatibility with CMOS technology [55]. Significantly, there have been successful instances of Excitatory Postsynaptic Current (EPSC) and Spike-Timing Dependent Plasticity (STDP) demonstrated in memristor-based networks [49][56].

Memristors, the fourth fundamental circuit element, form a distinct category within resistive memory devices, adding to the existing triad of resistors, capacitors, and inductors [57]. They are characterized by their resistive switching behavior, a unique feature that allows the device to 'remember' its historical input data based on the application of voltage or current. This trait enables the memristor to exhibit two or more distinguishable resistance states, often referred to as the High Resistance State (HRS) and Low Resistance State (LRS). When the applied voltage or current surpasses a particular threshold, the memristor can transition between its HRS and LRS. Consequently, the current output at a fixed 'Read' voltage provides an indication of past inputs. However, if the device operates within the voltage threshold, the resistance state of the memristor can be read as an output current under a fixed voltage. The types of memristors can vary, primarily dependent on the working mechanism of resistive switching and the material composition of the device.

3.3.1 Phase-change memories

Phase-Change Memories (PCMs) constitute a subset of memristors that operate based on the different conductivity between amorphous and crystalline states of certain materials, commonly GeTe based materials [58][59] . PCMs are predominantly controlled by current pulses. These devices typically feature two electrodes, separated by a layer of phase-change material, with a heater connected to the bottom electrode.

To transition these devices to a Low Resistance State (LRS), a moderate electrical current pulse is applied to the heater. This anneals the phase-change material to a temperature exceeding its crystallization threshold, but remaining below its melting point. A longer pulse is generally required to complete the recrystallization process and establish the device in LRS.

Resetting the device to a High Resistance State (HRS) necessitates a brief, high electrical current pulse to elevate the local temperature above the melting point of the phase-change material. This increased temperature melts and subsequently rapidly quenches the phase-change material, transitioning it to an amorphous phase and resetting the device. Device sensing is accomplished by applying a small current that does not exceed the crystallization temperature, thereby measuring the resistance state of the device.

PCM devices provide non-volatile memory characterized by exceptional switching speed (<10ns) [59], prolonged retention time (over 10 years) [60], and commendable endurance ($> 10^7$) [61][62][63]. In neuromorphic applications, PCM devices can emulate synaptic plasticity during the SET process when the phase-change materials transition from the amorphous to the crystalline phase [51].

However, during the RESET process, the phase-change material undergoes rapid melting and quenching. Achieving refined control over the resistance state changes can prove challenging with a single PCM device. Furthermore, this process often leads to high peak power and energy consumption due to the elevated temperatures (500 to 700 degrees for most phase-change materials) necessary for the melting process [64].

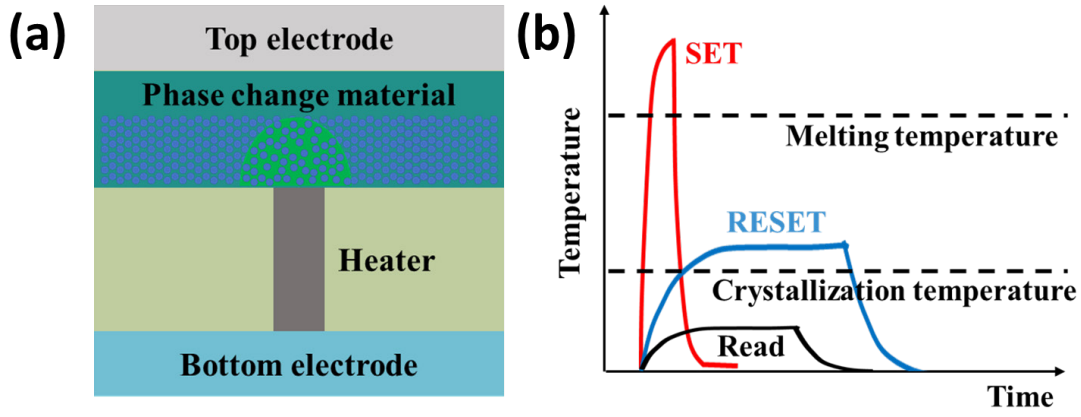


Fig 3.8 (a) Schematic structure diagram and (b) illustration of the SET/RESET process of the Phase-change memories.

3.3.2 Oxide based random-access memories

Oxide-based Random-Access Memories (OxRAM) constitute a large category of memristors. These devices have a simple, capacitor-like structure, featuring two metal electrodes separated by a thin layer of insulating material. The resistive switching behavior of most OxRAMs is based on a filamentary process, where the formation and rupture of conductive filaments drive the resistive switching.

Three different mechanisms have been proposed for the filamentary resistive switching in OxRAMs: the Electrochemical Metallization Mechanism (ECM), the Valence Change Mechanism (VCM), and the Thermochemical Mechanism (TCM). ECM often

occurs in memristors with an electrochemically active metal electrode, known as Conductive Bridge Random-Access Memories (CBRAM). These devices typically use electrochemically active metals like Ag [65], Au [66], or Cu [66] as anodes. Under a positive bias, metal atoms from the active contact oxidize into mobile metal cations within the solid electrolyte thin film. Driven by the applied electric field, these cations drift towards the counter electrode, often an inert metal such as W [67] or Pt [66]. The resistance state changes through the reduction of these metal cations and the formation of metal clusters or a conductive path between the two electrodes.

The VCM operates on the movement of oxygen anions (often described using the movement of oxygen vacancies) in defect-rich transitional metal oxides, such as TaOx [68], TiO_2 [69], and HfOx [70]. The formation of the conductive filament in VCM devices is often related to the movement of oxygen vacancies, a result of redox reactions and the transport of oxygen ions. Both ECM and VCM cells can offer bipolar SET/RESET, and in devices with both transitional metal oxides and an electrochemically active metal contact, both mechanisms can coexist in a single device [71].

The TCM occurs almost exclusively during the RESET process, especially during unipolar reset. Under TCM, the rupture of the conductive filaments can be attributed to thermally induced stoichiometry variations and redox reactions caused by significant joule heat as a high current flows through the narrow part of the conductive filament. However, because of the difficulty in precisely controlling the RESET process and the Joule heat generated, TCM is often not favored in oxide-based memristors.

Oxide-based Random-Access Memories have emerged as a promising candidate for neuromorphic applications, offering high endurance of over 10^{12} [72] and high

switching speeds at sub-ns (0.3nS) [73], reported in multiple instances [74][75]. Successful demonstrations of OxRAM emulating biological synaptic behavior, such as excitatory post synaptic currents (EPSC) [49] and spike-timing-dependent plasticity (STDP) [76], make these devices highly competitive as synapses in neuromorphic hardware.

Despite their rapid development, challenges persist in developing energy-efficient hardware neural networks. The Voltage–Time Dilemma is a key challenge for energy-efficient OxRAM-based neural networks, i.e. maintaining high switching speed and long retention time often requires high operation voltage [77]. This, coupled with thin oxide thickness, often leads to high power density during writing operation. Additionally, the nonlinearity of weight updates in OxRAM-based neural networks poses another problem for synaptic applications. Cycle-to-cycle and device-to-device variations are also major concerns for large-scale OxRAM neural networks.

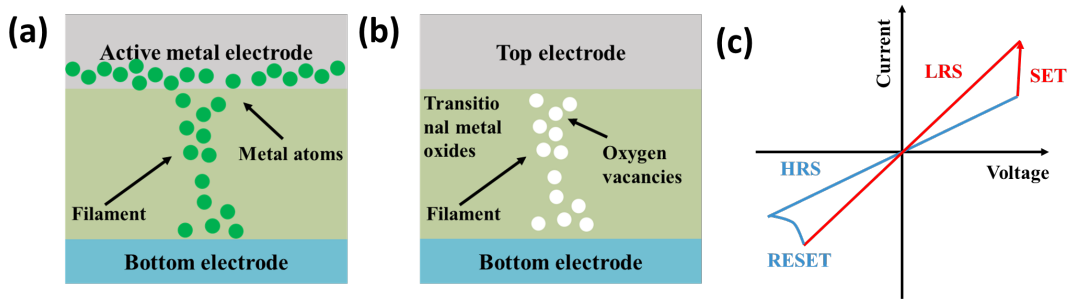


Fig 3.9 Schematic structure diagram of ReRAM based on (a) active metal contact (CBRAM) and (b) transitional metal oxides (RRAM). (c) illustration of the SET/RESET process of the Oxide based random-access memories.

3.3.3 Ferroelectric memories

Ferroelectric RAM (FeRAM), also known as Ferroelectric Capacitor (FeCAP), employs a capacitor-like structure, consisting of two metal electrodes separated by a layer

of ferroelectric (FE) material. The memory function in these devices is achieved by switching electrical dipoles within the FE material under an external bias. However, the read operation of earlier FeRAMs relied on the discharge currents induced by the polarization switching from the memorized state, resulting in a destructive read operation. This complex process, combined with limited CMOS compatibility and scaling issues, hindered the development of capacitor-based FeRAM for many years [42].

Recently, however, with the discovery of ferroelectricity in hafnium oxide [43] and the development of new device structures like Ferroelectric Field Effect Transistors (FeFET), FeRAM has seen renewed interest in the fields of compute-in-memory and neuromorphic applications. Hafnium oxide, a well-established CMOS-compatible material, and the adoption of FeFET have addressed issues related to destructive read operations.

In the context of artificial synapses, successful implementation of EPSC and STDP has been demonstrated, utilizing the multi-domain characteristics of ferroelectric material [44][45]. However, the implementation of FeFET potentiation and depression often necessitates non-identical pulses due to the domain switching mechanism [78]. Consequently, the conductance state of a FeFET is not truly analog, but is limited by multi-domain characteristics.

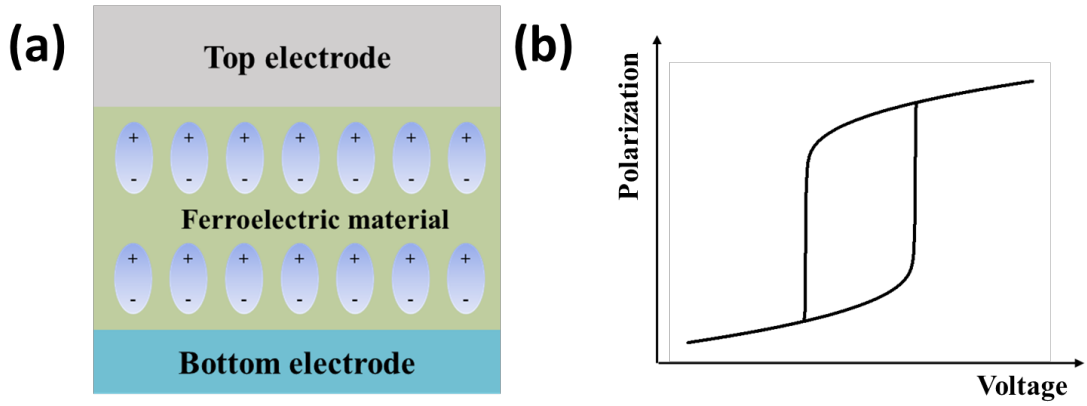


Fig 3.10 (a) Schematic structure diagram of a FeRAM and (b) the polarization charge of the ferroelectric material.

3.3.4 Spin torque magnetic random access memories

Spin-Transfer Torque Magnetic Random Access Memories (STT-MRAM) is one of the more competitive technologies among memristors for the development of neuromorphic hardware. An STT-MRAM consists of a Metal-Insulator-Metal (MIM) structure with two layers of ferromagnetic electrodes separated by a thin layer of insulating material, acting as the tunneling layer. Among these two ferromagnetic layers, there is typically a relatively thick layer where its spin polarization is fixed, known as the pinned layer. There is another thinner ferromagnetic layer where magnetization direction can be altered by an external current, termed the free layer. Depending on the polarization of the free layer, the combined magnetic polarization of the two ferromagnetic layers can switch between parallel (P) and antiparallel (AP) states. As electrons have a higher probability of tunneling through the insulating layer under the parallel state and a lower probability under the antiparallel state, changing the magnetic polarization from an antiparallel (AP) to the parallel (P) state results in a resistance state change from High Resistance State (HRS) to

Low Resistance State (LRS). Commercial development of STT-MRAM technology, one of the more mature among memristor devices, began as early as 2002 [79].

Contrary to other memristor technologies, the resistance switching of STT-MRAM is purely electronic, requiring no atomic level changes. This mechanism results in a fast switching speed, reaching the single-digit nanosecond range (< 3 ns) [73][80][81] with very low power consumption, as low as 6fJ/switch on state-of-the-art STT-MRAM devices [80]. However, compared to other memristors, one of the main drawbacks of STT-MRAM is its relatively poor on/off ratio [82]. This low on/off ratio results in a weak reading signal and limits the device's potential in certain synaptic applications.

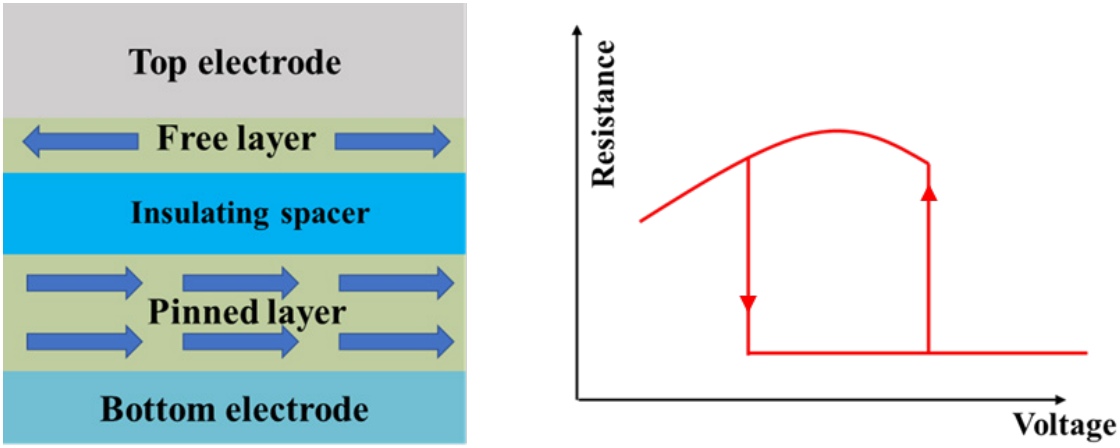


Fig 3.11 (a) Schematic structure diagram of STT-MRAM based on (b) illustration of the R-V curve of the STT-MRAM.

3.3.5 Electrolytes gate transistors

While different types of memristors can mimic biological synaptic plasticity to some degree, the working mechanisms of memristors and biological synapses fundamentally differ. Biological synapses transmit information through the release of chemicals and the diffusion of ions. Given the ionic nature of synapses, diffusive

electrolyte gate transistors might serve as a viable alternative for developing physical synaptic devices and are believed to better mimic biological synaptic plasticity [83]. These devices usually adopt a transistor structure with an ionic electrolyte in the form of a polymer [84], gel [85], or liquid [83] as the gate dielectric. Their working mechanisms differs from case to case but commonly based on redox reactions [83] or double-layer charge accumulation [84] causing change of channel conductivity.

Incorporating a third terminal into a conventional memristor provides benefits of compliance control, instantaneous feedback, and continuous learning, which are not achievable with a two-terminal device. However, the third terminal complicates interconnect and poses integration and footprint challenges, particularly in the context of the crossbar array. Three-terminal TFT memory devices, commonly formulated from ferroelectric materials, are used for storing synaptic weights. Herein, the switching of dipoles between two stable states generates a memory transistor [86]. Beyond ferroelectric materials, oxides such as Ta_2O_5, TiO_2, or Gd_2O_3 can be embedded within the gate insulator of a TFT memory [87]. A cellular neural network employing IGZO thin-film devices capable of simple logic functions has been demonstrated by Mutsumi Kimura et al. [88]. The highest speed Static Random Access Memory (SRAM) in IGZO TFTs, operating at 7V at 140kb/s with only 0.55mW of power consumption, was demonstrated using a diode load by F. De Roose et al. [88].

Liquid gated 3-terminal memory devices have been demonstrated in Tungsten Oxide (WO_3) by Jing-Ting Yang [89], or ZnO TFT with W probed liquid gated TFT [90]. Comparatively, liquid gating is more prevalent in Organic Thin-Film Transistors (OTFTs) for biosensing applications.

In general, TFT memory devices in both AOS and organic materials still have a significant path to traverse in comparison to their CMOS-based counterparts. The structure of conventional electrochemical transistors makes it challenging for them to exhibit non-volatility, making them more suitable for applications where volatility can be leveraged in novel concepts such as reservoir computing.

Thanks to their diffusive mechanism, devices of this nature typically demonstrate true analog behavior with robust excitatory post-synaptic current (EPSC) performance [83]–[85]. Despite these advantages and their true analog nature, these devices face challenges related to scalability and stability. The use of electrolytes, especially in gel [85] or liquid [83] forms, may present a significant issue, though this could potentially be addressed by using solid-state electrolytes such as Ta_2O_5 [84].

Although they present advantages of integration and compatibility with solution sensing and processing, and are sorely needed in Internet of Things (IoT) platforms, they are still in the early stages of development. Their limitations can be categorised into two main topics: (i) device performance and (ii) integration of different functionality. Device performance issues mainly involve stability and scalability [89], while integration poses the challenge of merging sensing, memory, and processing [91].

3.3.6 Perovskite in memory applications

Hybrid perovskite-based resistive random access memory (ReRAM) is an emerging field, and it is rapidly gaining traction due to the unique properties of perovskite materials. These materials demonstrate promising performance in memory devices due to their high on/off ratio (up to 10^6), small operation window (±0.14 V) [92], and potential for flexible, low-cost solution-processed applications [93].

The first perovskite-based ReRAM was reported by E. Yoo et al. in 2015, with a device structure of Au/$CH_3NH_3PbI_3$/FTO [94]. Subsequent demonstrations of long-term memory in $CH_3NH_3PbI_3$-based memristors were achieved with similar structures, such as Ag/$CH_3NH_3PbI_3$/FTO [92] and Au/$CH_3NH_3PbI_3$/ITO [95].

The underlying principles of hybrid perovskite-based ReRAM vary between different devices. Two primary types of resistive switching have been identified: filamentary-type switching and interface-type switching.

Filamentary-type switching, as observed in some hybrid perovskite-based ReRAMs, is thought to function in a way similar to oxide-based ReRAM or conductive-bridging random access memory (CBRAM). Here, the formation and disruption of conductive filaments, which may occur through the migration of metal ions or defects in the halide, cause changes in the device's resistance [95][96].

On the other hand, non-filamentary or interface-type switching operates by ion transfer near the interface between the perovskite layer and a more active electrode such as Ag [80]. It is suggested that the distribution and rotation of the MA^+ within the perovskite structure also contributes to this process.

These findings show promising potential for the further development of perovskite-based memory devices. Further research is needed to fully understand the complex mechanisms underlying their operation, as well as to optimize and enhance their performance [93].

3.3.7 Amorphous Oxide Semiconductors in memristor applications

Amorphous Oxide Semiconductors (AOS) have been increasingly recognized for their potential in memory applications, commonly used in RRAM, PCM, and FeRAM

discussed earlier. For example, control over oxygen vacancies in IGZO can be achieved via manipulation of the In-to-Ga ratio, which has led to the exploration of IGZO for applications in ReRAMs [97]. A Gallium proportion of x=2 has been determined as suitable for resistive switching (RS) in a TiN/Ti/IGZO/Pt structure, with the additional benefit of multilevel capabilities [98][99]. Other techniques can also be employed to influence the properties of IGZO films. For instance, Titanium serves as an oxygen scavenger that assists in increasing oxygen vacancies. A Pt/a-IGZO/TaOx/Al2O3/W/Ti cell showed self-rectification due to resistive switching at the interfacial TaOx layer [100]. Oxygen plasma treatment applied to the bottom electrode in an Al/IGZO/Al structure improved resistive switching [101]. An IGZO ECM cell with a Ag/IGZO/MnO/Pt structure demonstrated a transition from volatile to non-volatile behaviour upon adjusting the current compliance, rendering it appealing for both ReRAM and selector applications [102]. In addition, the integration of a TFT with a ReRAM (1T-1R) using an IGZO/Al2O3 layer has been exhibited [103].

Despite these initial evidences of IGZO-based memory devices, issues affecting the technology include device variability, cycle-to-cycle variability, a limited memory window, and a reduced capability for multilevel logic. More recently, ZnO-based ReRAM has been demonstrated by U. B. Isyaku and S. Member [104]. The advantages of ZnO over IGZO are motivated by the scarcity of materials; the simpler binary compound ZnO is abundant on Earth and has been deposited both in amorphous and nanocrystalline forms [105]. While Kamiya et al. have reported a high concentration of residual free electrons ($>10^{17}$ cm^{-3}) in ZnO due to native defects such as zinc interstitials and oxygen vacancies [106], we have demonstrated on/off ratios of 10^9 in amorphous ZnO TFTs [105].

3.4 Summary

In this chapter, we have reviewed the fundamental principles of brain-inspired computing, which seeks to emulate artificial neural networks. We have introduced the basics of artificial neural networks, ranging from Feed-Forward Neural Networks (FNNs) to Recurrent Neural Networks (RNNs), offering a brief history of the evolution of Reservoir Computing (RC) networks and outlining their primary concepts. These networks leverage synapses and neurons to perform the required functions. We have highlighted the key characteristics necessary for the development of physical synaptic devices and explored the emerging memory devices commonly employed as physical synapses in neuromorphic systems. Owing to their unique switching mechanism and working principle, each of these devices have their own strengths and weaknesses and might be preferred depending on the requirement of the specific application and network. Our understanding of synapses and human learning are still limited and large scale hardware array that could mimic the complex neuron models are still under development. This chapter provide a concise introduction of the physical synaptic devices as a foundation for the subsequent work in neuromorphic applications.

3.5 References

[1] W. Teka *et al.*, "Neuronal Spike Timing Adaptation Described with a Fractional Leaky Integrate-and-Fire Model," *PLoS Comput. Biol.*, vol. 10, no. 3, p. e1003526, Mar. 2014.

[2] A. E. Pereda, "Electrical synapses and their functional interactions with chemical synapses," *Nat. Rev. Neurosci.*, vol. 15, no. 4, pp. 250–263, Apr. 2014.

[3] M. Stampanoni Bassi *et al.*, "Synaptic Plasticity Shapes Brain Connectivity:

Implications for Network Topology," *Int. J. Mol. Sci.*, vol. 20, no. 24, p. 6193, Dec. 2019.

[4] H. D. Mansvelder *et al.*, "Synaptic plasticity in human cortical circuits: cellular mechanisms of learning and memory in the human brain?," *Curr. Opin. Neurobiol.*, vol. 54, pp. 186–193, Feb. 2019.

[5] R. S. Zucker and W. G. Regehr, "Short-Term Synaptic Plasticity," *Annu. Rev. Physiol.*, vol. 64, no. 1, pp. 355–405, Mar. 2002.

[6] M. A. Lynch, "Long-Term Potentiation and Memory," *Physiol. Rev.*, vol. 84, no. 1, pp. 87–136, Jan. 2004.

[7] H. Markram *et al.*, "Spike-Timing-Dependent Plasticity: A Comprehensive Overview," *Front. Synaptic Neurosci.*, vol. 4, no. JULY, pp. 2010–2012, 2012.

[8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[9] G. H. A Krizhevsky *et al.*, "Imagenet classification with deep convolutional neural networks," pp. 1–1432, 2017.

[10] G. Tanaka *et al.*, "Recent advances in physical reservoir computing: A review," *Neural Networks*, vol. 115, pp. 100–123, Jul. 2019.

[11] Z. C. Lipton *et al.*, "A Critical Review of Recurrent Neural Networks for Sequence Learning," pp. 1–38, 2015.

[12] K. M.Tarwani and S. Edem, "Survey on Recurrent Neural Network in Natural Language Processing," *Int. J. Eng. Trends Technol.*, vol. 48, no. 6, pp. 301–304, Jun. 2017.

[13] W. Yin *et al.*, "Comparative Study of CNN and RNN for Natural Language

Processing," 2017.

[14] T. Hori *et al.*, "END-TO-END SPEECH RECOGNITION WITH WORD-BASED RNN LANGUAGE MODELS Mitsubishi Electric Research Laboratories (MERL) Center for Language and Speech Processing , Johns Hopkins University," *2018 IEEE Spok. Lang. Technol. Work.*, pp. 389–396, 2018.

[15] A. Graves *et al.*, "Speech recognition with deep recurrent neural networks," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, vol. 3, no. 3, pp. 6645–6649.

[16] R. Madan and P. S. Mangipudi, "Predicting Computer Network Traffic: A Time Series Forecasting Approach Using DWT, ARIMA and RNN," in *2018 Eleventh International Conference on Contemporary Computing (IC3)*, 2018, no. 1, pp. 1–5.

[17] B. B. Sahoo *et al.*, "Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting," *Acta Geophys.*, vol. 67, no. 5, pp. 1471–1481, Oct. 2019.

[18] P. J. Werbos, "Backpropagation Through Time : What It Does and How to Do It," vol. 78, no. October, pp. 1550–1560, 1990.

[19] K. Nakajima, "Physical reservoir computing — an introductory perspective," 2020.

[20] D. E. Rumelhart *et al.*, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, Oct. 1986.

[21] P. J. Werbos, "Backpropagation through time: what it does and how to do it," *Proc. IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.

[22] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Comput. Sci. Rev.*, vol. 3, no. 3, pp. 127–149, Aug. 2009.

[23] J. F. Kolen and S. C. Kremer, “Gradient Flow in Recurrent Nets: The Difficulty of Learning LongTerm Dependencies,” *A F. Guid. to Dyn. Recurr. Networks*, no. November, 2010.

[24] Y. Bengio *et al.*r, “Learning Long-Term Dependencies with Gradient Descent is Difficult,” vol. 5, no. 2, 1994.

[25] H. Jaeger and H. Haas, “Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication,” *Science (80-.).*, vol. 304, no. 5667, pp. 78–80, Apr. 2004.

[26] H. Jaeger, “The ‘echo state’ approach to analysing and training recurrent neural networks – with an Erratum note1,” *Nurse Educ. Pract.*, vol. 1, no. 4, pp. 221–223, Dec. 2001.

[27] H. M. Natschläger *et al.*, “The" liquid computer": A novel strategy for real-time computing on time series.”

[28] A. A. Ferreira *et al.*, “An approach to reservoir computing design and training,” *Expert Syst. Appl.*, vol. 40, no. 10, pp. 4172–4182, Aug. 2013.

[29] Y. Liu *et al.*, “Online Adaptation and Energy Minimization for Hardware Recurrent Spiking Neural Networks,” *ACM J. Emerg. Technol. Comput. Syst.*, vol. 14, no. 1, pp. 1–21, Mar. 2018.

[30] J. C. Coulombe *et al.*, “Computing with networks of nonlinear mechanical oscillators,” pp. 1–13, 2017.

[31] Q. Vinckier *et al.*, “High-performance photonic reservoir computer based on a coherently driven passive cavity,” *Optica*, vol. 2, no. 5, p. 438, May 2015.

[32] D. Brunner *et al.*, “Parallel photonic information processing at gigabyte per second

data rates using transient states," *Nat. Commun.*, vol. 4, no. 1, p. 1364, Jan. 2013.

[33] K. Fujii and K. Nakajima, "Harnessing Disordered-Ensemble Quantum Dynamics for Machine Learning," *Phys. Rev. Appl.*, vol. 8, no. 2, p. 024030, Aug. 2017.

[34] C. Du *et al.*, "Reservoir computing using dynamic memristors for temporal information processing," *Nat. Commun.*, vol. 8, no. 1, p. 2204, Dec. 2017.

[35] M. S. Kulkarni and C. Teuscher, "Memristor-based reservoir computing," *Proc. 2012 IEEE/ACM Int. Symp. Nanoscale Archit. NANOARCH 2012*, pp. 226–232, 2012.

[36] A. M. Hassan *et al.*, "Hardware implementation of echo state networks using memristor double crossbar arrays," *Proc. Int. Jt. Conf. Neural Networks*, vol. 2017-May, pp. 2171–2177, 2017.

[37] C. Merkel *et al.*, "Memristive Reservoir Computing Architecture for Epileptic Seizure Detection," *Procedia Comput. Sci.*, vol. 41, pp. 249–254, 2014.

[38] C. Donahue *et al.*, "Design and analysis of neuromemristive echo state networks with limited-precision synapses," in *2015 IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA)*, 2015, pp. 1–6.

[39] J. Bürger and C. Teuscher, "Volatile memristive devices as short-term memory in a neuromorphic learning architecture," in *Proceedings of the 2014 IEEE/ACM International Symposium on Nanoscale Architectures*, 2014, pp. 104–109.

[40] R. Midya *et al.*, "Reservoir Computing Using Diffusive Memristors," *Adv. Intell. Syst.*, vol. 1, no. 7, p. 1900084, Nov. 2019.

[41] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Trans. Neural Networks*, vol. 14, no. 6, pp. 1569–1572, Nov. 2003.

[42] A. N. Burkitt, “A Review of the Integrate-and-fire Neuron Model: I. Homogeneous Synaptic Input,” *Biol. Cybern.*, vol. 95, no. 1, pp. 1–19, Jul. 2006.

[43] V. P. Calabrese, “Principles of Neural Science,” *JAMA J. Am. Med. Assoc.*, vol. 268, no. 23, p. 3381, Dec. 1992.

[44] M. Payvand *et al.*, “Dendritic Computation through Exploiting Resistive Memory as both Delays and Weights,” *Proc. Make sure to enter correct Conf. title from your rights confirmation emai (Conference Acron. ’XX)*, vol. 1, no. 1, pp. 1–4, May 2023.

[45] S. Gerber *et al.*, “Neuromorphic implementation of ECG anomaly detection using delay chains,” *BioCAS 2022 - IEEE Biomed. Circuits Syst. Conf. Intell. Biomed. Syst. a Better Futur. Proc.*, vol. 724295, no. 724295, pp. 369–373, 2022.

[46] H. Markram, “Biology---The blue brain project,” in *Proceedings of the 2006 ACM/IEEE conference on Supercomputing - SC ’06*, 2006, vol. 7, no. February, p. 53.

[47] S. Dai *et al.*, “Recent Advances in Transistor-Based Artificial Synapses,” *Adv. Funct. Mater.*, vol. 29, no. 42, p. 1903700, Oct. 2019.

[48] R. Ananthanarayanan *et al.*, “The cat is out of the bag: cortical simulations with 109 neurons, 1013 synapses,” *High Perform. Comput. Networking, Storage Anal. Proc. Conf.*, no. c, pp. 1–12, 2009.

[49] G. Indiveri *et al.*, “Integration of nanoscale memristor synapses in neuromorphic computing architectures,” *Nanotechnology*, vol. 24, no. 38, p. 384010, Sep. 2013.

[50] M. Jerry *et al.*, “Ferroelectric FET analog synapse for acceleration of deep neural network training,” *Tech. Dig. - Int. Electron Devices Meet. IEDM*, vol. 6, no. c, pp. 6.2.1-6.2.4, 2018.

[51] S. R. Nandakumar *et al.*, “Experimental Demonstration of Supervised Learning in Spiking Neural Networks with Phase-Change Memory Synapses,” *Sci. Rep.*, vol. 10, no. 1, p. 8080, Dec. 2020.

[52] J. Wang *et al.*, “A Two-terminal Electric-double-layer Synaptic Device with Short-term Plasticity,” *2018 IEEE Electron Devices Technol. Manuf. Conf. EDTM 2018 - Proc.*, pp. 157–159, 2018.

[53] S. Chen *et al.*, “Wafer-scale integration of two-dimensional materials in high-density memristive crossbar arrays for artificial neural networks,” *Nat. Electron.*, vol. 3, no. 10, pp. 638–645, Oct. 2020.

[54] Y. Chen *et al.*, “Polymer memristor for information storage and neuromorphic applications,” *Mater. Horizons*, vol. 1, no. 5, p. 489, Jun. 2014.

[55] Q. Xia *et al.*, “Memristor-CMOS hybrid integrated circuits for reconfigurable logic,” *Nano Lett.*, vol. 9, no. 10, pp. 3640–3645, 2009.

[56] D. S. Jeong *et al.*, “Memristors for Energy-Efficient New Computing Paradigms,” *Adv. Electron. Mater.*, vol. 2, no. 9, pp. 1–27, 2016.

[57] D. B. Strukov *et al.*, “The missing memristor found,” *Nature*, vol. 453, no. 7191, pp. 80–83, May 2008.

[58] K. Stern *et al.*, “Uncovering Phase Change Memory Energy Limits by Sub-Nanosecond Probing of Power Dissipation Dynamics,” vol. 2100217, pp. 1–9, 2021.

[59] Y. Yin *et al.*, “Sub 10 ns fast switching and resistance control in lateral GeTe-based phase-change memory,” *Jpn. J. Appl. Phys.*, vol. 55, no. 6S1, p. 06GG07, Jun. 2016.

[60] A. Redaelli *et al.*, “Material and process engineering challenges in Ge-rich GST for embedded PCM,” *Mater. Sci. Semicond. Process.*, vol. 137, p. 106184, Jan. 2022.

[61] P. Wang *et al.*, "Picosecond amorphization of chalcogenides material: From scattering to ionization," *Appl. Phys. Lett.*, vol. 102, no. 11, p. 112108, Mar. 2013.

[62] D. Loke *et al.*, "Ultrafast phase-change logic device driven by melting processes," *Proc. Natl. Acad. Sci.*, vol. 111, no. 37, pp. 13272–13277, Sep. 2014.

[63] W. C. Chien *et al.*, "Endurance Evaluation on OTS-PCM Device using Constant Current Stress Scheme," in *2022 IEEE International Reliability Physics Symposium (IRPS)*, 2022, pp. P7-1-P7-4.

[64] M. H. R. Lankhorst, "Modelling glass transition temperatures of chalcogenide glasses . Applied to phase-change optical recording materials," vol. 297, pp. 210–219, 2002.

[65] J. R. Jameson *et al.*, "Conductive-bridge memory (CBRAM) with excellent high-temperature retention," in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 30.1.1-30.1.4.

[66] M. Saadi *et al.*, "On the mechanisms of cation injection in conducting bridge memories: The case of HfO 2 in contact with noble metal anodes (Au, Cu, Ag)," *J. Appl. Phys.*, vol. 119, no. 11, p. 114501, Mar. 2016.

[67] A. Belmonte *et al.*, "A thermally stable and high-performance 90-nm Al2O3 \ Cu-based 1T1R CBRAM cell," *IEEE Trans. Electron Devices*, vol. 60, no. 11, pp. 3690–3695, Nov. 2013.

[68] J. Joshua Yang *et al.*, "Engineering nonlinearity into memristors for passive crossbar applications," *Appl. Phys. Lett.*, vol. 100, no. 11, p. 113501, Mar. 2012.

[69] I. T. Wang *et al.*, "3D Ta/TaO x /TiO2/Ti synaptic array and linearity tuning of weight update for hardware neural network applications," *Nanotechnology*, vol. 27,

no. 36, 2016.

[70] S. Kumar *et al.*, "Oxygen migration during resistance switching and failure of hafnium oxide memristors," *Appl. Phys. Lett.*, vol. 110, no. 10, p. 103503, Mar. 2017.

[71] J. H. Yoon *et al.*, "Truly Electroforming-Free and Low-Energy Memristors with Preconditioned Conductive Tunneling Paths," *Adv. Funct. Mater.*, vol. 27, no. 35, p. 1702010, Sep. 2017.

[72] M.-J. Lee *et al.*, "A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta2O5−x/TaO2−x bilayer structures," *Nat. Mater.*, vol. 10, no. 8, pp. 625–630, Aug. 2011.

[73] H. Y. Lee *et al.*, "Evidence and solution of over-RESET problem for HfOx based resistive memory with sub-ns switching speed and high endurance," in *2010 International Electron Devices Meeting*, 2010, pp. 19.7.1-19.7.4.

[74] A. C. Torrezan *et al.*, "Sub-nanosecond switching of a tantalum oxide memristor," *Nanotechnology*, vol. 22, no. 48, p. 485203, Dec. 2011.

[75] U. Böttger *et al.*, "Picosecond multilevel resistive switching in tantalum oxide thin films," *Sci. Rep.*, vol. 10, no. 1, p. 16391, Dec. 2020.

[76] S. Kim *et al.*, "Comprehensive Physical Model of Dynamic Resistive Switching in an Oxide Memristor," *ACS Nano*, vol. 8, no. 3, pp. 2369–2376, Mar. 2014.

[77] P. Huang *et al.*, "Analysis of the Voltage–Time Dilemma of Metal Oxide-Based RRAM and Solution Exploration of High Speed and Low Voltage AC Switching," *IEEE Trans. Nanotechnol.*, vol. 13, no. 6, pp. 1127–1132, Nov. 2014.

[78] Y. Sun *et al.*, "Multilevel memory and artificial synaptic plasticity in P(VDF-TrFE)-

based ferroelectric field effect transistors," *Nano Energy*, vol. 98, no. December 2021, p. 107252, Jul. 2022.

[79] A. V. Khvalkovskiy *et al.*, "Basic principles of STT-MRAM cell operation in memory arrays," *J. Phys. D. Appl. Phys.*, vol. 46, no. 13, p. 139601, Apr. 2013.

[80] S. Kanai *et al.*, "Electric-field-induced magnetization switching in CoFeB/MgO magnetic tunnel junctions with high junction resistance," *Appl. Phys. Lett.*, vol. 108, no. 19, p. 192406, May 2016.

[81] K. Ando *et al.*, "Spin-transfer torque magnetoresistive random-access memory technologies for normally off computing (invited)," *J. Appl. Phys.*, vol. 115, no. 17, p. 172607, May 2014.

[82] A. D. Kent and D. C. Worledge, "A new spin on magnetic memories," *Nat. Nanotechnol.*, vol. 10, no. 3, pp. 187–191, 2015.

[83] J. Shi *et al.*, "A correlated nickelate synaptic transistor," *Nat. Commun.*, vol. 4, pp. 1–9, 2013.

[84] J. Jiang *et al.*, "2D MoS2 Neuromorphic Devices for Brain-Like Computational Systems," *Small*, vol. 13, no. 29, 2017.

[85] J. Zhu *et al.*, "Ion Gated Synaptic Transistors Based on 2D van der Waals Crystals with Tunable Diffusive Dynamics," *Adv. Mater.*, vol. 30, no. 21, 2018.

[86] Y. Nishitani *et al.*, "Three-terminal ferroelectric synapse device with concurrent learning function for artificial neural networks," *J. Appl. Phys.*, vol. 111, no. 12, p. 124108, Jun. 2012.

[87] Y. Kaneko *et al.*, "Ferroelectric Artificial Synapses for Recognition of a Multishaded Image," *IEEE Trans. Electron Devices*, vol. 61, no. 8, pp. 2827–2833,

Aug. 2014.

[88] F. De Roose *et al.*, “12.3 Memory Solutions for Flexible Thin-Film Logic: up to 8kb, >105.9kb/s LPROM and SRAM with Integrated Timing Generation Meeting the ISO NFC Standard,” in *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, 2019, vol. 31, pp. 206–208.

[89] J.-T. Yang *et al.*, “Artificial Synapses Emulated by an Electrolyte-Gated Tungsten-Oxide Transistor,” *Adv. Mater.*, vol. 30, no. 34, p. 1801548, Aug. 2018.

[90] C. J. Wan *et al.*, “Short-Term Synaptic Plasticity Regulation in Solution-Gated Indium–Gallium–Zinc-Oxide Electric-Double-Layer Transistors,” *ACS Appl. Mater. Interfaces*, vol. 8, no. 15, pp. 9762–9768, Apr. 2016.

[91] C.-T. Lee *et al.*, “Metal–oxide–semiconductor devices using Ga2O3 dielectrics on n-type GaN,” *Appl. Phys. Lett.*, vol. 82, no. 24, pp. 4304–4306, Jun. 2003.

[92] J. Choi *et al.*, “Organolead Halide Perovskites for Low Operating Voltage Multilevel Resistive Switching,” *Adv. Mater.*, vol. 28, no. 31, pp. 6562–6567, Aug. 2016.

[93] C. Gu and J.-S. Lee, “Flexible Hybrid Organic–Inorganic Perovskite Memory,” *ACS Nano*, vol. 10, no. 5, pp. 5413–5418, May 2016.

[94] E. J. Yoo *et al.*, “Resistive Switching Behavior in Organic-Inorganic Hybrid CH 3 NH 3 PbI 3 −x Cl x Perovskite for Resistive Random Access Memory Devices,” *Adv. Mater.*, vol. 27, no. 40, pp. 6170–6175, Oct. 2015.

[95] B. Hwang *et al.*, “Effect of halide-mixing on the switching behaviors of organic-inorganic hybrid perovskite memory,” *Sci. Rep.*, vol. 7, no. 1, p. 43794, Mar. 2017.

[96] E. Yoo *et al.*, “Bifunctional resistive switching behavior in an organolead halide

perovskite based Ag/CH3NH3PbI3-xClx/FTO structure," *J. Mater. Chem. C*, vol. 4, no. 33, pp. 7824–7830, 2016.

[97] T. Kamiya and H. Hosono, "Material characteristics and applications of transparent amorphous oxide semiconductors," *NPG Asia Mater.*, vol. 2, no. 1, pp. 15–22, 2010.

[98] D. Y. Kang *et al.*, "Improved characteristics of amorphous indium-gallium-zinc-oxide-based resistive random access memory using hydrogen post-annealing," *Appl. Phys. Lett.*, vol. 109, no. 7, p. 073105, Aug. 2016.

[99] C.-H. Hsu *et al.*, "Multilevel resistive switching memory with amorphous InGaZnO-based thin film," *Appl. Phys. Lett.*, vol. 102, no. 6, p. 062905, Feb. 2013.

[100] H.-M. Kwon *et al.*, "Current hysteresis by oxygen vacancy exchange between oxides in Pt/a-IGZO/TaOx/W," *Appl. Surf. Sci.*, vol. 293, pp. 220–224, Feb. 2014.

[101] P. Ma *et al.*, "High-Performance InGaZnO-Based ReRAMs," *IEEE Trans. Electron Devices*, vol. 66, no. 6, pp. 2600–2605, Jun. 2019.

[102] H. Abbas *et al.*, "Reversible transition of volatile to non-volatile resistive switching and compliance current-dependent multistate switching in IGZO/MnO RRAM devices," *Appl. Phys. Lett.*, vol. 114, no. 9, p. 093503, Mar. 2019.

[103] C.-C. Chang *et al.*, "Solving the integration problem of one transistor one memristor architecture with a Bi-layer IGZO film through synchronous process," *Appl. Phys. Lett.*, vol. 112, no. 17, p. 172101, Apr. 2018.

[104] U. B. Isyaku *et al.*, "ZnO Based Resistive Random Access Memory Device: A Prospective Multifunctional Next-Generation Memory," *IEEE Access*, vol. 9, pp. 105012–105047, 2021.

[105] P. B. Pillai *et al.*, "Diffusion-Controlled Faradaic Charge Storage in High-

Performance Solid Electrolyte-Gated Zinc Oxide Thin-Film Transistors," *ACS Appl. Mater. Interfaces*, vol. 10, no. 11, pp. 9782–9791, Mar. 2018.

[106] T. Kamiya *et al.*, "Present status of amorphous In–Ga–Zn–O thin-film transistors," *Sci. Technol. Adv. Mater.*, vol. 11, no. 4, p. 044305, Feb. 2010.

CHAPTER 4

Solid Electrolyte-Gated Ta_2O_5/ZnO Field-Effect transistor: Fabrication, modeling, and device mechanism

4.1 Introduction

This chapter elucidates the fabrication procedures and inherent properties of Tantalum Pentoxide (Ta_2O_5) and Zinc Oxide (ZnO) Field-Effect Transistors. Here we introduce a more accurate device model, that is utilized for simulation of the crossbar in subsequent chapters.

Section 4.2 details the fabrication process of the devices, focusing on the intrinsic attributes of Ta_2O_5/ZnO thin films. Sections 4.3 and 4.4 delve into the operational principles of the Ta_2O_5 and ZnO Solid Electrolyte-Gated Field-Effect Transistors (SE-FET), and depict a Simulink model of the SE-FET. This model is based the drift-diffusion model established in previous works [1][2]. The key contribution lies in the integration of a gate current model, which had not been addressed in the earlier studies.

4.2 Device fabrication

The configuration of a bottom-gated three-terminal Tantalum Pentoxide (Ta_2O_5)/Zinc Oxide (ZnO) TFT is presented in Fig 4.1. The Ta_2O_5 and ZnO layers are deposited onto indium tin oxide (ITO) substrates using a radio frequency (RF) sputtering process, utilizing ceramic targets composed of 99.99% Ta_2O_5 and ZnO. The device is fabricated at room temperature, wherein the ITO and Ta_2O_5 serve as a unified bottom gate. ZnO is patterned employing AZ 1514 photoresist and subsequently etched utilizing a diluted solution of Hydrochloric (HCl) acid. The Aluminum (Al) contact is deposited via

a thermal evaporation technique, followed by a lift-off process. Detailed procedures regarding the device fabrication have been previously reported in [2][3].

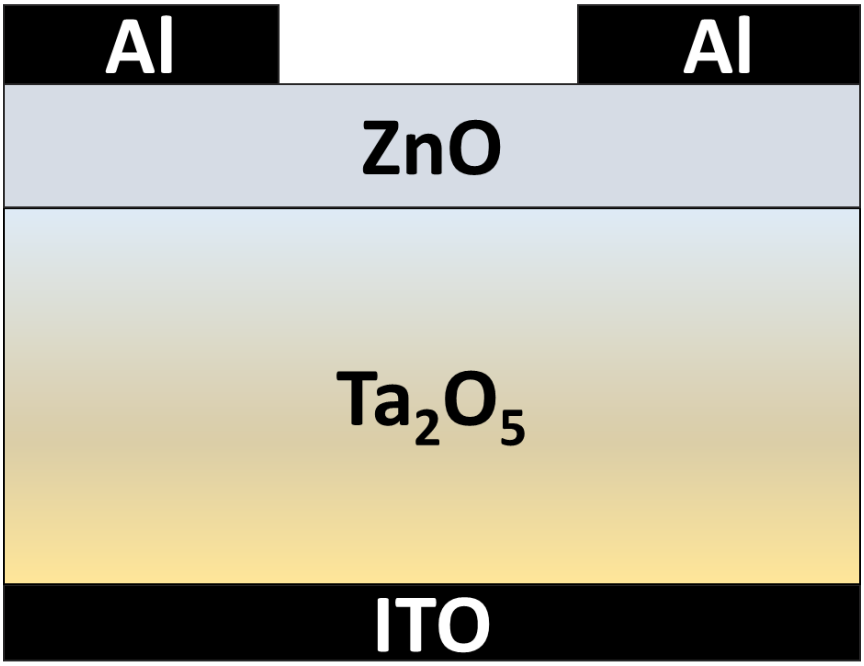


Fig 4.1 Structure of a Ta2O5/ZnO TFT

Fig 4.2 and Fig 4.3 exhibit the Atomic Force Microscopy (AFM) and X-ray diffraction (XRD) results for the amorphous Ta_2O_5/ZnO deposited via RF sputtering. The surface roughness, characterized by the average arithmetic value of roughness, for the film is recorded as 1.6 nm for Ta_2O_5 and 1.9 nm for ZnO. The XRD pattern of the Ta_2O_5 and ZnO films deposited on the ITO/glass substrate aligns closely with that of the ITO/glass substrate background, with no additional peak attribute to Ta_2O_5 and ZnO films, suggesting the absence of significant crystalline structures during the room-temperature deposition process.

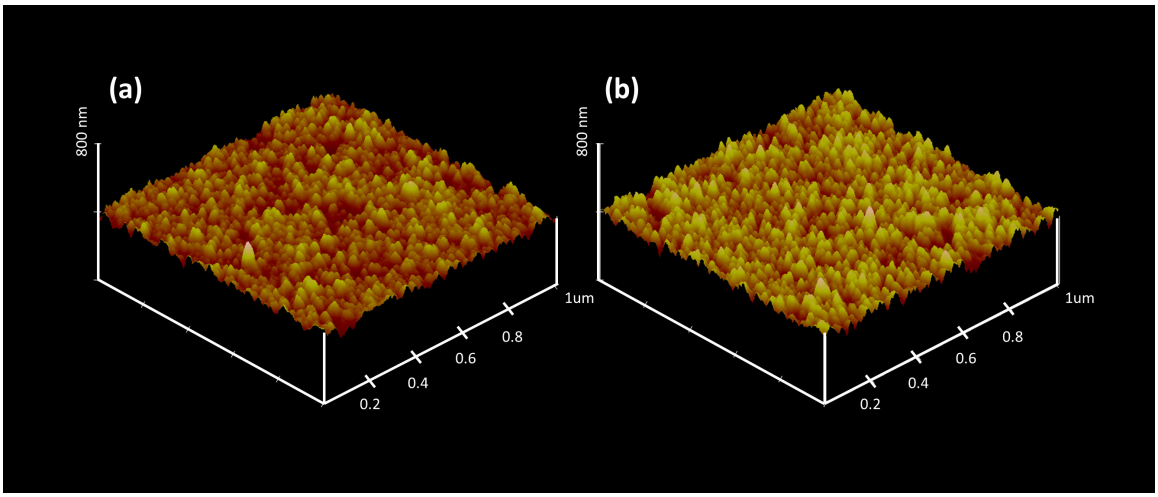


Fig 4.2 AFM images of (a) 275nm Ta_2O_5 and (b) 40nm ZnO deposited on ITO substrates.

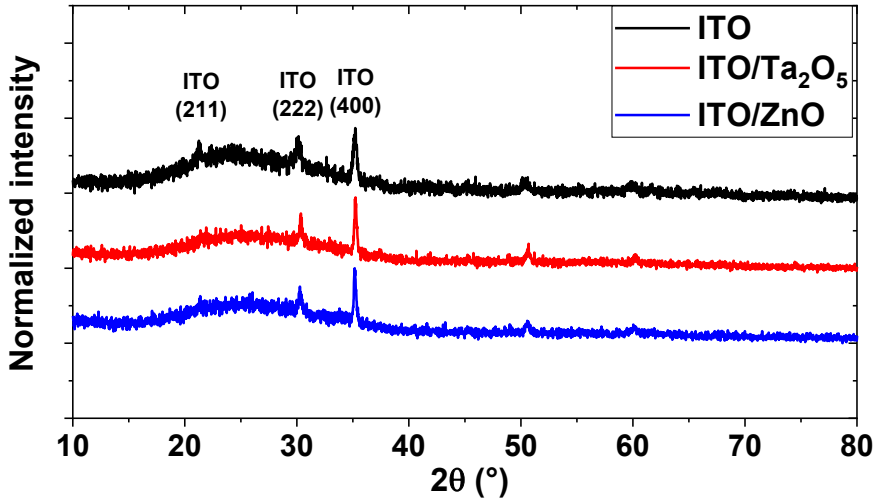


Fig 4.3 XRD of Ta_2O_5 and ZnO films and background ITO substrate.

4.3 Physics of Ta_2O_5/ZnO SE-FETs

Fig 4.4 illustrates the transfer current-voltage (I-V) curves for the Ta_2O_5-based FETs. During this process, a static voltage of 0.1 V is applied to the drain source of the device. The gate voltage is varied from -5 V to +5 V, subsequently sweeping back to -5 V

at a constant scan rate of 140 mV/s. The I_{DS}-V_{GS} transfer characteristics, as depicted in Fig 4.4 (a), exhibit a counterclockwise hysteresis with an on/off ratio exceeding 10^5. Notably, devices with a thicker oxide display a wider hysteresis window and higher on-current. The I_{GS}-V_{GS} transfer characteristics (Fig 4.4 (b)) are separately recorded with both drain sources of the device grounded to minimise the noise encountered for gate current recording in sub-nA range. The gate current of the device exhibits a clockwise hysteresis, with a maximum gate current reaching only 1nA for devices equipped with a 275nm gate oxide.

The operational mechanism of this device can be described as follows: the Ta_2O_5 serves as a high-k ($\varepsilon_{Ta2O5} = 20.8$ [1]) dielectric material. Owing to the inherent characteristics of the material under the deposition conditions, this material often exhibits high concentrations of mobile oxygen vacancies [4]–[6] with relatively low activation energies (1-2 eV) [7][8]. Consequently, during the forward scan, a positive voltage instigates a movement of highly mobile oxygen vacancies (V^{2+}) and oxygen ions (O^{2-}) towards the opposite interfaces within the dielectric layer. Under positive bias, the positively charged oxygen vacancies (V^{2+}) drift toward the dielectric/semiconductor interface, culminating in increased carrier concentration in the channel and, thus, augmenting channel conductivity (Fig 4.5 (a)). The ion accumulation during this process establishes a local concentration gradient of vacancies (V^{2+}) and oxygen ions (O^{2-}). This concentration gradient facilitates the diffusion of ionic species and generates an internal electric field in the direction opposite to the applied electric field.

During the reverse scan, as the applied voltage starts to reduce the buildup of ions, it ultimately reaches a steady state where the applied electric field is neutralized by diffusion, and the device attains its peak current (Fig 4.5 (b)). Beyond this point, though

the gate voltage continues to reduce, the already built-up oxygen vacancies at the Ta_2O_5/ZnO interface are yet to diffuse away and maintain the high carrier concentration in the ZnO channel (Fig 4.5 (c)).

As the gate voltage is swept to negative values, the applied electric field changes polarity but the concentration gradient built up during the positive scan is yet to diffuse away. Both the applied electric field and diffusion promote the recombination of oxygen vacancies (V^{2+}) and oxygen ions (O^{2-}) (Fig 4.5 (d)). Driven by an escalating negative electric field, the oxygen vacancies (V^{2+}) and oxygen ions (O^{2-}) drift towards each other and are forced to recombine. This phenomenon leads to a distinctive redox peak observed in the gate current characteristics during the reverse scan (marked on the gate current peak in Fig 4.4 (b)) [2]. The sudden decline in oxygen vacancies at the interface results in a dramatic reduction in the electron concentration in the ZnO channel thus the devices enters into the off state.

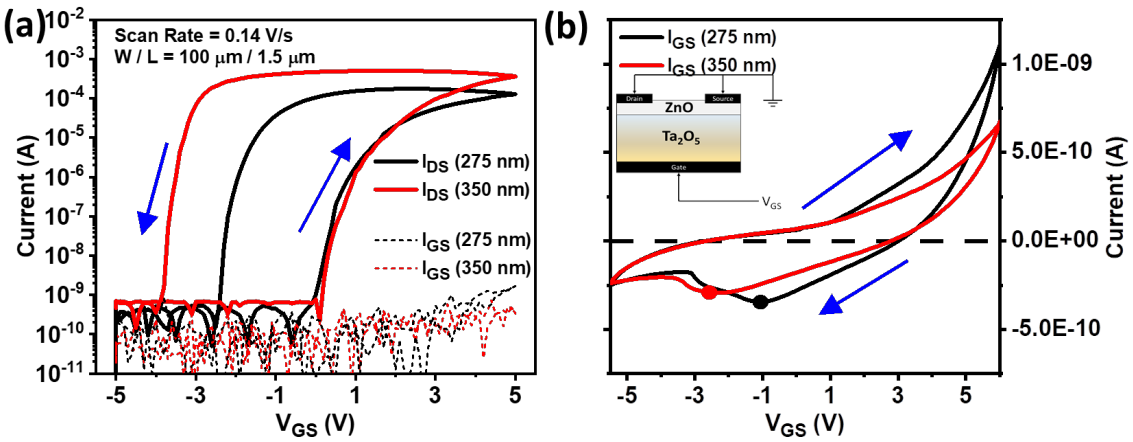


Fig 4.4 (a) I_{DS}-V_{GS} and (b) I_{GS}-V_{GS} characteristics of the Ta_2O_5/ZnO SE-FET at room temperature. I_{GS}-V_{GS} data were measured when both drain sources of the device were grounded.

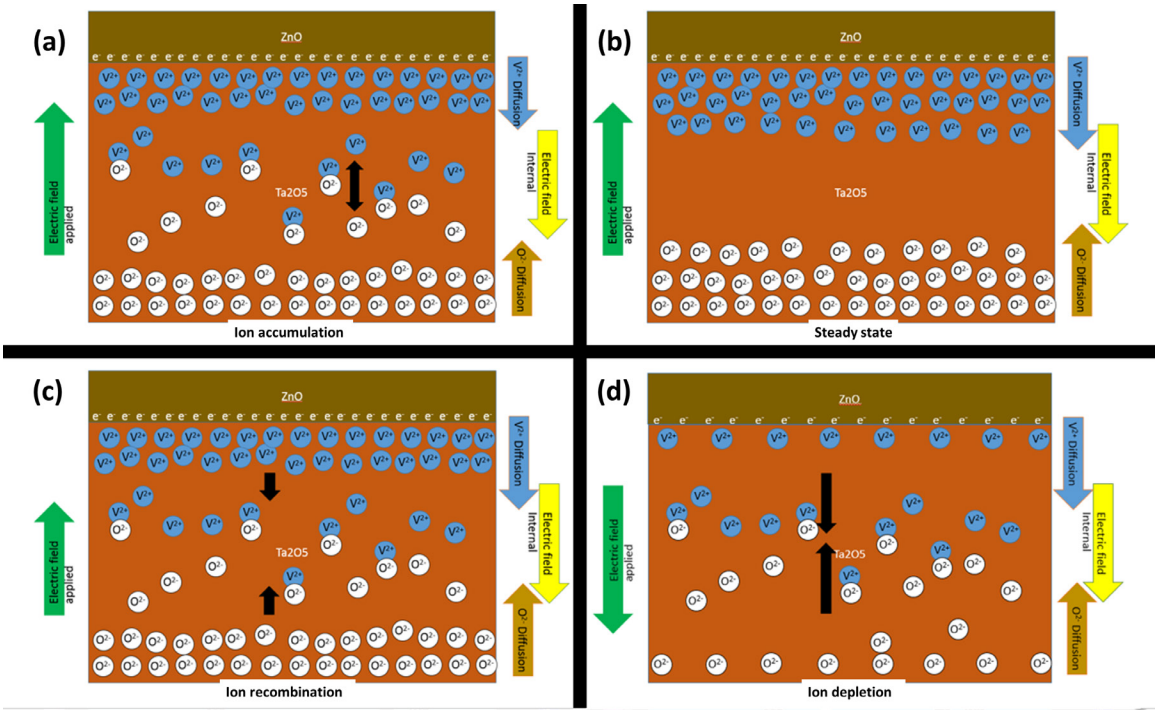


Fig 4.5 The schematic depicts the motion of ions within the insulator of the device during the dual sweep.

It is indeed important to differentiate between hysteresis attributed to ionic migration and that which is induced by trapping of charge at Ta_2O_5/ZnO interface. Fig 4.6 (a) and (b) illustrate the disparity between the I-V curves of a Ta_2O_5/ZnO Solid Electrolyte Field Effect Transistor (SE-FET) and a standard TFT, where Silicon Nitride (SiN) is utilized as a dielectric with a thin Ta_2O_5 layer.

As depicted in Fig 4.6 (a), the Ta_2O_5/ZnO SE-FET device demonstrates a high on-off ratio of 10^6 when the scan rate is below 1.2 V/s. Driven by this ion accumulation process, the drain-source current of the device primarily manifests an anti-clockwise hysteresis that is scan-rate dependent. The main memory function of the device is reflected by the hysteresis width [9][10], which is defined as the voltage difference between the forward and reverse scan. Fig 4.6 (c) plots this scan-rate-dependent hysteresis width at 100 nA. Comparatively, without the performance enhancement afforded by the ionic dielectric layer, the SiN/ZnO device displays a significantly lower on/off ratio of 10^5, despite having a substantially broader voltage window of 25 V (-10 V to +15 V). Contrasting with the ionic Ta_2O_5/ZnO SE-FET, the SiN-based device presents a clockwise hysteresis, a phenomenon frequently observed in TFTs based on ZnO [11] or InGaZnO [12], suggesting carrier trapping. This finding is further corroborated by the clockwise hysteresis observed in the Ta_2O_5/ZnO SE-FET at a nearly steady state (3.45×10^{-5} V/s scan rate, as shown in Fig 4.6 (d)).

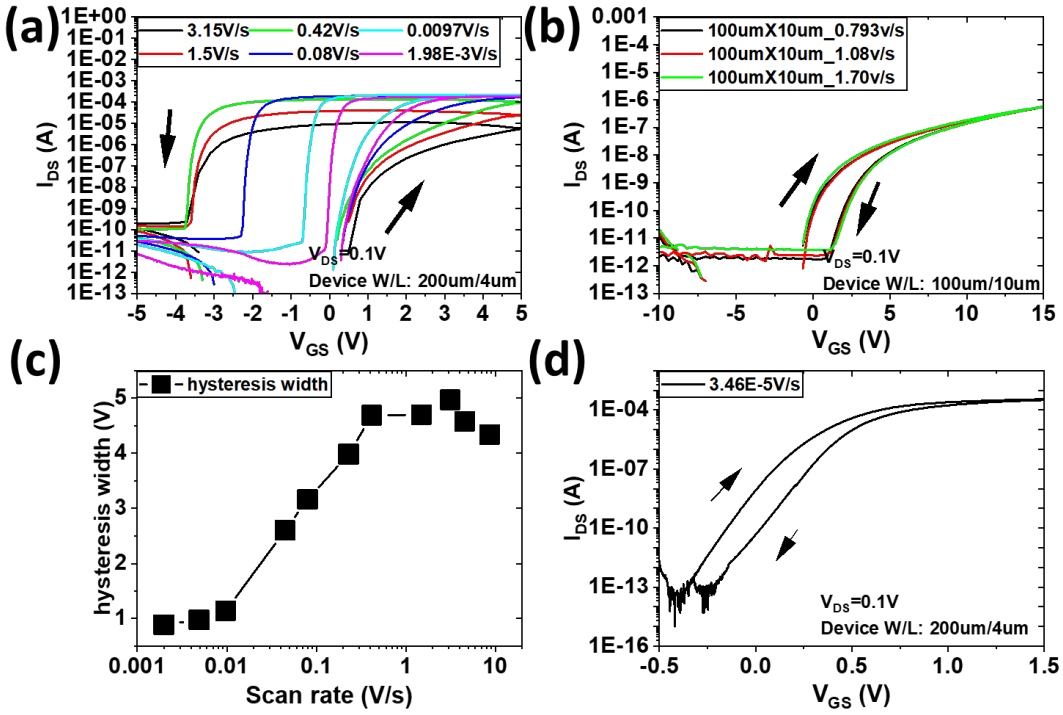


Fig 4.6 Dual-sweep transfer characteristics I_{DS} vs. V_{GS} (a) SE-FET with 275-nm Ta_2O_5 as gate oxide, and (b) devices with 60-nm SiN + 30-nm Ta_2O_5. (c) Hysteresis width vs. scan rate at 100 nA. (d) Dual-sweep I_{DS} vs. V_{GS} of a 275-nm Ta_2O_5 device at near steady state.

Given a constant gate bias, the device characteristics display a significant time-dependent behavior due to the continual generation of ions under the reversible redox reaction. As illustrated in Fig 4.7, when the device is subjected to prolonged +5V pulses, there exists an almost linear correlation between the device conductivity and pulse width. [2]. This dynamic alteration in conductivity enables the device to integrate and retain the input signal from the gate terminal of the device with considerable linearity under a fixed bias.

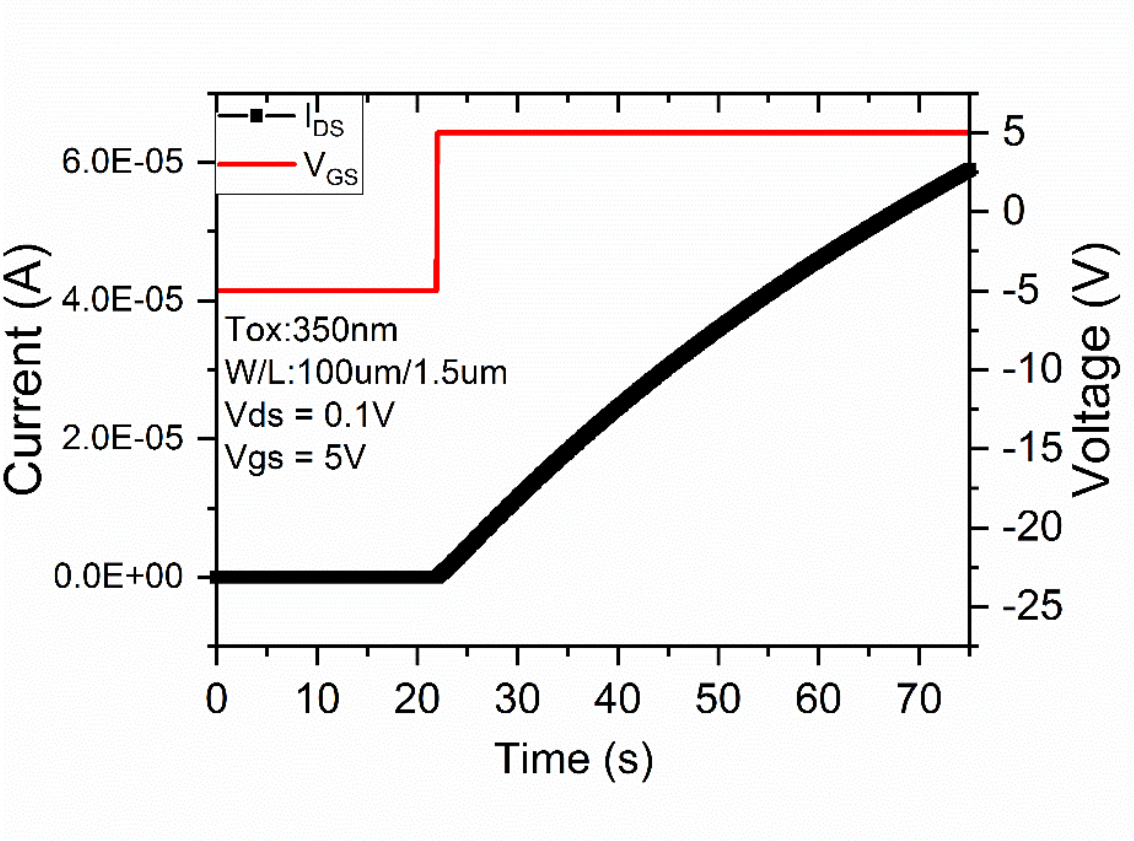


Fig 4.7 Time-dependent I_{DS} when switching from a constant V_{GS} of -5 V to a constant V_{GS} of +5 V.

The retention of the SE-FET is influenced by both the environmental temperature and the thickness of the Ta_2O_5 layer. For devices with equivalent gate insulator thickness (275 nm as shown in Fig 4.8 (a)), escalating the temperature from 20°C to 60°C triggers the decay of I_{DS} from 5×10^{-5} A to below 1×10^{-9} A. This leads to a decrease in retention time from 230 seconds to 50 seconds. This disparity in device retention time is likely attributable to the expedited diffusion of ions at the elevated temperature.

At a constant temperature (60°C as indicated in Fig 4.8 (b)), the retention time is dictated by the thickness of the gate insulator. As illustrated in Fig 4.8 (b), the devices with 350-nm gate insulators exhibit a decay in I_{DS} from 5×10^{-5} A to 1×10^{-5} A in the initial 200

seconds, followed by a constant I_{DS} value. In contrast, the devices with a 275-nm gate insulator layer show a considerably faster decay from 5×10^{-5} A to 1×10^{-9} A within a span of 60 seconds. Moreover, the device with a 120-nm gate insulator displays an even more accelerated decline from 5×10^{-5} A to 1×10^{-9} A within merely 5 seconds. This variation in retention time is likely linked to the difference in the concentration of oxygen vacancies in oxide layers of differing thicknesses.

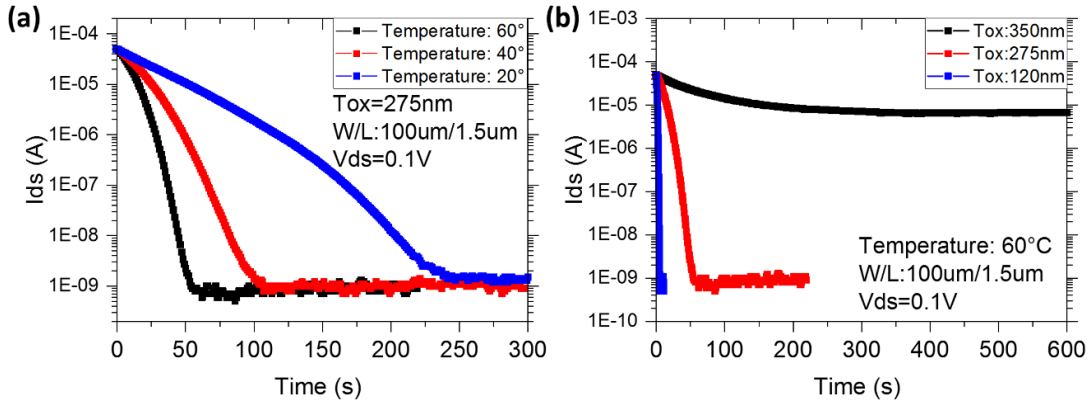


Fig 4.8 (a) Ids decay of devices with a gate insulator thickness of 275 nm at 20°C, 40°C, and 60 °C. (b) Ids decay of devices with a gate insulator thickness of 120 nm, 275 nm, or 350 nm at 60°C.

4.4 Modeling of the SE-FET

4.4.1 Basic of the SE-FET model

The synaptic SE-FET model described here draws upon the drift-diffusion model by Dr. Ashwani Kumar in [1]. The model utilized the voltage on the SE-FET gate as the input, simulating the drain-source current of each synaptic SE-FET by self-consistently integrating the drift-diffusion equation of the mobile ions in Ta_2O_5 and a Poisson solver for the ZnO channel. As shown in Fig 4.9, a Simulink model based on the drift-diffusion approach only models the drain source as physical resistance. Despite considering ionic

movement in the form of the drift-diffusion equation, the model lacks a physical representation of the gate, leading to an absence of a physical gate terminal as well as the leakage current (Fig 4.9). This restricts the use of the model in circuit simulation, as the gate of the device cannot physically connect to other devices, nor can it simulate leakage during operation.

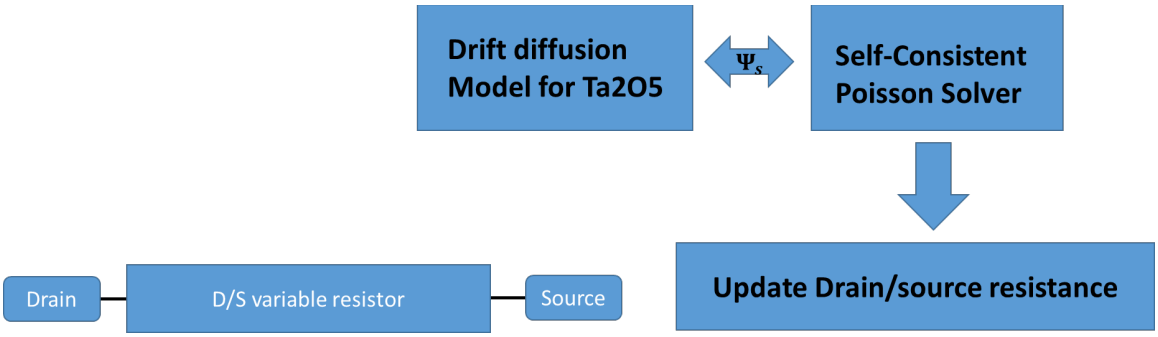


Fig 4.9 Scheme of the Simulink model based on the original SE-FET proposed by Dr. Ashwani Kumar in [1].

In the model reported in [1], our focus was directed towards exploring the subthreshold slope, which has been recorded at values as low as 26 mV/dec during the reverse scan in SE-FET devices. Given the Boltzmann limit at 60 mV/dec and the ionic, non-ferroelectric nature of Ta_2O_5, our objective was to investigate the potential negative capacitance induced by ionic accumulation.

The initial model assumed an equivalent response from oxygen ions and vacancies to both positive and negative voltage, and permitted free movement of ions from the ITO/Ta_2O_5 to Ta_2O_5/ZnO interface. From this premise, the drift velocity of the set oxygen ions and vacancies (v_d) was defined as:

$$v_d = \mu_{ion} \times \varepsilon_{ox} \quad \text{(Eq 4.1)}$$

where ε_{ox} is the electric field and the mobility of oxygen ions μ_{ion} is defined as:

$$\mu_{ion} = Q_{vac} a^2 \, f exp(-\frac{E_a}{KT})/KT \quad \text{(Eq 4.2)}$$

Here, Q_{vac} is the charge on the oxygen ions at 2e, a is the effective hopping distance, E_a is the height of the potential barrier, and f is the attempt frequency. From this expression, the dynamics of the sheet charge density (Q_{ox}) at the Ta_2O_5/ZnO interface is governed by a balance between the drift and diffusion of the oxygen ions:

$$\frac{dQ_{ox}}{dt} = Q_{vac} n_{ion} \mu_{ion} \varepsilon_{ox} - Q_{vac} D \Delta n_{ion} \quad \text{(Eq 4.3)}$$

Here, the first term describes the drift of oxygen ions/vacancies under an applied electric field. The second term is the diffusion of oxygen ions/vacancies driven by the concentration gradient. n_{ion} is the ion concentration, and D is the diffusion constant, written as:

$$D = \frac{KT}{Q_{vac}} \mu_{ion} \quad \text{(Eq 4.4)}$$

To simplify the diffusion term, a linear approximation to the oxide ion decay was assumed (Fig 4.10) where a fixed average diffusion length (x_D) and the oxide ion concentration (n_{peak}) decays linearly from the interface.

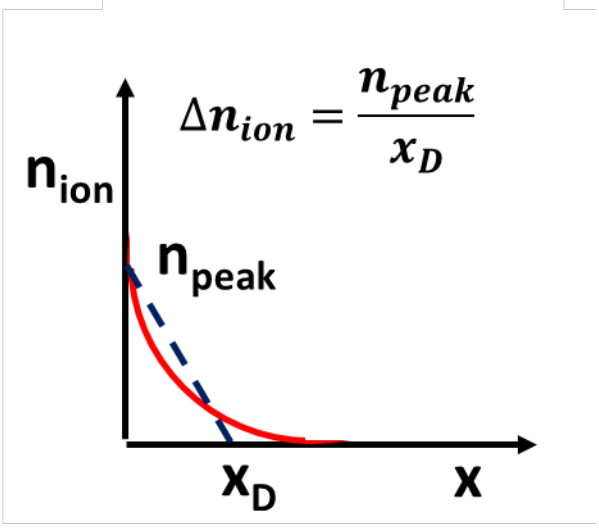


Fig 4.10 Scheme for linearly approximating the oxide ion concentration (n_{peak}) decay [1].

The drift-diffusion equation was coupled and self-consistently solved with a one-dimensional Poisson solver for ZnO. From this solution, the surface potential, and the net charge density in the channel Q_{ch}, the charge density in the conduction band Q_{free} was calculated, and the drain-source current is then obtained using the equation below:

$$I_{ds} = \mu_{ZnO} \frac{Q_{free}}{Q_{ch}} \frac{W}{L} Q_{free} V_{ds} \quad \text{(Eq 4.5)}$$

4.4.2 Proposed gate current model

For a comprehensive understanding of device performance in neuromorphic applications, it is crucial to conduct circuit-level simulations. To encapsulate the device's response within a circuit, we translated the model into MATLAB Simulink and enriched it by integrating a gate model that allows the simulation of the voltage distribution and current flow through the gate terminal. The gate current is modelled using the Ψ_s and Q_{ox} values, which were computed in Dr. A. Kumar's original model reported in [1]. The

parameters were extracted by fitting them to experimental data associated with the gate current during steady state and dual sweep conditions.

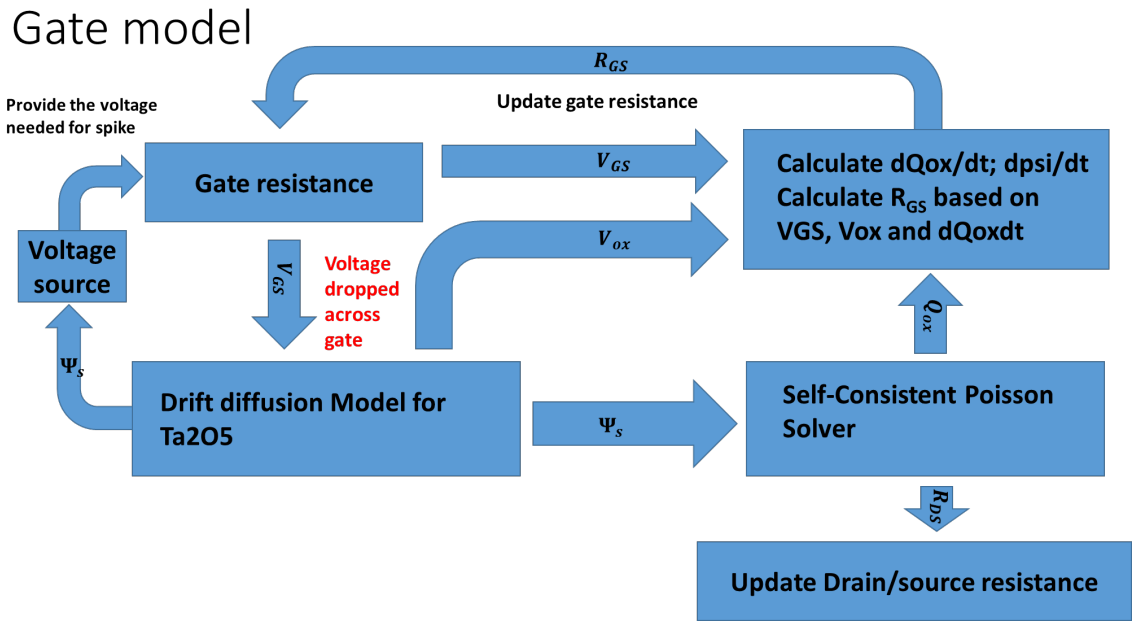


Fig 4.11 Scheme of the Simulink model of the SE-FET with the gate current model

The gate current model we derived integrates the charge storage properties linked to the redox reaction in Ta_2O_5 as a function of gate bias, as outlined in [2]. This inclusion of the redox process within our model is crucial to accurately reproduce the temporal decay of the gate current and the emergence of the redox peak during a dual sweep. Utilizing the parameters determined through our fitting process, the gate current of the device is expressed as:

$$I_{GS} = n \times \left(a_1 \times {V_{ox}}^3 + a_2 \times V_{ox} + c \times \frac{d\Psi_s}{dt}\right) + b \times \frac{dQ_{OX}}{dt} \qquad \text{(Eq 4.6)}$$

where V_{ox} is the voltage drop across the gate oxide (Ta_2O_5), Ψ_s is the surface potential, Q_{OX} is the sheet charge density in the semiconductor, $a_1 = 2.056 \times 10^{-12}$, $a_2 =$

8.1075×10^{-12}, $b = 1.897 \times 10^{-3}$, $c = 1.0076 \times 10^{-10}$, and $n = -1.968 + 1.968 \times \exp(V_{GS}/1.307)$.

The equivalent circuit of our Simulink model is illustrated in Fig 4.12. This model is comprised of a conductor in series with a battery, which symbolizes the rate of change of the oxide charge (dQ_{OX}/dt). Parallel to this, there is a steady-state gate conductor, which corresponds to leakage current across the oxide layer, and an additional oxide capacitor due to surface potential alterations ($d\Psi s/dt$) associated with ion migration in the insulator.

During steady state, the V_{OX} term primarily governs the leakage across the oxide. Conversely, the term dQ_{OX}/dt plays a central role in accounting for gate current during charging and discharging periods.

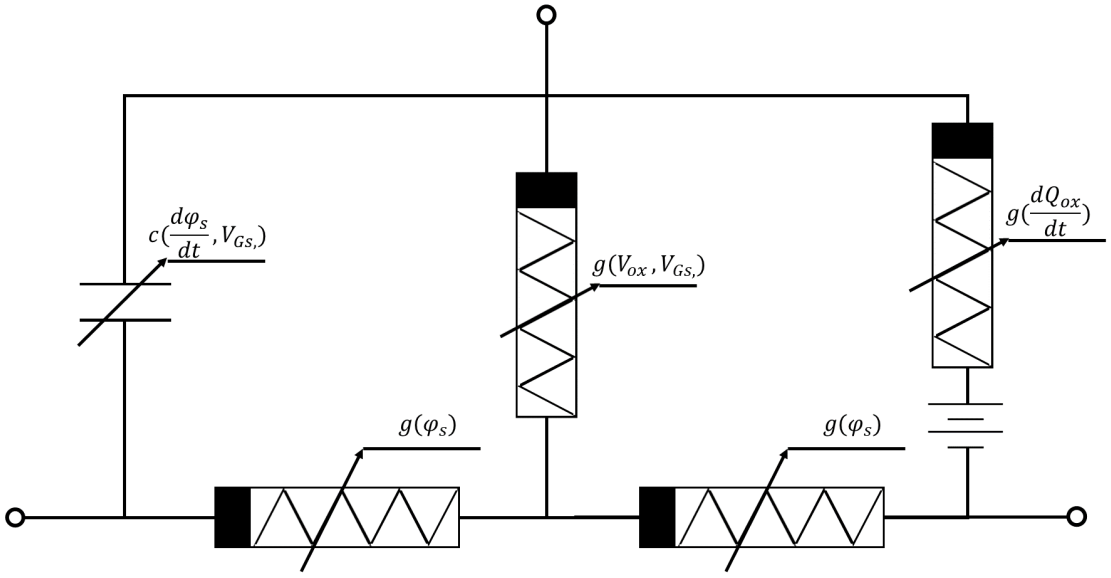


Fig 4.12 Equivalent circuit model of the SE-FET.

Fig 4.13 provides a comparative view of the measured and simulated I-V curves derived from our model. One particular point of interest in our investigation, as elaborated

in reference [1], is the understanding of the origin of steep switching observed during the reverse scan, denoted as point 2 in Fig 4.13 (a).

The work in [2] underscores the manifestation of this redox reaction as a peak in the gate current. This peak is dependent on the scan rate, as seen at point 2 Fig 4.13 (b). The reaction induces a negative capacitance during the reverse scan, allowing the device to exhibit steep switching within a certain scan rate range [1]. Understanding this process enhances our comprehensive knowledge of the SE-FET devices and their complex operational characteristics.

Another notable point is the increase of the gate current when the gate voltage is removed (as seen in Fig 4.13 (d) after 350s) and the negative differential resistance during the reverse scan (observed in Fig 4.13 (b) between 5V to 0V). Both are related to the ionic charging and discharging, and the opposite polarity between V_{GS} and V_{OX} causes this phenomenon.

In Fig 4.13 (b), during the reverse scan, as the applied electric field decreases, the built-up charge at the opposite polarity diffuses away. Similarly, the electrons in the channel also diffuse away. We now reveal that the increase of the current can be attributed to the dQ_{OX}/dt term. The reducing applied electric field and the strong internal field, caused by ions remaining at the opposite polarity, could also cause a temporary opposite polarity between V_{GS} and V_{OX}. This contributes to an opposite polarity between the applied voltage and gate current, resulting in a negative differential resistance in the transfer characteristics. Similar to the negative capacitance [1], this negative differential resistance is only observable in device transfer characteristics when a temporary opposite polarity is created between the applied electric field and internal field.

In the steady state depicted in Fig 4.13 (d) (before 350s), under the constant applied voltage, the applied electric field, ionic diffusion, and the strong internal field reach a dynamic equilibrium. However, when the applied voltage is abruptly removed after the 350s mark, the ionic diffusion and internal field opposing the applied field assume control. The ions of the opposite polarity begin to depart from the interface and engage in recombination. As the positively charged vacancies vacate the interface, the electrons in the channel, initially attracted to the positive charge, also disperse. This chain of events culminates in the generation of a negative current.

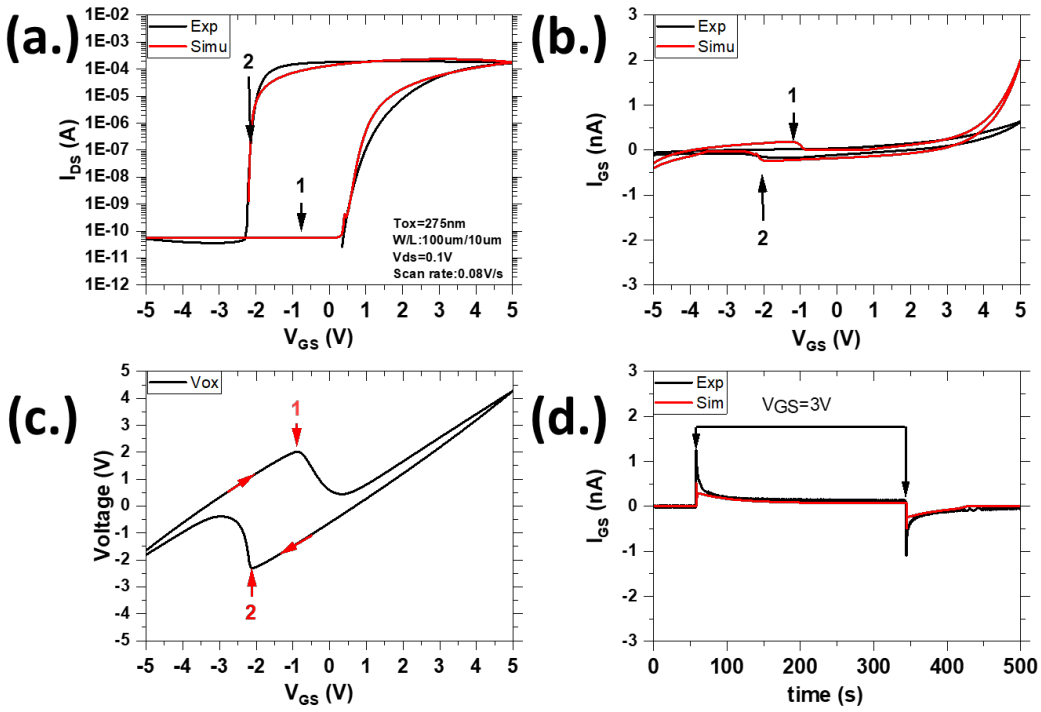


Fig 4.13 Comparison of (a) measured/simulated transfer I-V curves at a scan rate of 0.08 V/s and (b) gate current characteristics of the device at steady state. (c) Simulated Vox corresponding to the transfer I-V curve. (d) Comparison of the

measured/simulated time-dependent gate current characteristics at a constant voltage of 3 V.

In the model, as outlined in [1], we simplified the redox reaction as a symmetrical process in both forward and reverse scans, based on the assumptions of a linear approximation to the ionic concentration [1]. This suggests that a similar redox phenomenon should occur during both the forward scan and the reverse scan, as marked at points 1 and 2 in Fig 4.13 (a), causing redox peaks in the gate current. Theoretically, this redox reaction could potentially lead to a negative capacitance in both the forward and reverse scans. At the time, our understanding was that negative capacitance did indeed exist during the forward scan. However, its effect on the drain-source channel was deemed negligible, given that during this period, the device operates within the off region, where the drain-source resistance measures in the Gohm range and the current is less than 1 nA. The negative capacitance would therefore not have any observable impact during the off state of the devices. Consequently, we proposed that the redox reaction occurs in both scan directions, but observable steep switching occurs only during the reverse scan.

However, this assumption for the forward scan poses certain challenges when modeling the gate current. As illustrated in Fig 4.13 (b) and (c), our hypothesis predicted a redox peak not only in the reverse scan but also in the forward scan for both V_{ox} and I_{GS}. Contrarily, the comparison of the measured and simulated I_{GS} in Fig 4.13 (b) revealed a discrepancy, as the redox peak is observed only in the reverse scan. This inconsistency is likely a result of the influence of the electronic chemical potential difference of the redox reaction, which is attributable to the material differences between the ITO/Ta_2O_5 and ZnO/Ta_2O_5 interfaces [15].

The root cause of this discrepancy during the forward scan is likely traced back to the drift-diffusion equation where x_D was assumed as a constant, facilitating a linear approximation of the oxide ion concentration. This assumption results in a drift-diffusion model that symmetrically presents accumulation and depletion, thereby predicting two redox peaks, one each during the forward and reverse scans. This interpretation suggest a more dynamic approach by adjusting the oxygen ion mobility (μ_{ion}) and the average diffusion length (x_D) as follows [13]:

$$\mu_{ion} = \frac{\mu_{ion,0}}{\left[1+\left(\frac{\mu_0 \varepsilon_{ox}}{v_{sat}}\right)^{\beta}\right]^{\frac{1}{\beta}}} \quad \text{(Eq 4.7)}$$

$$x_D = \frac{x_{D,0}}{1-\left(\frac{Q_{ox,0}}{Q_{ox}}\right)^2} \quad \text{(Eq 4.8)}$$

By incorporating the dynamic oxygen ion mobility and diffusion length, we modelled the gate current by refitting the experimental data with the updated V_{ox} and Q_{ox} values. Fig 4.14 (a) and (c) demonstrate the impact of this modification, with the redox peak during the forward scan of the simulated V_{ox} noticeably reduced, while the drain-source current remains largely unaffected. This is due to the fact that these alterations predominantly influence the device when it is operating within the off region (highlighted at point 1 in Fig 4.14 (a)).

Consequently, the following improvements have been integrated in our gate model:

$$I_{GS} = [a_1 \times \sinh(a_2 \times V_{ox}) + b \times \frac{dQox}{dt} \times \frac{1}{n} + c \times V_{ox} + d \times \frac{d\Psi_s}{dt}] \times n \quad \text{(Eq 4.9)}$$

Where $a_1 = 1.258 \times 10^{-11}$, $a_2 = 0.7869$, $b = 8.45 \times 10^{-4}$, $c = 1.1075 \times 10^{-11}$, and $d = 2.0076 \times 10^{-11}$. In this enhanced model, the cubic term of Vox is replaced with a hyperbolic sine function (sinh) to effectively counteract the excessive gate current

observed at elevated voltages, as depicted in Fig 4.13 (b). The sinh function is often employed to model the drift-diffusion process of oxygen vacancies in Resistive Random-Access Memory (ReRAM) [14][15]. A voltage-dependent scaling factor, denoted as 'n', is also introduced to further align the model predictions with the experimental data. The modified equation reads as follows:

$$n = 2.35 \times \frac{1}{1+e^{(-abs(1.22 \times V_{ox}))}} \quad \text{(Eq 4.10)}$$

As depicted in Fig 4.14 (b) and (c), the modified model successfully corrected the excessive gate current during the transfer scan without overcorrecting the steady state condition. This enhanced drift-diffusion model yields a closer match with the observed gate current, while leaving the simulated drain-source current unaffected. However, the model currently operates within a constrained voltage range for a specific gate oxide thickness. Furthermore, the redox peak generated by the simulation remains more pronounced than the measured data due to the sharp Vox peak produced by the dynamic drift-diffusion model. A more holistic approach could take into account the differences between the ITO/Ta_2O_5 and Ta_2O_5/ZnO interfaces, where the ion drift diffusion might be modeled differently at each interface, and the oxygen and vacancy species could be modeled separately.

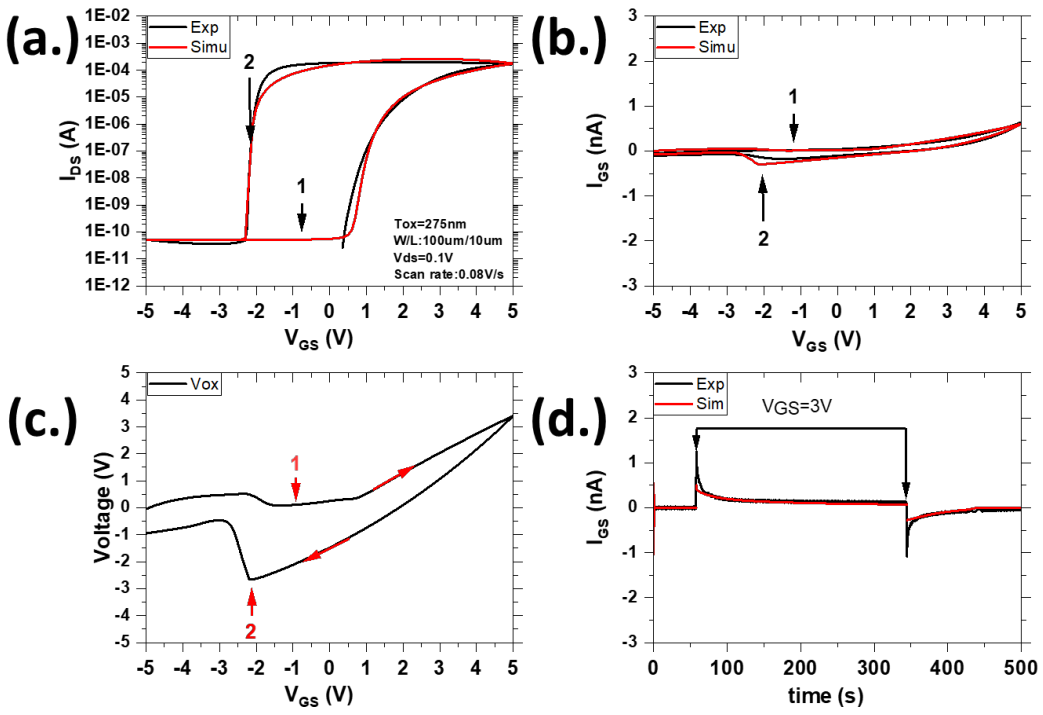


Fig 4.14 Comparison of (a) measured/simulated transfer I-V curves at a scan rate of 0.08 V/s and (b) gate current characteristics of the device at steady state. (c) Simulated Vox and Ψs for the transfer I-V curve. (d) Comparison of measured/simulated time-dependent gate current characteristics at a constant voltage of 3 V.

4.5 Summary

This chapter outlines the fabrication process of the SE-FET devices demonstrated in this thesis and elucidated their operational mechanism. The gate current model, which enables the device simulation presented in Chapter 4, details the updated physical interpretation of the drift-diffusion model originally reported in [1]. The initial model anticipated a symmetrical ion migration process where a redox reaction takes place during both the forward and reverse scans. However, this prediction does not align with the empirically measured gate current, as a redox peak is discernible only during the reverse

scan. This issue is mitigated, in part, by formulating the dynamic drift-diffusion model that curbs ion migration during the forward scan. This refined methodology demonstrates closer accordance with the measured gate current, providing a physical interpretation and representation that lays the foundation for the subsequent chapters.

4.6 References

[1] A. Kumar *et al.*, "Negative Capacitance beyond Ferroelectric Switches," *ACS Appl. Mater. Interfaces*, vol. 10, no. 23, pp. 19812–19819, Jun. 2018.

[2] P. Balakrishna Pillai *et al.*, "Diffusion-Controlled Faradaic Charge Storage in High-Performance Solid Electrolyte-Gated Zinc Oxide Thin-Film Transistors," *ACS Appl. Mater. Interfaces*, vol. 10, no. 11, pp. 9782–9791, Mar. 2018.

[3] P. Balakrishna Pillai and M. M. De Souza, "Nanoionics-Based Three-Terminal Synaptic Device Using Zinc Oxide," *ACS Appl. Mater. Interfaces*, vol. 9, no. 2, pp. 1609–1618, Jan. 2017.

[4] R. Nakamura *et al.*, "Diffusion of oxygen in amorphous Al2O3 , Ta2O5, and Nb2O5," *J. Appl. Phys.*, vol. 116, no. 3, p. 033504, 2014.

[5] M.-J. Lee *et al.*, "A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta2O5−x/TaO2−x bilayer structures," *Nat. Mater.*, vol. 10, no. 8, pp. 625–630, 2011.

[6] H. Jiang and D. A. Stewart, "Enhanced oxygen vacancy diffusion in Ta2O5 resistive memory devices due to infinitely adaptive crystal structure," *J. Appl. Phys.*, vol. 119, no. 13, 2016.

[7] K. Ratzke and F. Faupel, "Diffusion in metallic glasses and undercooled metallic melts," *Zeitschrift Fur Met.*, vol. 95, no. 10, pp. 956–960, 2004.

[8] U. Geyer *et al.*, “Atomic diffusion in the supercooled liquid and glassy states of the Zr41.2Ti13.8Cu12.5Ni10Be22.5 alloy,” *Phys. Rev. Lett.*, vol. 75, no. 12, pp. 2364–2367, 1995.

[9] C. Baeumer *et al.*, “Spectromicroscopic insights for rational design of redox-based memristive devices,” *Nat. Commun.*, vol. 6, p. 8610, 2015.

[10] Y. Kuo and H. Nominanda, “Nonvolatile hydrogenated-amorphous-silicon thin-film-transistor memory devices,” *Appl. Phys. Lett.*, vol. 89, no. 17, pp. 1–4, 2006.

[11] Z. Ye *et al.*, “Mechanism and Origin of Hysteresis in Oxide Thin-Film Transistor and Its Application on 3-D Nonvolatile Memory,” *IEEE Trans. Electron Devices*, vol. 64, no. 2, pp. 438–446, 2017.

[12] Y.-J. Chen and Y.-H. Tai, “Hysteresis of Transistor Characteristics of Amorphous IGZO TFTs Studied by Controlling Measurement Speed,” *ECS Solid State Lett.*, vol. 4, no. 4, pp. Q10–Q12, 2015.

[13] A. Kumar, “Addressing the Challenges of Future Electronics,” no. August, 2018.

[14] T. Chang *et al.*, “Synaptic behaviors and modeling of a metal oxide memristive device,” *Appl. Phys. A*, vol. 102, no. 4, pp. 857–863, Mar. 2011.

[15] C. Du *et al.*, “Reservoir computing using dynamic memristors for temporal information processing,” *Nat. Commun.*, vol. 8, no. 1, p. 2204, Dec. 2017.

CHAPTER 5

Off-state operation of SE-FETS

5.1 Introduction

The exceptional characteristics of ReRAM, such as incredibly low current consumption, a switching time of 10–100 ns, and CMOS compatibility, have positioned it as the preferred technology for compute-in-memory and neuromorphic applications. Nonetheless, ReRAM devices encounter difficulties due to variability when subjected to extreme scaling at dimensions below 10 nm, as they rely on filamentary processes. Furthermore, ReRAM typically necessitates write voltages (Vwrite) of at least ~1V, leading to a current Iwrite ranging from 1 μA to 1 mA, thereby resulting in a typical power consumption of 1 μW-1 mW, though exceptions are documented in [1].

ReRAMs are usually incorporated into crossbar arrays in two main configurations: a MOSFET-addressed structure featuring one transistor and one resistor (1T1R), or a traditional cross-point structure composed of one selector and one resistor. The latter design facilitates extremely compact scaling ($4F^2$), where F represents the footprint, but suffers from sneak current leakage paths when trying to address a single array element. Conversely, the MOSFET-addressed design faces challenges in scalability due to the substantial footprint required to deliver high write currents for ReRAM. These limitations have stimulated scientists to explore new physical mechanisms to bridge the gap between current computing performance and the ultimate efficiency of the human brain [2].

The concept of SE-FETs introduces a different approach. Here, the set or reset process is simply a result of ionic movement within the gate of a three-terminal device. This procedure is facilitated by separating the programming within the gate from the

conduction path in the channel during the read operation. Our channel merely serves as a sensor of the state written into the gate insulator, making our three-terminal device essentially a transistor with an integrated resistive memory within its gate insulator. In this chapter, we not only experimentally demonstrate logic operations at the single-device level but also assess the potential for scaling and applications of this concept.

5.2 Off-state operation for logic in memory

5.2.1 Basic device characteristics and operation

Fig 5.1 (a) and (b) offer a fundamental comparison between our Solid Electrolyte Field-Effect Transistor (SE-FET) and conventional Resistive Switches (RS). The latter are two-terminal devices that consist of metal contacts separated by an approximately 10 nm thick dielectric material. The resistance within this dielectric layer is subject to change via a filamentary process, which implies higher energy consumption due to the increased current flowing through the device. In these two-terminal RSs, both 'Write' and 'Read' operations are executed by applying voltage across the device. Although the 'Read' operation occurs at a lower voltage (Vread) compared to 'Write', any reduction in the write voltage tends to prolong the writing time, thereby leading to increased energy consumption [3]. Typically, the write energy is technology-dependent and presents scaling challenges due to variability and issues with the reliability associated with the size of the conductive filament [4].

Fig 5.1 (b) underscores the synaptic characteristics of our three-terminal SE-FET. In this configuration, 'Read' and 'Write' operations transpire across two separate terminals of the transistor—Drain-Source and Gate-Source, respectively. As outlined in Chapter 4, the device operates based on a diffusion mechanism, where redox-generated oxygen

vacancy ions accumulate at the Ta_2O_5/ZnO interface, causing an increase in carrier concentration. Naturally, it is expected that the write voltage (Vwrite) would be applied across the gate terminal, and the read voltage (Vread) through the Drain-Source. This arrangement allows the device to execute 'Write' and 'Read' operations either simultaneously or to perform 'Write' operations with minimal energy consumption when the device is switched off. As expounded in Chapter 4, the write operation primarily depends on ion accumulation and is almost independent of the Drain-Source terminal. Consequently, the device can be written even when both Drain-Source terminals are grounded.

Owing to the considerable thickness of the gate oxide (>100 nm), which is markedly thicker than the ~10 nm dielectric in ReRAM, even with a substantial gate-source voltage (V_{GS}) of 5V (as seen in Fig 5.1 (c) abd (d)), the write current (Iwrite) is constrained by the gate leakage current (I_{GS}) remaining in the sub-nanoampere range. Moreover, the incorporation of our device in a crossbar array does not necessitate a large drive current, as is required in a 1T1R configuration. The primary requirement for our device is a suitable gate voltage that is proportional to the thickness of the gate insulator and influences the switching speed—therefore, a higher voltage leads to a faster switching speed.

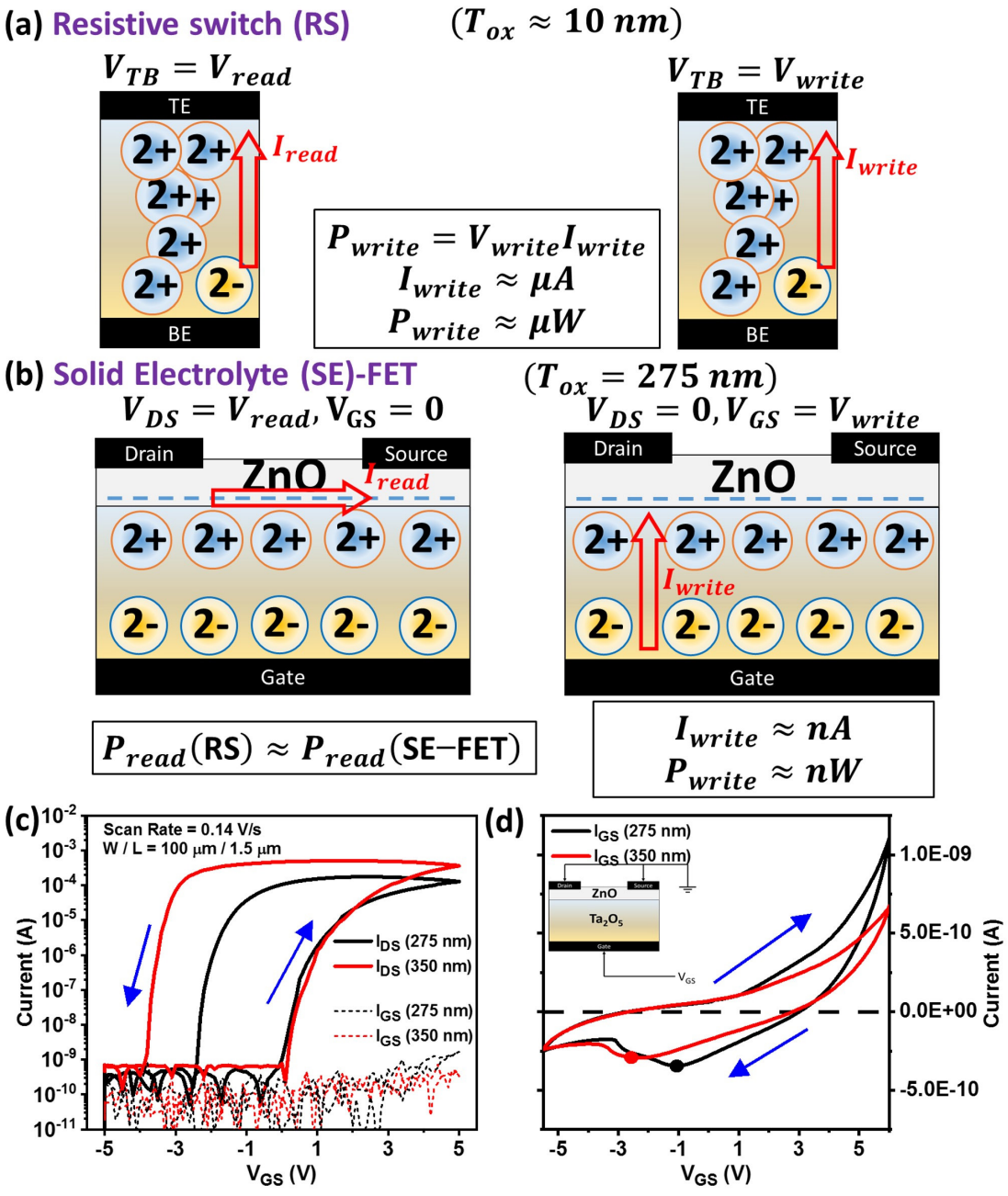


Fig 5.1 Mechanism of read and write operations in (a) resistive switching and (b) SE-FETs, demonstrating the possibility of low power consumption during writing in SE-FETs. (c) IDS-VGS and (d) IGS-VGS curves of the basic device at room temperature.

In the SE-FET, write and read operations can be conducted concurrently, as depicted in Fig 5.2. The drain-source current (I_{DS}) derived from the sequence of gate-source voltage (V_{GS}) pulses shown in Fig 5.2 (a) at a consistent drain-source voltage (V_{DS}) of 0.1 V is demonstrated in Fig 5.2 (b). There is a distinct point marked by a red circle between each pulse where V_{GS} equals 0 V. The I_{DS} characteristics exhibited in Fig 5.2 (b) corroborate that the device retains its current state even when the gate is turned off (V_{GS} equals 0 V). The polarity of I_{DS} determines the Low Resistance State (LRS) and High Resistance State (HRS), corresponding respectively to an accumulation and depletion of oxygen vacancies at the interface. This is elucidated in the inset of Fig 5.2 (b).

In this context, the device state mirrors the dual-sweep transfer characteristics exhibited at V_{GS} equals 0 V under a forward and reverse sweep as shown in Fig 5.2 (a). However, contrary to the dual sweep at a low scan rate of 7 mHz, the square wave illustrated in Fig 5.2 (a) at 1 Hz is incapable of accumulating or depleting adequate charge after each cycle. Consequently, this causes the ratio of LRS to HRS resistance (R_{LRS}/R_{HRS}) to decrease by three orders of magnitude, as highlighted in Fig 5.2 (b), compared with the six-order-of-magnitude R_{LRS}/R_{HRS} difference achieved in the dual sweep at low frequency, as depicted in Fig 5.1 (c).

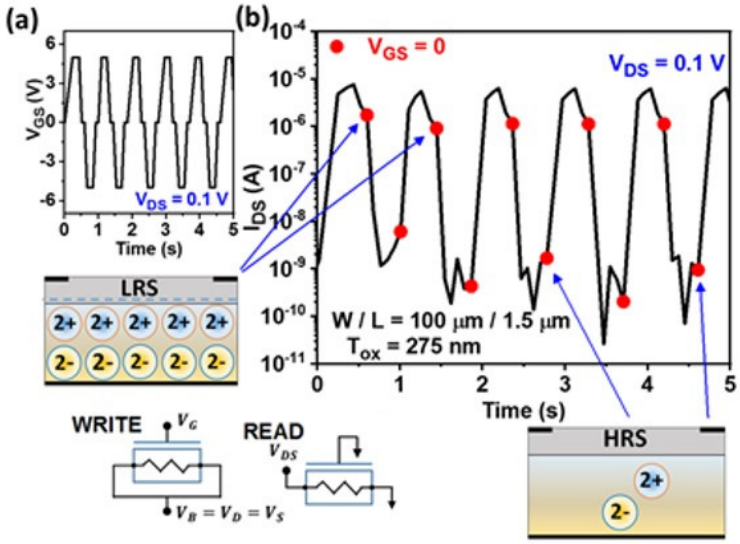


Fig 5.2 (a) Gate bias pulses at 1 Hz applied to the device with a constant $V_{DS} = 0.1\ V$. (b) Measured drain current characteristics as a function of time, exhibiting two resistance states (HRS and LRS) as the gate bias crosses zero, indicated by red circles.

While the SE-FET is capable of performing write and read operations concurrently, this feature could potentially result in significant energy consumption during the writing process. This challenge primarily emanates from the need to keep the drain-source terminal 'on' during the write operation, which leads to a substantial writing current of approximately 2 μA. In the context of two-terminal ReRAM, efforts have been focused on reducing the write time and voltage to address this issue. Nevertheless, in the SE-FET introduced here, which is a three-terminal device, the write and read operations occur at two distinct terminals. As a result, the high-energy consumption during the write operation can be mitigated by employing off-state operation.

In this context, the device operates between two terminals, V_G and V_B ($V_B = V_D = V_S$) during the writing process, and with V_{DS} at 0.1 V and V_{GS} at 0 during the reading

process. This arrangement is implemented via alternating pulses, as shown in Fig 5.3 (a). This unique feature suggests that the lateral scaling of the device may not be impeded by the voltage capability of the drain terminal.

Fig 5.3 (b) demonstrates the resulting measured drain-source current (I_{DS}) for a gate oxide thickness (T_{OX}) of 275 nm and a channel length of 1.5 µm during the reading process. A discernible Low Resistance State (LRS) and High Resistance State (HRS), distinguished by at least three orders of magnitude, are evident. An elevated resistance of approximately 1 MΩ in the LRS ensures a low drain current, thereby minimizing power consumption during the reading phase. The inset highlights the maximum and minimum variations of each state, which are less than one order of magnitude. As expected, the standard deviation of the HRS (0.3 of an order) exceeds that of the LRS (0.06 of an order), given the higher impact of noise on the HRS where the current is low.

Fig 5.3 (c) displays the gate current characteristics as a function of time. Here, the magnitude is confined to 10 nA, indicating that the maximum power consumption during the write operation is less than 50 nW for V_{GS} at 5 V. Fig 5.4 (a) portrays the behaviour of a device with a thicker oxide layer of 350 nm, subjected to the same input waveform as depicted in Fig 5.3 (a). In this case, the LRS and HRS differ by only one order of magnitude as the negative bias is incapable of resetting the device effectively (in the HRS, I_DS escalates from approximately 1 nA to roughly 1 µA). This challenge can be resolved by employing asymmetric gate pulses with magnitudes of +5V and -6V, respectively, alongside an enhanced gate bias for the reset process. This setup, as shown in Fig 5.4 (b), results in a difference of at least two orders of magnitude between the resistance states, as demonstrated in Fig 5.4 (c).

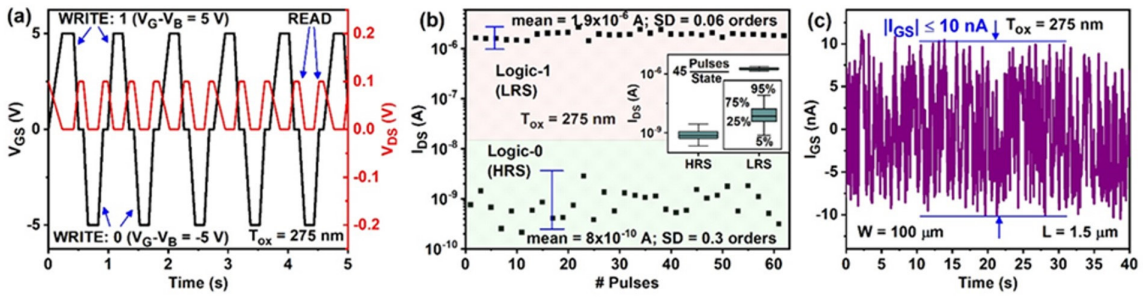


Fig 5.3 (a) Scheme of 1-Hz bias pulses for V_{GS} and V_{DS}, where $V_{DS} = 0$ and $V_{GS} \neq 0$ during the write operation and vice versa during the read operation, ensuring that the write and read operations are exclusive. (b) Drain current measured during the read operation ($V_{DS} = +0.1\ V$), indicating the presence of two resistance states (HRS and LRS) separated by three orders of magnitude, with a small variance (inset). (c) Measured gate current characteristics showing that the power consumption remains limited to the nanowatt level during the write operation, as the drain current remains zero.

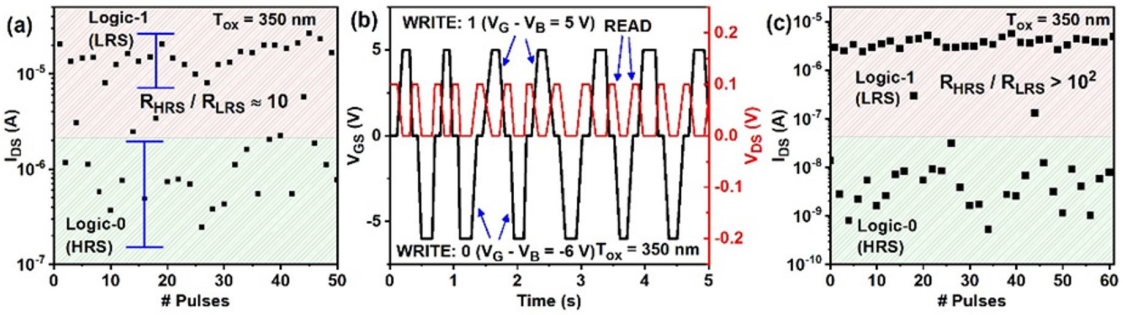


Fig 5.4 (a) Drain current measurements of a device with a 350-nm-thick oxide layer for the pulse scheme shown in Fig. 4.3(a). (b) Alternate pulse scheme at 1.5 Hz, with V_{GS} equal to $+5\ V$ and $-6\ V$ for switching between the LRS and HRS and (c) corresponding drain current measurements.

In the SE-FET, a trade-off exists between the retention time, HRS/LRS ratio, and switching speed as a function of the insulator thickness, as illustrated in Fig 5.5. In this figure, the devices are set to the LRS with a drain-source current (I_{DS}) of 5×10^{-5} A at a temperature of 60°C and a gate-source voltage (V_{GS}) of 5 V.

In comparison to the device with a 275-nm-thick gate insulator, the device equipped with a 350-nm-thick insulator displays stable retention behavior, sustaining the LRS at $I_{DS} = 1 \times 10^{-5}$ A for more than 900 seconds. Meanwhile, the device featuring a 275-nm-thick gate insulator declines to $I_{DS} = 1 \times 10^{-9}$ A after 60 seconds at a temperature of 60°C, as demonstrated in Fig 4.7.

This observation is consistent with the broader memory window presented in Fig 5.1 (a) for a thicker insulator. The longer ion travel distance and the smaller electric field during the set/reset operations at the same gate voltage account for this difference. For the same reasons, this setup results in a higher switching speed and R_{HRS}/R_{LRS} ratio for the device with a 275-nm-thick insulator at an identical frequency. However, as the switching speed increases, the on/off ratio correspondingly decreases.

The switching speed is determined by several factors, including ion diffusivity, recombination time, and oxide thickness. In this case, due to the overlapping of drain/source and gate contacts, both devices are unable to respond to switching times less than 150 ms, regardless of the insulator thickness. This limitation results in a lower R_{HRS}/R_{LRS} ratio at 13 Hz (corresponding to a switching time of 76 ms) for this particular device configuration.

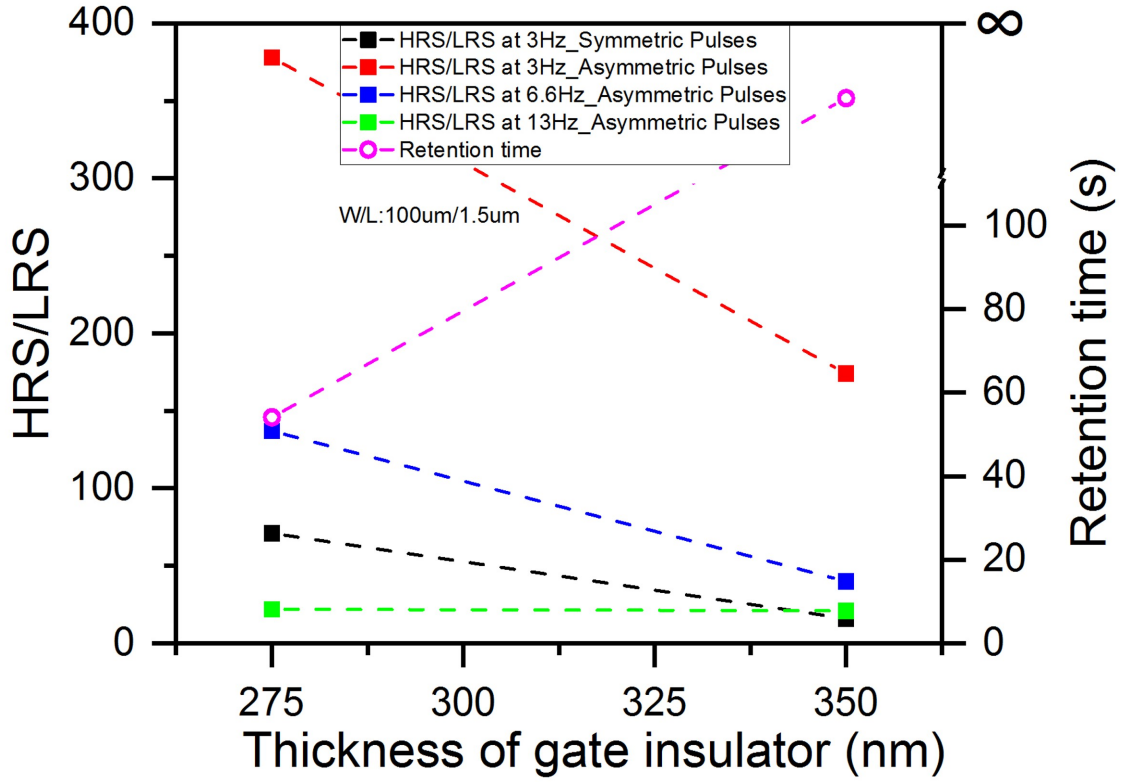


Fig 5.5 Gate insulator thickness vs. HRS/LRS ratio for asymmetric pulses (275 nm: +6 V/-5 V for set/reset; 350 nm: +5 V/-6 V for set/reset) (left axis) and retention time at 60°C (right axis).

5.2.2 Logic operations

Similar to a bipolar or complementary resistance switch [5], an example of the SE-FET operating as a finite-state machine is depicted in Fig 5.6 (a). The application of V_G or V_B = 1 V or 0, respectively, toggles the device between the LRS and HRS. Here, S represents the state of the device, with 0 or 1 indicating whether the device is in the HRS or LRS, respectively.

Fig 5.6 (b) provides a corresponding truth table for the state diagram. Here, the previous state of the device acts as an additional input, informing the following state. From this truth table, we derive the equation of state:

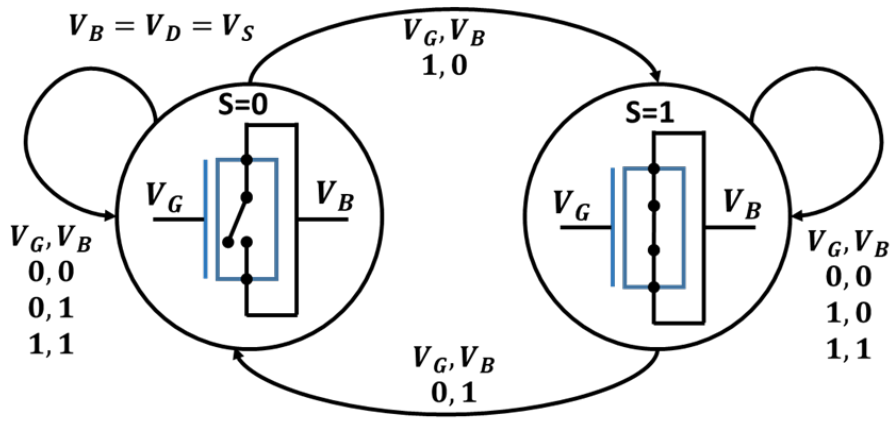
$$S = S \cdot V_G + S \cdot V_B' + V_G \cdot V_B' \quad \text{(Eq 5.1)}$$

Here, the operators ·, +, and ' correspond to AND, OR, and NOT logic functions, respectively. A single device can execute 14 out of a possible 16 logical functions of two inputs. The remaining two functions, XOR and XNOR, necessitate an additional device [5]. If the operations involving two or more inputs can be decomposed into a sequence of steps in accordance with Eq. 5.1, a single device can be utilized to perform these logic operations.

Fig 5.7 (a) illuminates the two steps required to realize the material implication (IMP) operation, A'+B, also known as "if A then B". This operation plays a significant role in completing a logic set when combined with the reset process [6]. Step 1 sets the state S to 1. In Step 2, inputs A and B are applied to V_B and V_G respectively, with all possible outcomes summarized in the corresponding truth table.

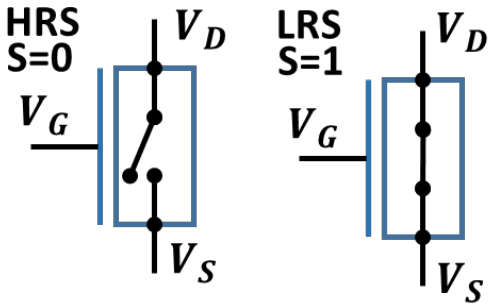
Fig 5.7 (b) illustrates the sequence of V_{GS} pulses generated in response to the inputs V_B and V_G, as specified in each row of the truth table in Fig 5.7 (a). The measured I_{DS} for a device with T_{OX} = 275 nm is plotted in Fig 5.7 (c), demonstrating a close alignment with the expected state in the last column of the truth table in Fig 5.7 (a) (as highlighted by colors).

(a)



(b)

S	V_G	V_B	S (Read)
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	1



$S = f(V_G, V_B, S)$

$S = S \cdot V_G + S \cdot V_B' + V_G \cdot V_B'$

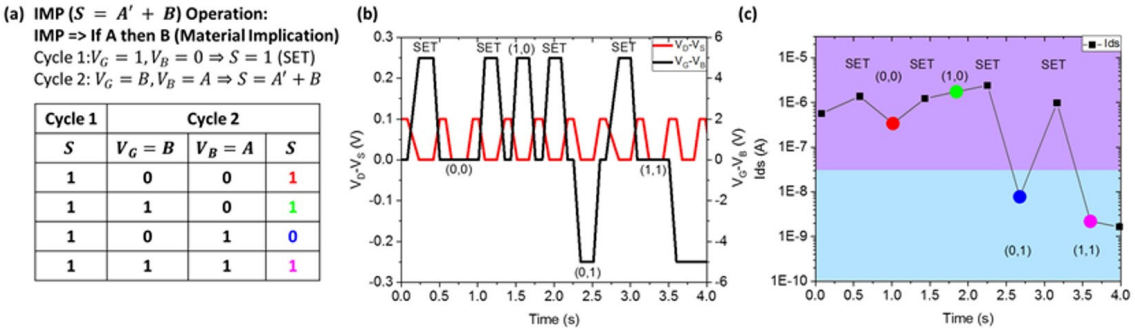
$V_B = V_S = V_D$

Fig 5.6 (a) Finite-state diagram of the device, where 1 for either V_G or V_B indicates application of a positive bias and S = 0 or 1 indicates the LRS or HRS, respectively. (b) Truth table describing the state transition.

Similarly, if we redefine A and B as V_B and V_G, respectively, we can perform the Not IMPlication (NIMP) operation, as demonstrated in Fig 5.8.

The corresponding realizations of two-input NAND and NOR operations are depicted in Fig 5.9 and Fig 5.10, respectively. Each operation must be divided into three

cycles, which include the initial set and reset [5]. State initialization takes place in Step 1, followed by the replication and storage of A' in Step 2. In Step 3, signals '1' & 'B' or '0' & 'B' are applied to V_B and V_G to perform the NAND or NOR operation, respectively. The measured drain currents for T_{OX} = 275 nm in Fig 5.9 (c) and Fig 5.10 (c) align with the state values in the truth table for each step in Fig 5.9 (a) and Fig 5.10 (a), respectively.



Cycle 1	Cycle 2		
S	$V_G = B$	$V_B = A$	S
1	0	0	1
1	1	0	1
1	0	1	0
1	1	1	1

Fig 5.7 (a) Steps to perform IMP (if A then B) logic operation and a corresponding truth table. (b) Applied gate and drain bias pulses. (c) Measured drain current for each step indicated in (a). $W \times L = 100 \times 1.5\ \mu m^2, T_{ox} = 275\ \mu m, frequency = 2\ Hz$.

(a) NIMP ($S = A \cdot B'$) Operation:
NIMP = Not IMP
Cycle 1: $V_G = 1, V_B = 0 \Rightarrow S = 0$
Cycle 2: $V_G = A, V_B = B \Rightarrow S = A \cdot B'$

Cycle 1	Cycle 2		
S	$V_G = A$	$V_B = B$	S
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0

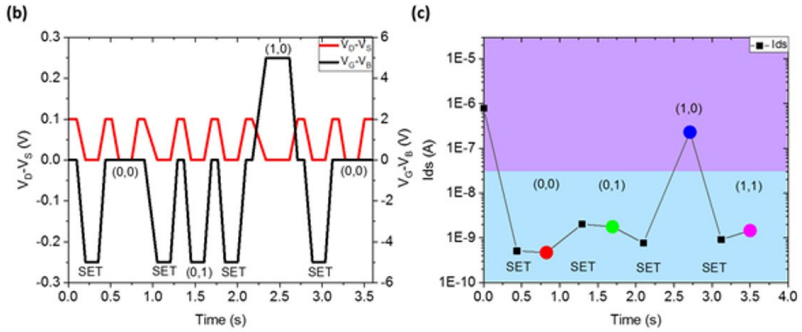


Fig 5.8 (a) Steps to perform the NIMP logic operation (S=A·B') and a corresponding truth table. (b) Applied gate and drain bias pulses. (c) Measured drain current for each step indicated in (a). $W \times L = 100 \times 1.5\ \mu m^2, T_{ox} = 275\ \mu m, frequency = 2\ Hz$.

(a) NAND ($S = (A \cdot B)' = A' + B'$):
Cycle 1: $V_G = 0, V_B = 1 \Rightarrow S = 0$ (RESET)
Cycle 2: $V_G = 1, V_B = A \Rightarrow S = A'$
Cycle 3: $V_G = 1, V_B = B \Rightarrow S = A' + B'$

Cycle 1	Cycle 2			Cycle 3		
S	$V_G = 1$	$V_B = A$	S	$V_G = 1$	$V_B = B$	S
0	1	0	1	1	0	1
0	1	0	1	1	1	1
0	1	1	0	1	0	1
0	1	1	0	1	1	0

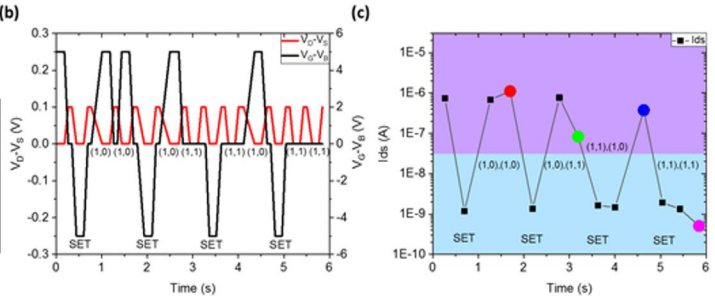
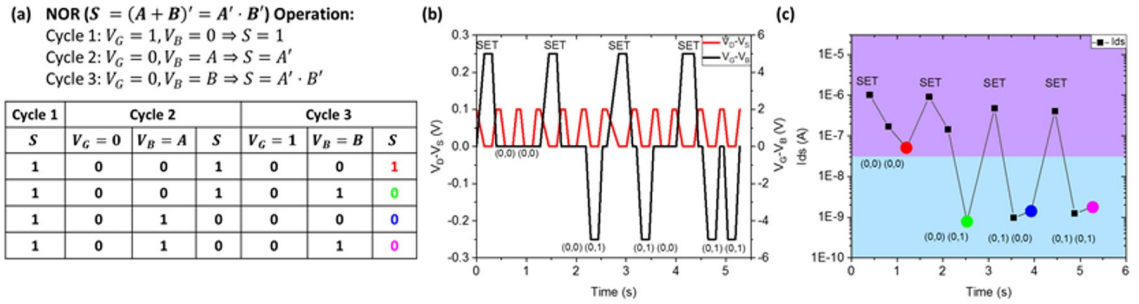


Fig 5.9 (a) Steps to perform a two-input NAND ($S = A' + B'$) logic operation and a corresponding truth table. (b) Applied gate and drain bias pulses. (c) Drain current measured for each of the steps shown in (a). $W \times L = 100 \times 1.5\ \mu m^2, T_{ox} = 275\ \mu m, frequency = 2\ Hz$.



Cycle 1	Cycle 2			Cycle 3		
S	$V_G = 0$	$V_B = A$	S	$V_G = 1$	$V_B = B$	S
1	0	0	1	0	0	1
1	0	0	1	0	1	0
1	0	1	0	0	0	0
1	0	1	0	0	1	0

Fig 5.10 (a) Steps to perform a two-input NOR ($S = A' \cdot B'$) logic operation and a corresponding truth table. (b) Applied gate and drain bias pulses. (c) Measured drain current for each step shown in (a). $W \times L = 100 \times 1.5\ \mu m^2, T_{ox} = 275\ \mu m, frequency = 2\ Hz$.

5.2.3 Benchmark

Table I summarizes the characteristics of state-of-the-art devices in comparison with our experimental results, which show a switching time of 250 ms at V_{GS} = 5 V. Magnetic tunnel junctions demonstrate the fastest switching, but OxRAM also switches on the nanosecond level, with energy scaling on the order of nanojoules. For reported non-filamentary ReRAMs, Table I presents evidence of a programming voltage–switching speed–current dilemma, especially for non-filamentary ReRAMs operating on the time scale of seconds or even hours [7], as current levels are reduced in the quest for truly neuromorphic systems [8]. Other non-filamentary devices scaled by a factor of 1000, as reported in [9], for example, have power consumption in the microwatt range. In comparison, the current device shows a promising low power density, despite having a 1.5-μm channel length and a common gate that extends over the entire chip, resulting in 100% overlap capacitance with the source/drain pads and underlying gate oxide.

A plot of our experimentally obtained switching time versus channel length (Fig 5.11 (a)) shows the lowest power consumption/unit area reported to date. However, despite its low power consumption, the device currently has a limited switching frequency, which is partially due to the large active area of the device. As shown in Fig 5.11 (b) for a fixed pulse amplitude, when the input pulse frequency increases, the on/off ratio between the LRS and HRS decreases.

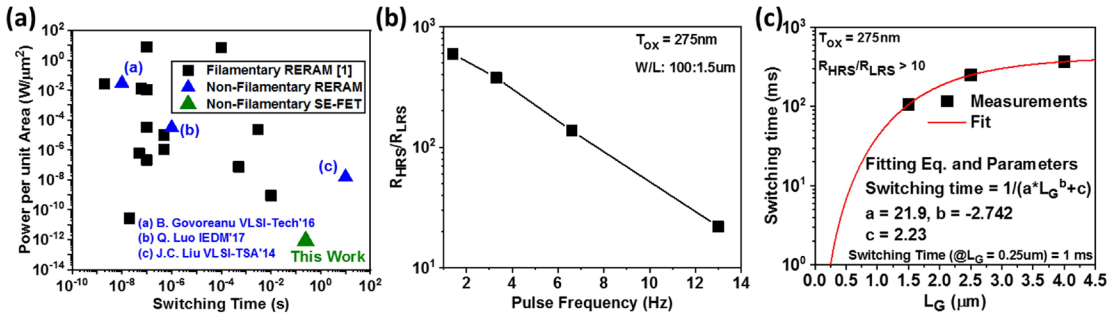


Fig 5.11 (a) Benchmarking the power consumption per unit device area vs. switching time based on experimental data reported in [1] and our current measurements. (b) Measured HRS/LRS ratio vs. input pulse frequency and (c) switching time vs. gate length L. The fit to data predicts a switching time of 150 ns for L = 20 nm.

In context of the compute in memory application, as a result of the device current dimension, the operation speed of the device is limited. However, the switching speed can be further optimized by scaling the channel length, as shown in Fig 5.11 (c). This result occurs because ion accumulation is expected to initiate from the overlapping regions under the drain/source contacts and subsequently spreads to the center of the channel. Although the switching time is relatively poor at 100 ms for the current device with a 1.5-μm channel length, by applying our measured data as a function of channel length to extrapolate to dimensions similar to those of contemporary ReRAMs, the potential of a 150-ns switching

time for a 20-nm channel length is well within reach of a vertical crossbar array if nanowires are used for this device, as shown in Fig 5.11 (c). However, scaling in the vertical direction is more challenging and must be addressed via control of the diffusivity and ion concentration in the insulator.

TABLE I
COMPARISON OF THE Ta_2O_5/ZnO-BASED SE-FET WITH STATE-OF-THE-ART TWO-TERMINAL DEVICES REPORTED TO DATE. DESPITE A LARGER PULSE WIDTH, THE MEASURED POWER CONSUMPTION DURING TRANSITION IS ON THE ORDER OF NANOWATTS, ACHIEVED BY WRITING IN THE OFF-STATE IN THE SE-FET.

Description	**R_{LRS} (kΩ)**	**R_{HRS}/R_{LRS} (10^x)**	**Switching time**	**Pulse amp. (V)**	**Energy consumption**	**Approx. power in a transition**	**Ref.**
Al_2O_3/aSi resistive switch (RS)	3.6×10^6	2	10 ms	3.6	36 pJ	3.6 nW	[8]
HfO_2 complementary RS (CRS)	$10 - 100$	3	0.5μs	1	0.5 nJ	1 mW	[10]
Ta_2O_5 CRS	4	2–3	3 ms	2.4	3 μJ	0.6 mW	[11]
HfO_X/AlO_Y RS	1	2	100 ns	1.2/-1.8	0.324 nJ	3.24 mW	[12]
Magnetic tunnel junction	6.4	0.5	2 ns	0.8	0.067 pJ	33.5 μW	[13]
GeTe RS	10	3–4	60 ns	1.2	498 pJ	8.3 mW	[14]
Amorphous In-Zn-Sn-O RS	12	2	100 ns	3	75 pJ	0.75 mW	[15]
Nanohole graphene/HfO_2 RS	1	6	100 ns	4	1.6 nJ	16 mW	[16]
TiO_{2-x} + Pt nanocrystal RS	2×10^4	3	--	1	--	1 nW	[17]
Graphene/2D perovskite RS	2.8×10^8	2	--	2.8	--	28 pW	[18]
HfO_2 OxRAM and GeS_2/HfO_2 CBRAM	$1 - 10$	1 6	100 ns 100 μs	1 2.3	0.1 nJ 0.36 μJ	1 mW 3.6 mW	[19]
Ta_2O_5/ZnO SE-FET	**1 $M\Omega$**	**3**	**250 ms**	**5**	**8.1 nJ**	**32.4 nW**	**This work**

5.3 Using synaptic properties of SE-FETs for supervised learning in a crossbar

5.3.1 Crossbar implementation of the SE-FET

The supervised learning process of the SE-FET crossbar has been simulated using the MATLAB Simulink model outlined in Chapter 4. The schematic representation of a 2x2 crossbar, which includes four three-terminal Ta_2O_5/ZnO SE-FETs in combination with four MOSFETs, is depicted in Fig 5.12. This simulation allows for initial analysis of the device's behaviour, including factors such as the impact of specific configurations and biasing conditions on its performance. Moreover, by using this simulation approach, we can anticipate the device's responses to certain stimuli and further understand how it might be effectively integrated into larger systems or network configurations.

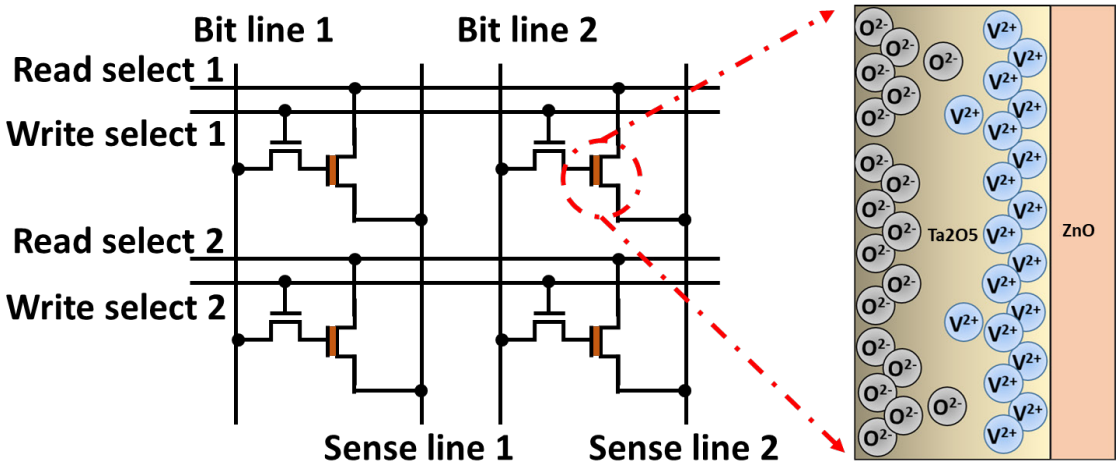


Fig 5.12 Schematic of a 2 × 2 crossbar consisting of an SE-FET and a depletion-mode MOSFET with two control lines for selecting individual cells and the Ta_2O_5/ZnO SE-FET.

Each SE-FET in the crossbar array is controlled by a depletion-mode MOSFET, with its drain-source terminal connected to the gate of the SE-FET. The 'write-select' and 'read-select' lines activate the corresponding functionality in the selected device within the crossbar.

During the read process, the output current from each device is added together at the sense line of each column. This summed output is then fed into an activation function, which is mathematically represented as $y_i = f_a(\sum_j w_i x_j)$. In this equation, f_a denotes a sigmoid activation function, given by:

$$f_a(\sum_j w_i x_j) = 2 \times \frac{1}{1+e^{(a\times(\sum_j w_i x_j - c))}} \quad \text{(Eq 5.2)}$$

Here 'a' is -1×10^{-6} and 'c' is 2×10^{-6}. After applying the activation function, the output is compared to the desired output to compute an error term δ_i. This error is used in the adjustment of the conductance of each device based on the delta rule. This is a fundamental concept in learning algorithms, where the weights (in this case, conductances) are adjusted in the direction that reduces the error. The delta rule is given by: $w_{ij} \leftarrow w_{ij} + \eta\delta_i x_j$ [20]. Here η is the learning rate (set to 1 in this case), δ_i is the error term, and x_j is the input. During the write process, the conductance update is performed with a gate voltage proportional to the error, where the sign of the gate voltage corresponds to the sign of the error.

The bias conditions of the crossbar array for read and write operations are provided in Fig 5.13 (a). Also, the truth table for OR and AND operations and the corresponding requirements on the sense line are depicted in Fig 5.13 (b). These diagrams and tables

would help in visualizing the operations and understanding the requirement of signal conditions for different logic operations.

(a)

	WS	RS	BL	SL
Write Select	0	0	V_{Write}	0
Write Unselect	V_{th}	0	V_{Write}	0
Read Select	0	V_{Read}	V_m	$f(\Sigma I_{SL})$
Read Unselect	0	0	V_m	$f(\Sigma I_{SL})$

(b)

X	Y	Logic	SL
0	0	OR: 0	$\Sigma I_{SL}<2uA$
		AND: 0	$\Sigma I_{SL}<2uA$
0	1	OR: 1	$\Sigma I_{SL}\geq 2uA$
		AND: 0	$\Sigma I_{SL}<2uA$
1	0	OR: 1	$\Sigma I_{SL}\geq 2uA$
		AND: 0	$\Sigma I_{SL}<2uA$
1	1	OR: 1	$\Sigma I_{SL}\geq 2uA$
		AND: 1	$\Sigma I_{SL}\geq 2uA$

Fig 5.13 (a) Bias conditions during the read and write operations. Here, V_{th} is the threshold voltage of the MOSFET, V_m is the refresh voltage (0.5 V), $f(\sum I_{SL})$ is the output of the activation function (sigmoid) (Eq. (5.2)), V_{Write} is the write voltage for potentiation (0 V + 4 V), and the read voltage is V_{Read} = 0.2 V. (b) Truth table for OR and AND and the corresponding condition of the sense line.

5.3.2 Synaptic properties of SE-FETs and supervised learning

The variation in channel conductance, subject to the application of gate bias pulses of both positive and negative polarities, has been experimentally evaluated and is demonstrated in Fig 5.14. This steady modulation in channel conductance exhibits characteristics analogous to the behavior of numerous reported memristor devices, underlining the fundamental mechanism of data storage in these systems through controlled shifts in conductance.

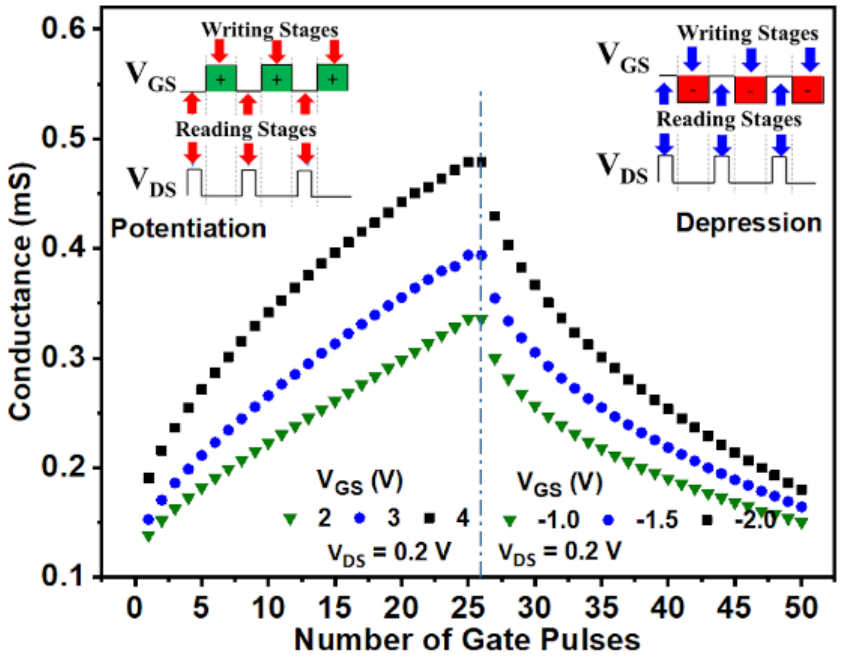


Fig 5.14 Conductance measured after sequential gate bias pulses of different amplitudes and polarities, as shown in the inset.

5.3.3 Crossbar arrangement of the SE-FET

The modelling results obtained from MATLAB Simulink simulation are illustrated in Fig 5.15. This figure represents the functionality of the device array under consideration, with devices in the first column (devices 1x and 1y in Fig 5.15 (a) and (b)) being trained to execute OR operations, while devices in the second column (devices 2x and 2y in Fig 5.15 (c) and (d)) are instructed to execute AND operations. The relevant inputs (X and Y) along with the resulting logical output from the crossbar array are depicted in Fig 5.15 (e).

In the initial state, all devices are configured in a low-conductance state, resulting in a diminished current reading at the sense line. Consequently, this is inadequate to match the desired reading current of 2 μA. This discrepancy triggers the generation of an error signal δ_{1i} for devices in the first column when inputs 01, 10, and 11 are presented, whereas

for the second column, an error signal δ_{2i} is generated only when input 11 is encountered. This prompts the application of a positive bias at the bit line to update the conductance of the associated device.

In response to this, an increase in the conductivity of the respective device is observed. Eventually, this leads to conductance levels for w_{1x} and w_{1y} surpassing 10 μS and for w_{2x} and w_{2y} stabilizing within the range of 5-10 μS. The net result is the achievement of expected logical OR and AND operation behaviors within five correction cycles.

After a duration of 20 continuous reading cycles, as highlighted by the green and brown squares in Figure 5.15, it is evident that the devices require a refresh cycle owing to ionic diffusion. This is to ensure the maintenance of the desired level of output. In general, the application of a single pulse of 0.5 V suffices for this refresh operation, since the weight of each device is observed to be relatively stable, remaining close to the conductance level necessary for accurate logical operation.

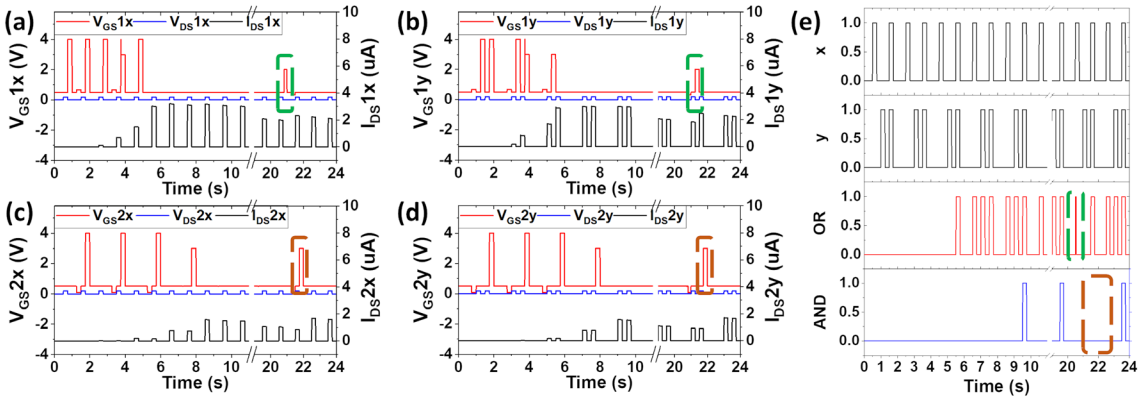


Fig 5.15 Simulated bias and read currents of devices (a) 1x and (b) 1y in the crossbar during training for OR operation and devices (c) 2x and (d) 2y for AND operation.

(e) Simulated output of the activation function for OR and AND logic operations, with respect to inputs X and Y. Errors at 20 and 22 s due to volatility are refreshed by application of a gate voltage of 2–3 V.

5.4 Conclusions

To conclude, we have conducted an in-depth examination of a three-terminal SE-FET's performance, a device with the ability to integrate memory and logic functions within a single transistor. Leveraging the non-filamentary mechanism and the inherent memory of this device, we demonstrated two-state logic operations, revealing a significant difference—equivalent to at least three orders of magnitude—between the High Resistance State (HRS) and the Low Resistance State (LRS).

Through computational simulations, we further highlighted the device's potential by implementing a basic 2 x 2 crossbar array designed for neuromorphic learning of the AND and OR logic operations, using the SE-FET as the core element. Despite its relatively large physical size and high operating voltage, the device maintains an impressively low power consumption, operating within the range of nanowatts per transition. This efficiency is primarily due to the innovative concept of off-state logic, indicating potential future applications of the SE-FET in power-efficient computation.

Owing to large physical size, currently the device still operating on 100s of milliseconds switching speed. Though the low speed could a pose a limit on application such as compute in memory, where the switching speed are important. For other application such as spike sorting in neural (or brain–computer) interfaces, where a limited communication bandwidth and energy, our potentially flexible platform technology is ideally suited for hardware-implementable spike recording and feature extraction, as it does

not require historical data storage, at rates of a few 100 kHz [21][22], even volatile memory has been shown to serve this purpose [23]. Due to its inherent diffusion-based switching mechanism, the SE-FET may be better suited for these types of applications.

5.5 References

[1] X. Wong *et al.*, "Stanford Memory Trends," *https://nano.stanford.edu/stanford-memory-trends*.

[2] E. J. Fuller *et al.*, "Li-Ion Synaptic Transistor for Low Power Analog Computing," *Adv. Mater.*, vol. 29, no. 4, pp. 1–8, 2017.

[3] W. Kim *et al.*, "Impact of oxygen exchange reaction at the ohmic interface in Ta2O5-based ReRAM devices," *Nanoscale*, vol. 8, no. 41, pp. 17774–17781, 2016.

[4] I. Daniele, "Resistive switching memories based on metal oxides: mechanisms, reliability and scaling," *Semicond. Sci. Technol.*, vol. 31, no. 6, p. 63002, 2016.

[5] E. Linn *et al.*, "Beyond von Neumann - Logic operations in passive crossbar arrays alongside memory operations," *Nanotechnology*, vol. 23, no. 30, 2012.

[6] J. Borghetti *et al.*, "'Memristive' switches enable 'stateful' logic operations via material implication," *Nature*, vol. 464, no. 7290, pp. 873–876, Apr. 2010.

[7] I. Sugiyama *et al.*, "A nonvolatile memory device with very low power consumption based on the switching of a standard electrode potential," *APL Mater.*, vol. 5, no. 4, 2017.

[8] J. Zhou *et al.*, "Very low-programming-current RRAM with self-rectifying characteristics," *IEEE Electron Device Lett.*, vol. 37, no. 4, pp. 404–407, 2016.

[9] B. Govoreanu *et al.*, "Advanced a-VMCO resistive switching memory through inner interface engineering with wide (>10^2) on/off window, tunable μA-range switching

current and excellent variability," in *2016 IEEE Symposium on VLSI Technology*, 2016, vol. 2016-Septe, no. 8, pp. 1–2.

[10] T. Breuer *et al.*, "A HfO2-Based Complementary Switching Crossbar Adder," *Adv. Electron. Mater.*, vol. 1, no. 10, pp. 1–10, 2015.

[11] T. Breuer *et al.*, "Realization of Minimum and Maximum Gate Function in Ta2 O5 -based Memristive Devices," *Sci. Rep.*, vol. 6, no. April, pp. 1–9, 2016.

[12] P. Huang *et al.*, "Reconfigurable Nonvolatile Logic Operations in Resistance Switching Crossbar Array for Large-Scale Circuits," *Adv. Mater.*, vol. 28, no. 44, pp. 9758–9764, 2016.

[13] A. Jaiswal *et al.*, "In-situ, In-Memory Stateful Vector Logic Operations based on Voltage Controlled Magnetic Anisotropy," *Sci. Rep.*, vol. 8, no. 1, pp. 1–12, 2018.

[14] M. Jin *et al.*, "Recon fi gurable logic in nanosecond Cu / GeTe / TiN filamentary memristors for energy- ef fi cient in-memory computing," 2018.

[15] B. C. Jang *et al.*, "Memristive Logic-in-Memory Integrated Circuits for Energy-Efficient Flexible Electronics," *Adv. Funct. Mater.*, vol. 28, no. 2, pp. 1–10, 2018.

[16] X. Zhao *et al.*, "Confining Cation Injection to Enhance CBRAM Performance by Nanopore Graphene Layer," *Small*, vol. 13, no. 35, 2017.

[17] M. Tsigkourakos *et al.*, "Ultra-Low Power Multilevel Switching with Enhanced Uniformity in Forming Free TiO2−x-Based RRAM with Embedded Pt Nanocrystals," *Phys. Status Solidi Appl. Mater. Sci.*, vol. 214, no. 12, pp. 1–5, 2017.

[18] H. Tian *et al.*, "Extremely Low Operating Current Resistive Memory Based on Exfoliated 2D Perovskite Single Crystals for Neuromorphic Computing," *ACS Nano*, vol. 11, no. 12, pp. 12247–12256, 2017.

[19] E. Vianello *et al.*, “Resistive Memories for Ultra-Low-Power embedded computing design,” *Tech. Dig. - Int. Electron Devices Meet. IEDM*, vol. 2015-Febru, no. February, pp. 6.3.1-6.3.4, 2015.

[20] J. Hertz et al., Introduction to the Theory of Neural Computation, vol. 1. 1991.

[21] H. G. Rey *et al.*, “Past, present and future of spike sorting techniques,” *Brain Res. Bull.*, vol. 119, pp. 106–117, Oct. 2015.

[22] D. Kuzum *et al.*, “Synaptic electronics: materials, devices and applications,” *Nanotechnology*, vol. 24, no. 38, p. 382001, Sep. 2013.

[23] I. Gupta *et al.*, “Sub 100 nW volatile nano-metal-oxide memristor as synaptic-like encoder of neuronal spikes,” *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 2, pp. 351–359, 2018.

[24] M. Payvand *et al.*, “Dendritic Computation through Exploiting Resistive Memory as both Delays and Weights,” *Proc. Make sure to enter correct Conf. title from your rights confirmation emai (Conference Acron. ’XX)*, vol. 1, no. 1, pp. 1–4, May 2023.

[25] W. Maass, “Networks of spiking neurons: The third generation of neural network models,” *Neural Networks*, vol. 10, no. 9, pp. 1659–1671, Dec. 1997.

CHAPTER 6

Reservoir computing using SE-FET as Reservoir

6.1 Introduction

In this chapter, we present a physical RC framework utilizing a three-terminal SE-FET as a reservoir, a project undertaken in collaboration with IIT Roorkee. My contribution consists of the design and execution of the experimental work, while Ankit Gaurav from IIT Roorkee carried out the reservoir output generation for the benchmark task using the experimental measurement, and the software based reservoir computing network training and system evaluation. By combining both experimental data and simulation results, we demonstrate a novel approach to improve learning efficiency based on desirable nonlinear dynamics and volatile memory of the SEFET. Our approach can augment ongoing efforts towards creating more energy-efficient and powerful computing systems. Further details about our experimental setup, results, and their implications are discussed in the sections that follow.

6.2 Reservoir computing using SE-FET

The structure of the solid electrode field effect transistor (SE-FET) based reservoir system is illustrated in Fig 6.1 [1]. The initial step involves processing the input information and converting it into a binary image. Subsequently, this binary image is transformed into a series of sequential voltage inputs [2]. The connection from the input layer to the reservoir, denoted as $u_n(t)$, remains fixed for each training and test set. Unlike in software-based recurrent neural networks, the reservoirs in this SE-FET based reservoir framework are not interconnected neural networks but rather simple SE-FET devices. Here,

the temporal input is processed through the inherent diffusion-based memory of the SE-FET.

In order to effectively process the sequential input and function properly as a reservoir, the reservoir state must fulfill certain criteria. First, the device must provide sufficient conductance state variation and memory depending on the input function $u_n(t)$, (or write voltage) and the sequence length of the sequential input. This is crucial to prevent information loss due to conductance state saturation or memory loss and to facilitate higher dimensionality where both current and past input can be stored [3]. Second, the input function must be captured through the variation of the conductance state and the sequence of the input processed through the time-dependent memory loss (or fading memory) [4]. Weak or no memory decay could result in the reservoir being unable to distinguish the sequence of the input. This short term memory decay is significant for representing sequential data with strong time dependencies [5].

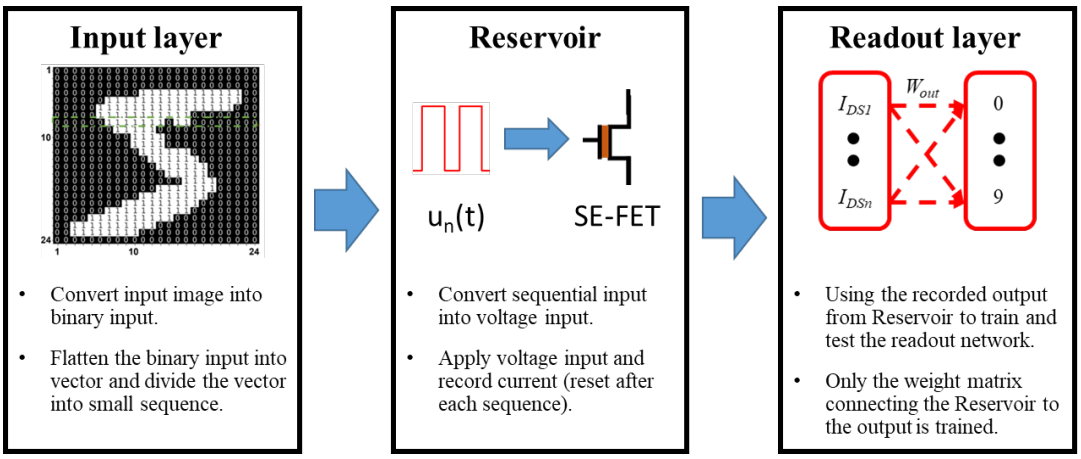


Fig 6.1 Scheme of the SE-FET-based reservoir framework [1].

In this work, we utilize the gate of the SE-FET as the input terminal for the reservoir. The transfer function between the input and the reservoir is denoted as $u_n(t)$. A reading voltage is supplied to the drain-source terminal to measure the conductance, defined as Idsn. The primary advantage of this three-terminal device is its ability to segregate reading and writing operations, thereby providing the potential to continuously read the SE-FET's response during sequential input, without the necessity for downsampling the input.

For this application, a modest reading voltage of -1V is supplied to the drain-source in order to read the device's conductance state (Fig 6.2). This careful selection of voltage is due to the bottom-gate structure of the device. As demonstrated in Fig 6.3, if a high positive reading voltage is applied during the off state, it effectively places the gate at a relatively low potential, which in turn shortens the retention time of the device's memory. On the other hand, applying a high negative voltage increases the gate's relative potential, risking unintentional writing onto the device during the reading process.

Therefore, a reading voltage of -1V has been carefully chosen as it helps maintain a readable memory state in the device without inadvertently triggering a writing action. This ensures that the device operates within its optimal conditions and that the reading process doesn't interfere with the device's memory state.

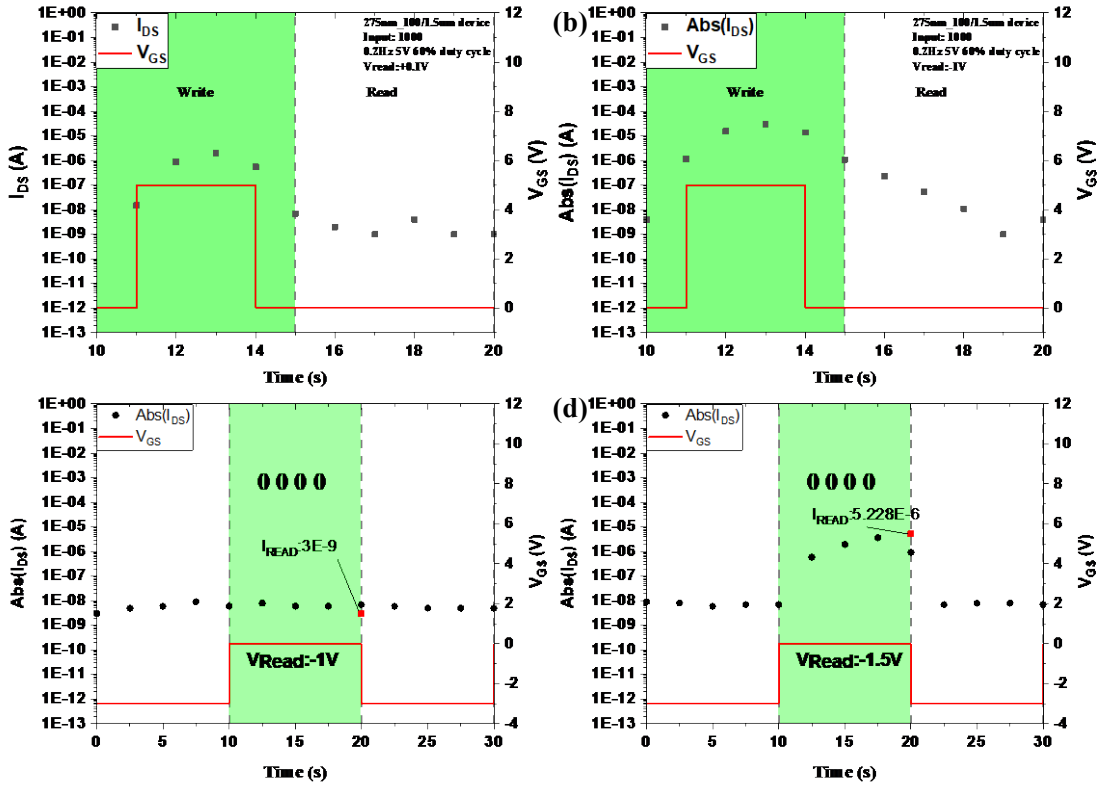


Fig 6.2 Read current of the SE-FET following a single pulse using +0.1V (a) and -1.0V (b) of read voltage. The -1.0V read voltage helps to maintain a more linear and enriched decay state, without offsetting the decay or writing to the devices. The read current of the SE-FET after reset with 0V gate votlage using -1V (c) and -1.5V (d) read voltage is also shown. Utilizing -1V read voltage keeps the device in the off state

after a reset without additional input from reading itself, while a -1.5V read voltage is sufficient to force the device into a potentiation state by the negative read voltage.

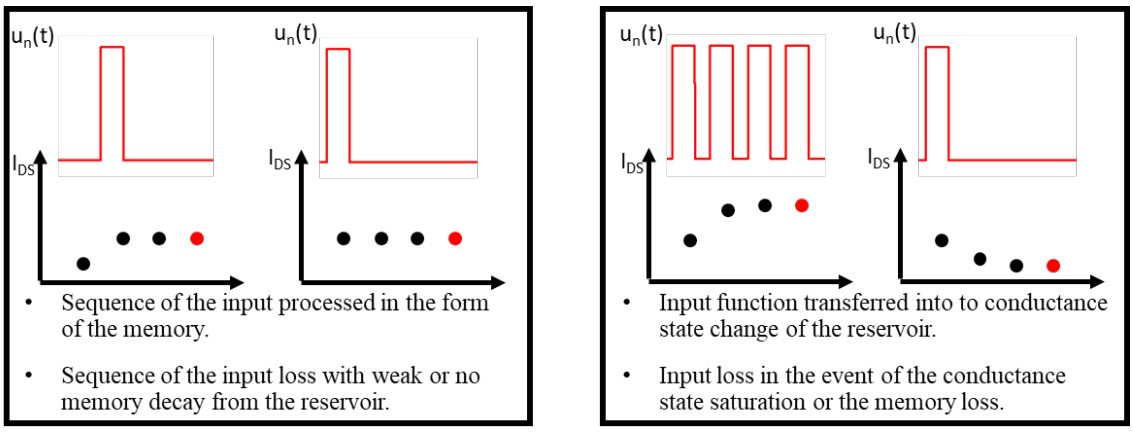


Fig 6.3 Conductance state saturation and complete memory loss results in information between input function and reservoir. Weak or no memory loss results a lost in temporal order of the sequential input.

In this experiment, we fed all potential input into multiple distinct SE-FET devices, and the output current from each device was recorded experimentally and stored independently. This stored data, encapsulating all potential outputs from each reservoir, serves as the database for training and testing the readout network, circumventing the need for physical measurements during these processes. In this setup, the readout network that connects the reservoir states to the output layer is the only layer subjected to training.

The training and testing of the readout are executed offline by Ankit Gaurav from IIT Roorkee, employing a software-based logistic regression model. He opted for a one-versus-rest scheme, a typical solution for image recognition tasks, which reconfigures the multi-class dataset into multiple binary classification problems. The class index that yields the highest probability is then selected as the system's final output.

The probabilities of the possible outcomes in a single trial are mapped using a sigmoid function, a common tool in both CNN and RNN networks:

$$f_p(x) = \frac{L}{1+e^{-k(x-x_0)}} \quad \text{(Eq 6.1)}$$

Here, L is the maximum output of the sigmoid function, k represents the logistic population growth of the curve, and x0 is the midpoint of the sigmoid function.

The implementation by Ankit Gaurav employed a gradient descent liblinear solver from Python's scikit-learn library for the Logistic Regression. To prevent overfitting to specific data selections, a 7-fold cross-validation strategy to train and test the network was implemented. Specifically, an image set of 70,000 images was used for training and testing. Each round, a total of 60,000 image samples from this set were used to train the readout function. After training, a separate sample set containing the remaining 10,000 images, which were not included in the training, was fed into the network for testing. Post-testing, these image sets are reassigned into 60,000 training images and 10,000 testing images. The readout function was then retrained and tested 7 times.

The raw images from the MNIST database was converted and processed into uniformly sized 24 × 24 binary images before being fed into the reservoir. Classification was performed based on the predicted output of the SE-FET, derived from the measured and stored reservoir dataset from multiple distinct devices.

6.3 Results and Discussion

In our experiment, we investigated the writing frequency of the SE-FET across a frequency range from 0.1Hz to 1Hz with 4V 60% duty cycle pulses. As demonstrated in Fig 6.4, for an identical total pulse duration of 30s (60% duty cycle pulses for 50s), it became clear that higher frequency pulses were more efficient at potentiating the devices

into a higher conductance state. With 0.1 Hz pulses, the device conductance increases from 2E-8S to 2E-4S after 50s, while with 1HZ pulses, the maximum conductance increases to 4.6E-4S within the same duration.

However, for a 4-bit sequence length, the conductance after 4 pulses remained below 1E-5A, posing a challenge to accommodate the necessary states for a 4-bit sequence (as a 4-bit binary sequence requires 16 readable conductance states). At 0.4Hz, the conductance state reaches 3.5E-5S after 4 pulses, providing enough room for a 4-bit sequence length.

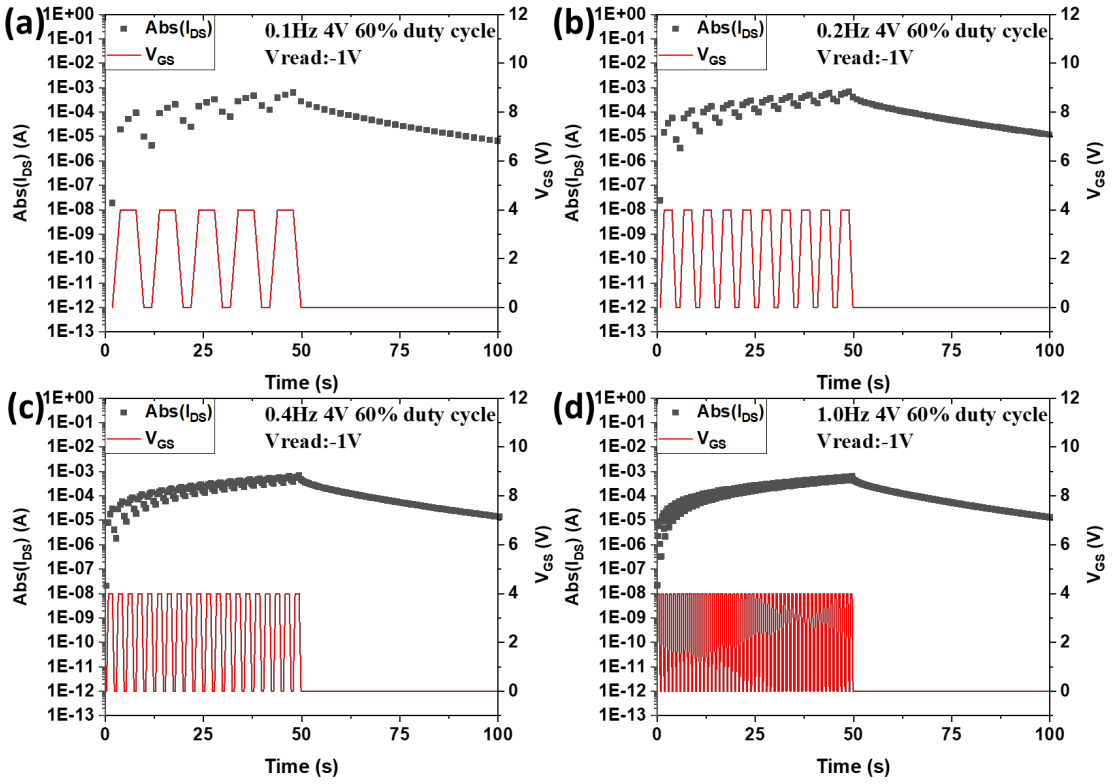


Fig 6.4 Potentiation conductance of the SE-FET at 0.1 (a), 0.2(b) 0.4(c) and 1Hz (d).

The decay of the conductance state and the change in conductance state per pulse record are plotted in Fig 6.5. Starting from 3.5E-5S, the device conductance state decays to 1E-7 level after 88 pulses (~220s). As the conductance state decays, the change in conductance state per pulse also reduces after each pulse, ranging from 5E-6 per pulse to below 1E-8 level after 70 pulses. After this point, the conductance state change per pulse becomes relatively noisy.

Considering the minimum conductance state required to process an x bit input with a sequential length of n equal to xn, with the potentiation and the decaying state of the SE-FET, the device, though unoptimized, has sufficient retention time and a sufficient number of distinguishable conductance states to reflect the temporal order of a 4 to 5 binary input. Given the retention time after potentiation, a reset is required between the sequences to avoid repeated potentiation and to maintain a consistent initial state.

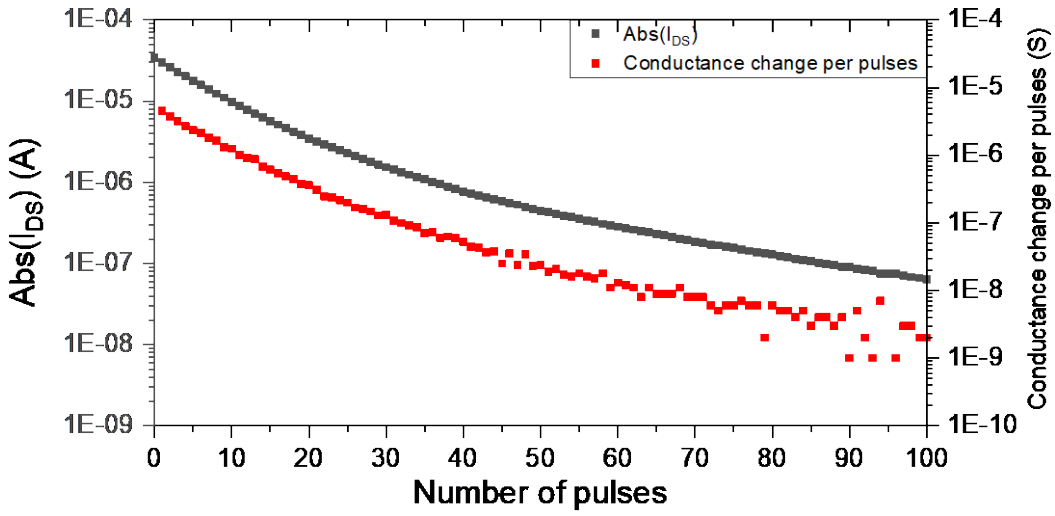


Fig 6.5 Conductance state decay and conductance state change per pulses record at 0.4 Hz frequency pulses.

In our experiment, the image was divided into 6 sub-sections, each consisting of 4 pixels. This 4-pixel sequence necessitated the devices to process a 4-bit binary input with 16 conductance states. Fig 6.6 illustrates the response of the SE-FET when subjected to all 16 combinations of temporal inputs, where the writing voltage is set to 3V for a binary input of 1 and 0V for 0. Fig 6.7 (data and results generated by Ankit Gaurav) presents an example of the output heatmap from the reservoir's response to digit 5 and a confusion matrix showcasing the experimentally obtained classification results of the SE-FET-based reservoir versus the correct outputs [6].

By converting the input into a higher dimensional current output using the SE-FET, we achieved an average accuracy of 91.19% across a 7-fold cross-validation of the test set, an improvement over the 90.82% accuracy achieved without the SE-FET. This proves that the SE-FET-based reservoir is capable of mapping binary input into a higher dimensional space in the form of the output current.

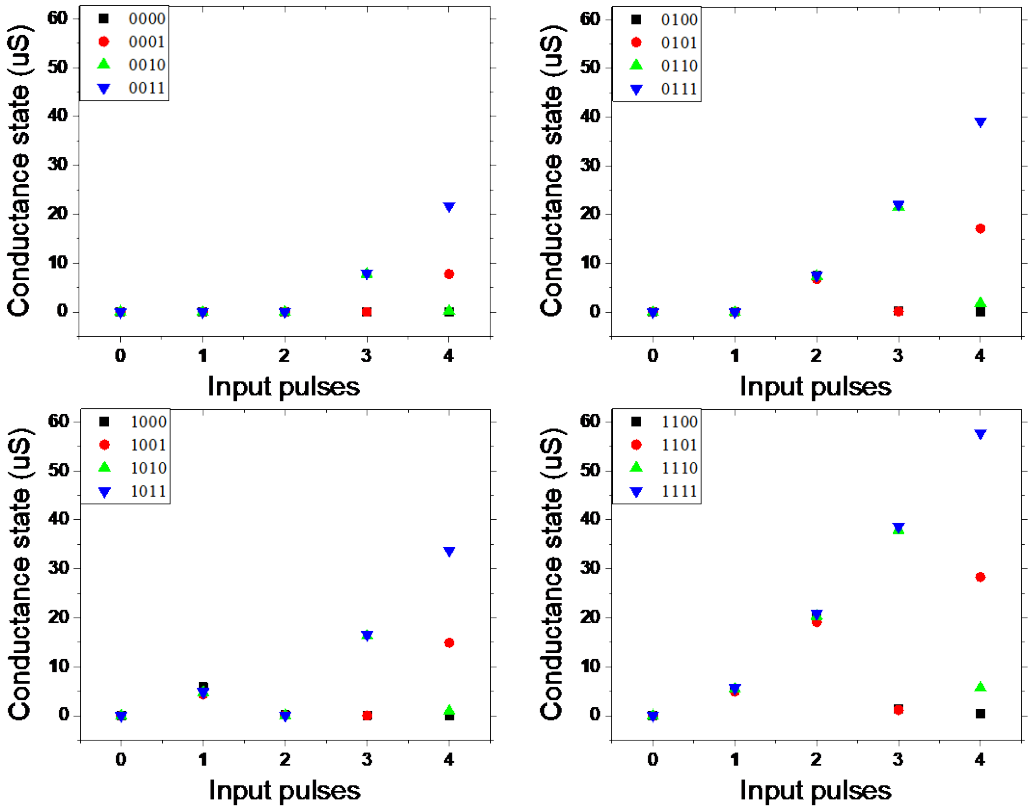


Fig 6.6 Example of the recorded SE-FET output from one of the devices used as the reservoirs output database to train and test the network. Devices recorded at 0.4Hz with 3V 60% duty cycle pulses.

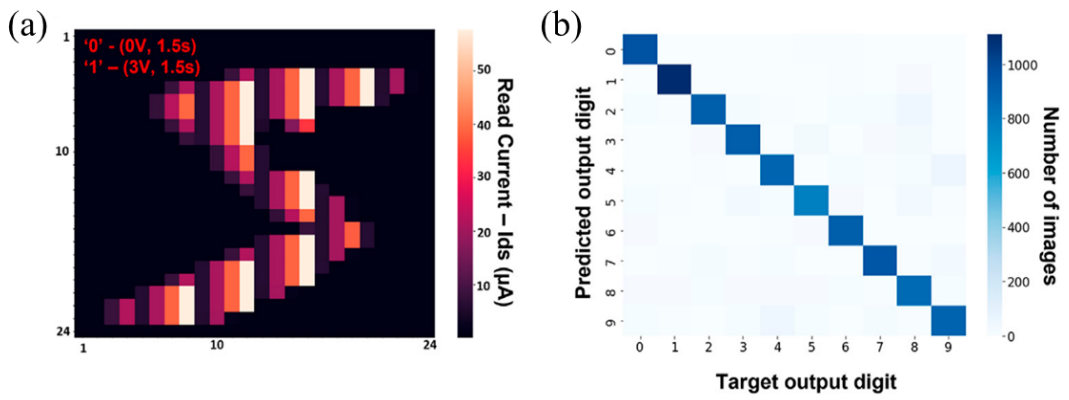


Fig 6.7 (a) Example of the heat map from the SE-FET response to digit 5 using data based on the measured results showing Fig 6.6. (b) Confusion matrix showing the experimentally obtained classification results of the SE-FET-based reservoir versus the correct outputs. Ankit Gaurav achieved an overall recognition rate of 91.19% using 7-fold cross-validation [6].

Typically, the inevitable device-to-device variation is not a desired characteristic for neuromorphic hardware. However, in the context of Reservoir Computing, a degree of device-to-device variation could actually be beneficial in certain scenarios. We tested the network using 3 different SE-FETs as the reservoirs, running in parallel. The final output with the highest probability is then used as the system's final output. The experiment setup was slightly adjusted to compensate for the devices' weaker performance. The image was fed into the system using a 3-bit binary input, with the writing voltage set to 3.5V for a binary input of 1 and 0.8V for 0. The conductance change of all three SE-FETs when these devices are subject to the same input is shown in Fig 6.8.

Due to device-to-device variation, when all three devices receive the same input of 111, the device conductance varies from 8.9uS to 27uS after the input. This level of difference could be problematic for CNN, as the conductance state is used directly for logic

operations. However, as a reservoir, such a difference in conductance state creates a more dynamic representation of the input. With three reservoirs running in parallel, Ankit Gaurav reported an improved mean recognition of 92.97%, an improvement over the 91.19% achieved using a single device. If the input image is converted into both horizontal and vertical sequences, an overall mean recognition rate of 94.97% was achieved. However, this does lead to an increased number of inputs that need to be processed.

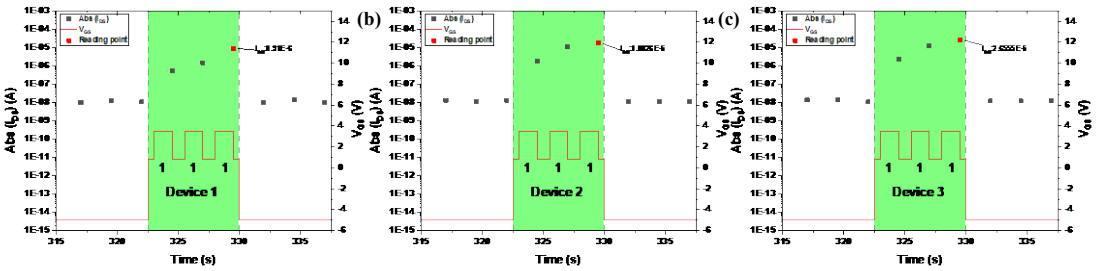


Fig 6.8 Response of device 1 (a), 2 (b) and 3 (c) when subjected to same input 111 with same 3.5V pulses. The devices conductivity varies due to inherent device-to-device variation.

These observations suggest that the variation between different reservoirs can also be manipulated and controlled by adjusting the input function. The network was tested in a different configuration with increased device-to-device variation by using distinct input functions for three different devices. A consistent base voltage of 0.8V was set for a binary input of 0, and the input voltages of 3.5V, 4.5V, and 5.5V were set for a binary input of 1 for devices 1, 2, and 3, respectively. An example of the recorded output from all three devices using their intrinsic device-to-device variation is shown in Fig 6.8, and the variation under input is plotted in Fig 6.9.

With these distinct input functions, the conductance state variation increased to between 8.9uS and 55uS after the same input of 111. These differences in the conductance state led to increased reservoir-to-reservoir output variation. When these three reservoirs were configured in the same parallel network, an improved mean recognition of 92.44% was achieved after training. The network using this configuration demonstrated improved recognition accuracy over the system using a single SE-FET but was slightly less accurate than the system operating with the SE-FET's inherent device-to-device variation, which had an accuracy of 92.97%.

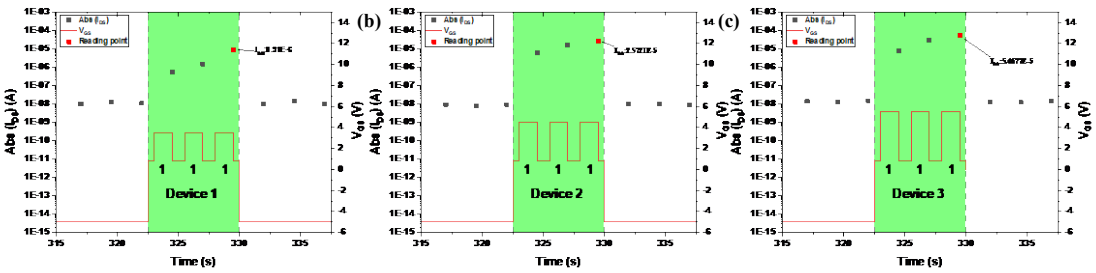


Fig 6.9 Response of device 1 (a), 2 (b) and 3 (c) when subjected to input 111 with 3.5V, 4.5V and 5.5V pulses. The devices conductivity varies due to inherent device-to-device variation and the different input function.

These results indicate the potential advantages of implementing a three-terminal SE-FET in a reservoir computing (RC) framework. Compared to a conventional network without the SE-FET, which exhibited an accuracy rate of 90.82%, the SE-FET-based RC network demonstrated a higher recognition accuracy of 91.19% when using a single device. Notably, this accuracy could be further improved to 94.44% by leveraging the inherent device-to-device variation and using three different SE-FETs as reservoirs.

The performance of the SE-FET-based RC network is the highest compared to other examples of physical RC networks that have utilized devices with diffusion-based mechanisms for the task of image processing. Most of these examples have used two-terminal memristors with fading memory. However, due to the two-terminal nature of these devices and the need for down sampling during processing, the recognition rates reported for these networks have been relatively low, at 88.1% [7] and 83% [8].

Overall, these results underscore the potential of SE-FET-based RC networks to outperform more traditional methods, demonstrating their viability and efficacy for image recognition tasks. Further optimization of the SE-FET device could potentially lead to even better results, making this a promising direction for future research in neuromorphic computing.

6.4 Conclusions

In collaboration with IIT Roorkee, we have experimentally demonstrated a three-terminal SE-FET-based reservoir network that exhibits enhanced learning efficiency of 95%. These devices are capable of converting the spatiotemporal dependencies of inputs into higher-dimensional outputs, leveraging the ionic migration intrinsic to their fading volatile memory. Our dynamic SE-FET-based reservoir computing (RC) system surpasses the performance of traditional networks, achieving an improvement of 3.62% in the same handwritten digit recognition task, and it does so without any reliance on down-sampling. The inherent diffusion-based mechanism of these devices allows the reservoir network to handle linearly non-separable problems, mapping input signals into a higher-dimensional space through a nonlinear relationship. Unlike the applications presented in Chapter 5, the SE-FET-based reservoir network does not necessitate a long-term or non-volatile memory.

Thus, devices with thinner oxide layers are no longer a constraint. The rapid writing and swift fading memory could potentially enhance the device's capability when processing inputs within an RC system. This work lays a solid foundation for the future development and optimization of SE-FET-based reservoir networks, promising to significantly advance the field of hardware-based reservoir computing.

6.5 References

[1] A. A. Ferreira *et al.*, "An approach to reservoir computing design and training," *Expert Syst. Appl.*, vol. 40, no. 10, pp. 4172–4182, Aug. 2013.

[2] C. Du *et al.*, "Reservoir computing using dynamic memristors for temporal information processing," *Nat. Commun.*, vol. 8, no. 1, p. 2204, Dec. 2017.

[3] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Comput. Sci. Rev.*, vol. 3, no. 3, pp. 127–149, Aug. 2009.

[4] L. Appeltant *et al.*, "Information processing using a single dynamical node as complex system," *Nat. Commun.*, vol. 2, no. 1, p. 468, Sep. 2011.

[5] H. Jaeger and H. Haas, "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication," *Science (80-.).*, vol. 304, no. 5667, pp. 78–80, Apr. 2004.

[6] A. Gaurav *et al.*, "Reservoir Computing for Temporal Data Classification Using a Dynamic Solid Electrolyte ZnO Thin Film Transistor," *Front. Electron.*, vol. 3, no. April, pp. 1–9, Apr. 2022.

[7] M. Hoffmann *et al.*, "Dynamic modeling of hysteresis-free negative capacitance in ferroelectric/dielectric stacks under fast pulsed voltage operation," *Device Res. Conf. - Conf. Dig. DRC*, vol. 2019-June, no. 2011, pp. 97–98, 2019.

[8] R. Midya *et al.*, “Reservoir Computing Using Diffusive Memristors,” *Adv. Intell. Syst.*, vol. 1, no. 7, p. 1900084, Nov. 2019.

CHAPTER 7

Seed-assisted method for <202>-oriented $CH_3NH_3PbI_3$ perovskite

7.1 Introduction

Despite the popularity of perovskite solar cells and the extensive toolbox of solution-based fabrication methods, the orientation of perovskite films has received little attention. This lack of attention is largely due to the fact that, irrespective of the formation method, most reported studies have focused on films with grains that are either randomly oriented or predominantly <110>-oriented. Brenner *et al.* demonstrated that a preferred orientation exists when PbI_2 is converted into perovskite through a topotactic reaction process. However, in most cases, the reaction between PbI_2 and CH_3NH_3 follows a dissolution–reconstruction process, and the preferred orientation is destroyed or the deposition orientation tends to be either random or predominantly <110>. Brenner *et al.* also demonstrated that even films initiated by a topotactic reaction with a strong relationship to the PbI_2 seed revert to dissolution–reconstruction behavior at longer time intervals, resulting in randomly oriented films [1]. This behavior may also extend to other reactions, such as the transition of the intermediate chloride-rich phase, $CH_3NH_3PbI_{3-x}Cl_x$, to $CH_3NH_3PbI_3$ under chlorine doping [2]. This finding has resulted in the proposition that chlorine doping results in improved uniformity [3], [4] and an increased carrier diffusion length from ~100 nm to >1 μm [5], which improves the performance [6].

Very few orientations besides (110) have been reported in the literature. Foley *et al.* demonstrated that their (100) tetragonal films were a result of interactions at the perovskite/solution interface of the cubic (110) phase, which was greatly stabilized by the solvent via a one-step solution process [7]. Their champion cells for (100) tetragonal

orientation showed a higher efficiency of 12.6%, despite having a carrier lifetime that is smaller than that of randomly oriented films by a factor of >20. Bae *et al.* reported efficiencies of 13.6% for (112)- and (200)-oriented I_3 films, with 11.26% for I_2 chloride films oriented parallel to (002) and (110) [8]. Tsai *et al.* obtained mesoscopic carbon electrode-based cells with (004)-oriented large grains by using N-Methyl-2-pyrrolidone as a crystal solvent with a champion device efficiency of 15%, which is approximately three-fold higher than that of <110>-oriented films with alternate precursors such as GBL, DMF, or DMSO [9]. This concept was recently adopted in the deposition of $FA_xMA_{1-x}PbI_{3-y}Cl_y$, resulting in a purely $<\bar{1}11>$-oriented film with a trap density lower (by a factor of 20) than that of the equivalent prepared using the solvent engineered technique [10]. A highly oriented film has been proposed to benefit from an enhancement of up to 300% in terms of carrier mobility, resulting in a higher fill factor and an efficiency of up to 19%.

In this section, we propose a novel fabrication method that produces, for the first time, highly oriented <202> perovskite layers by combining a two-step solution method with a solid-state reaction. The latter is a two-step method based on the reaction between separately prepared PbI_2 and CH_3NH_3I solid films that are subsequently brought into physical contact. As a relatively rare procedure for perovskite solar cells, this approach has previously resulted in efficiencies of ~10% [11]. We also investigate the impact of orientation on device performance.

7.2 Experimental methods

Our novel combined method is seed-assisted, as highlighted in Fig 7.1 (a). In the first stage, Step 1 consists of the formation of a PbI_2 layer via spin-coating of 460 mg/mL PbI_2/DMF solution at 2500 rpm for 30 s on an ITO substrate. In Step 2, a solution of 10

mg/mL $MAI_{1-x}Cl_x$ (9.5 mg MAI and 0.5 mg MACl per milliliter of 2-propanol) is spin-coated on the PbI_2 film at 1200 rpm for 40 s, followed by cleaning with IPA. In Step 3, the film is annealed at 90°C for 5 min to produce an intermediate perovskite phase that contains the desired seed layer. Stage 2 begins at Step 4 of the process, where 100 mg of $MAI_{1-x}Cl_x$ film is spin-coated on a separate substrate at 1500 rpm for 40 s. The spin-coated $MAI_{1-x}Cl_x$ film from Step 4 is placed face-to-face on top of the intermediate film from Step 3, and the two contacting films are annealed as a function of temperature for 45 min.

The films were compared with films produced by a conventional solid-state process whereby identical concentrations of PbI_2 and MAI were spin-coated at 2500 rpm for 30 s onto two separate substrates and brought into contact at 125°C for 45 min. Thicknesses in the range of ~200–500 nm are readily obtained via this method, with optimal current densities achieved for thicknesses of 300–350 nm.

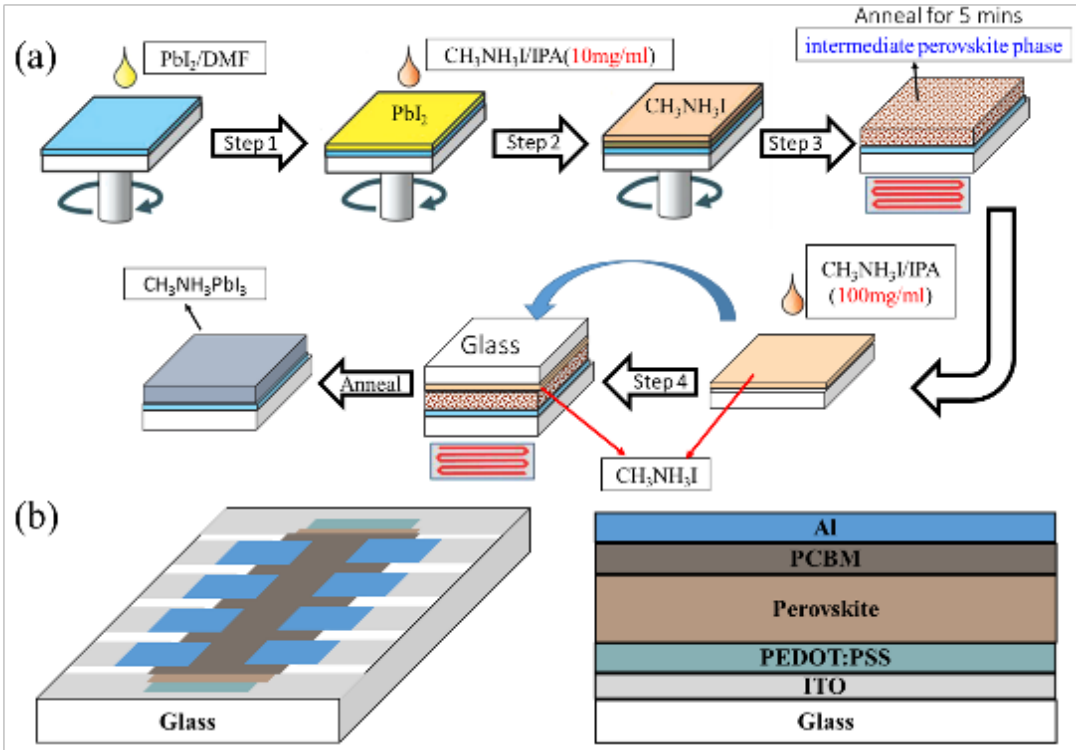


Fig 7.1 (a) Process flow diagram of the seed-assisted combined method for <202>-oriented perovskites. (b) Fabricated planar device structures.

Fig 7.1 (b) presents a schematic diagram of the cells reported in this work. Pre-patterned ITO substrates are cleaned via 1% Hellmanex in deionized (DI) water and IPA, dried under an N_2 flow, and placed on a hot plate at 120°C for 15 min. A 40-nm PEDOT: PSS layer is deposited via spin-coating and dried in air. The substrates are then transferred to a glove box, and the perovskite layers are deposited using either our novel combined seed-assisted method or the conventional solid-state method. A PCBM layer is deposited via spin-coating in chlorobenzene solution as the electron transport layer. Finally, the PCBM-deposited perovskite films are transferred to a high vacuum chamber for the deposition of Al electrodes and encapsulated with a glass cover using an ultraviolet epoxy sealant.

XRD measurements were performed in reflection geometry using a Bruker D2 Phaser operating at 30 kV and 10 mA. The simulated structure and XRD results were generated using CrystalMaker, and the effective electron and hole masses were extracted from band structures obtained from first-principle calculations within the generalized-gradient approximation (GGA-PBE) [12] of density-functional theory (DFT). The calculation employs a scalar-relativistic approximation with the pseudopotential projector-augmented wave method [13] as implemented in VASP [14]–[17]. Spin-orbit coupling is included to increase the accuracy of the band structure. A uniform mesh of 6x6x4 in the full Brillouin zone is used for the plane-wave basis set, with an energy cutoff of 500 eV.

Scanning electron microscopy (SEM) was performed on a FEGSEM RAITH SEM under operating conditions of 5 kV. The device measurements were performed using an Agilent B1500A semiconductor parameter analyzer and a LOT-oriel solar simulator under simulated global standard spectrum (AM1.5G) sunlight at 100 mW/cm^2 (Rera) and dark conditions. Raman spectroscopy was performed using a Renishaw inVia Raman microscope with 532-nm and 830-nm laser excitation. Measurements were conducted with backscatter geometry using a 50× objective lens (numerical aperture of 0.75) to focus and collect the laser and Raman-scattered light. Dielectric edge filters were used to block the Rayleigh-scattered light, and the Raman cutoff was set to 70 cm^{-1} at 532 nm and 100 cm^{-1} for the 830-nm configuration. For all measurements, the laser power was maintained at low levels to reduce/prevent laser-induced degradation of the samples.

7.3 Results and discussion

Fig 7.2 shows the normalized intensity of XRD peaks for the initial layer of spin-coated PbI_2 from Step 1 and the intermediate seed layer obtained from Step 3. The peaks

at 2θ = 12.7° correspond to the (001) plane of PbI_2 in both samples, confirming its only dominant orientation based on its known crystal behavior [6]. After spin-coating and annealing, the sample is still dominated by <001>-oriented PbI_2; however, a small fraction of $MAPbI_3$ with <202> orientation is also present on the surface, in agreement with the formation of an incomplete perovskite film [6], [18], [19]. Fig 7.2 (c)–(f) present XRD patterns of samples annealed at 120°C, 125°C, 130°C, and 135°C. The indexed XRD patterns confirm the presence of a pure β-phase tetragonal I4/mcm cell of $CH_3NH_3PbI_3$ [20], [21]. The XRD peaks in samples (c)–(f) located at 2θ = 14.21°, 24.57°, 28.55°, and 31.90° correspond to (110), (202), (220), and (310) planes of the $CH_3NH_3PbI_3$ perovskite. Despite the presence of 5% chlorine in the MAI solution, the formation of $CH_3NH_3PbCl_3$ is not detected at any stage of the reaction, indicating a loss of chlorine, as has been widely reported elsewhere [2], [22], [23]. Furthermore, Fig 7.2 reveals a narrow window at 125°C for the preferential formation of a <202>-oriented perovskite, with the same minimum peak intensity observed for other orientations. At temperatures higher or lower by 5°C, the (110) peak becomes stronger, and mixed orientations are observed, as commonly reported for the two-step solution method [18], [19], [24]. This result indicates a local minimum for the formation energy of this phase in contrast to the dissolution–reconstruction process.

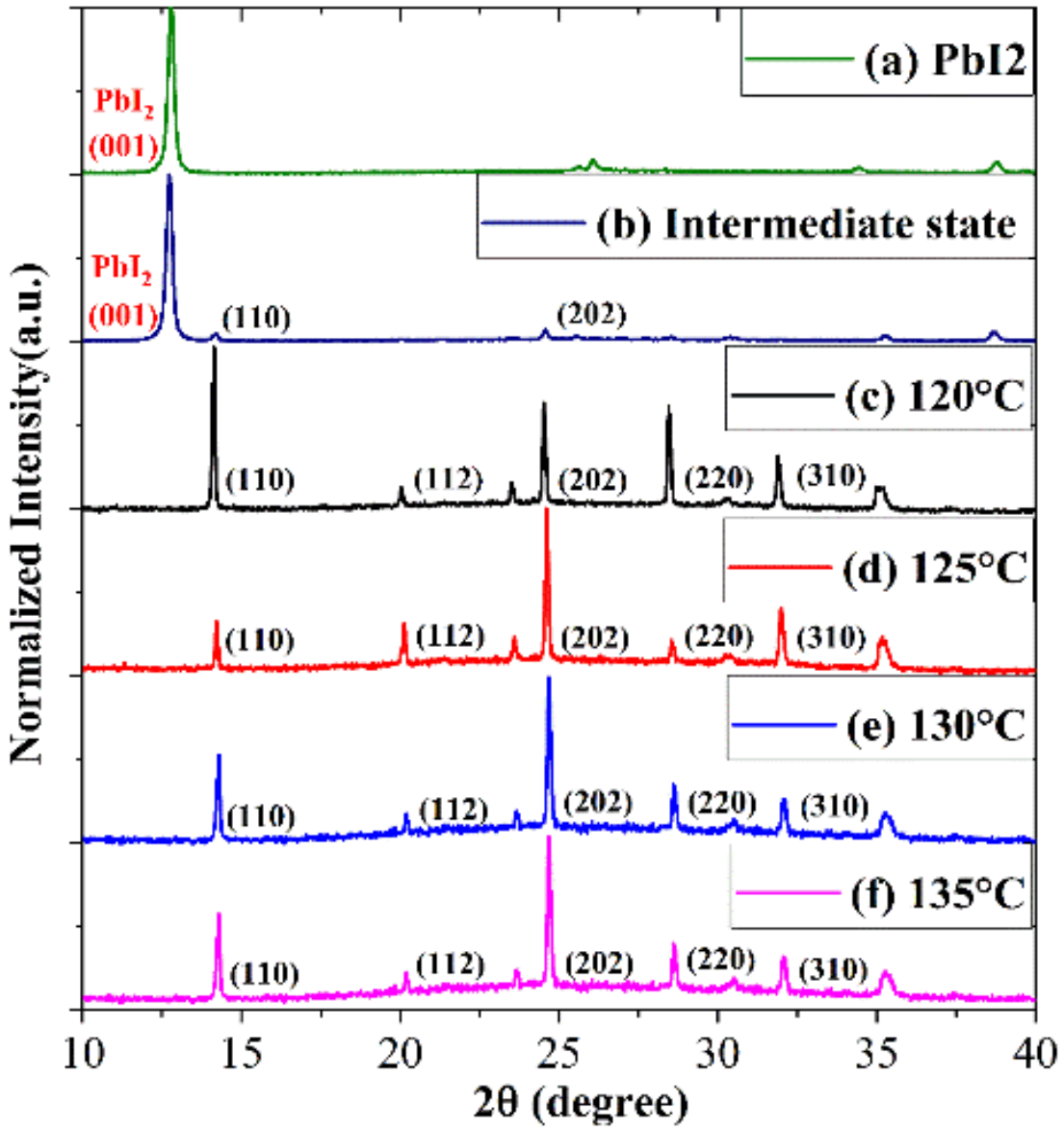


Fig 7.2 XRD of the (a) PbI_2 film, (b) intermediate seed layer, and (c–f) perovskite film after complete conversion from the intermediate film via the solid-state process at temperatures of 120°C to 135°C.

Despite both seeds appearing at a very early stage in the reaction, the formation processes of <110>- and <202>-oriented perovskite are rather different [6]. Brenner *et al.* demonstrated the reaction process of conversion from an <001> PbI_2 film to $CH_3NH_3PbI_3$

perovskite [1]. Depending on the structural relationship between the <001> PbI_2 film and the resultant $CH_3NH_3PbI_3$ perovskite, there are two different reaction processes for this conversion. If the reaction starts with dissolution of the Pb-I framework, i.e., a dissolution–reconstruction reaction process, the reaction tends to result in a randomly oriented $CH_3NH_3PbI_3$ film. In such a case, the Pb^{2+} ion and two I^- ions form the (110) plane, which is parallel to the substrate. Between each (110) plane, the MA^+ ion along with the third I^- ion forms another plane, which is also parallel to the substrate. Experimentally, this process tends to be dominated by the <110> orientation, where $CH_3NH_3PbI_3$ bears no relationship with the initial PbI_2 film.

In contrast, if the Pb-I framework retains its two-dimensional (2D) hexagonal packing, with the MA^+ ions and I^- ions diffused into the framework, the reaction is defined as topotactic, as illustrated in Fig 7.3. In the <001>-oriented PbI_2, each Pb ion is connected to six I^- ions, with each I^- shared by three Pb^{2+} ions (Fig 7.3 (a) and (b)). During the reaction with MAI, the distance between the Pb ions increases from 4.59 to 8.89 Å to form the (202) plane of $CH_3NH_3PbI_3$, while still retaining the original 2D hexagonal packing, as shown in Fig 7.3 (c). The MA^+ ions and one additional I^- ion from the MAI salt can diffuse into this Pb-I framework, resulting in a <202>-oriented $CH_3NH_3PbI_3$, where each I^- ion is now shared by two Pb^{2+} ions. The MA^+ compound is aligned in a plane at an angle of ~55° from the (202) plane (Fig 7.3 (c) and (d)). These diagrams demonstrate how the crystal retains a crystallographically equivalent orientational relationship with the original PbI_2 crystal, which is considered advantageous in terms of uniformity and trap state density [10] [25].

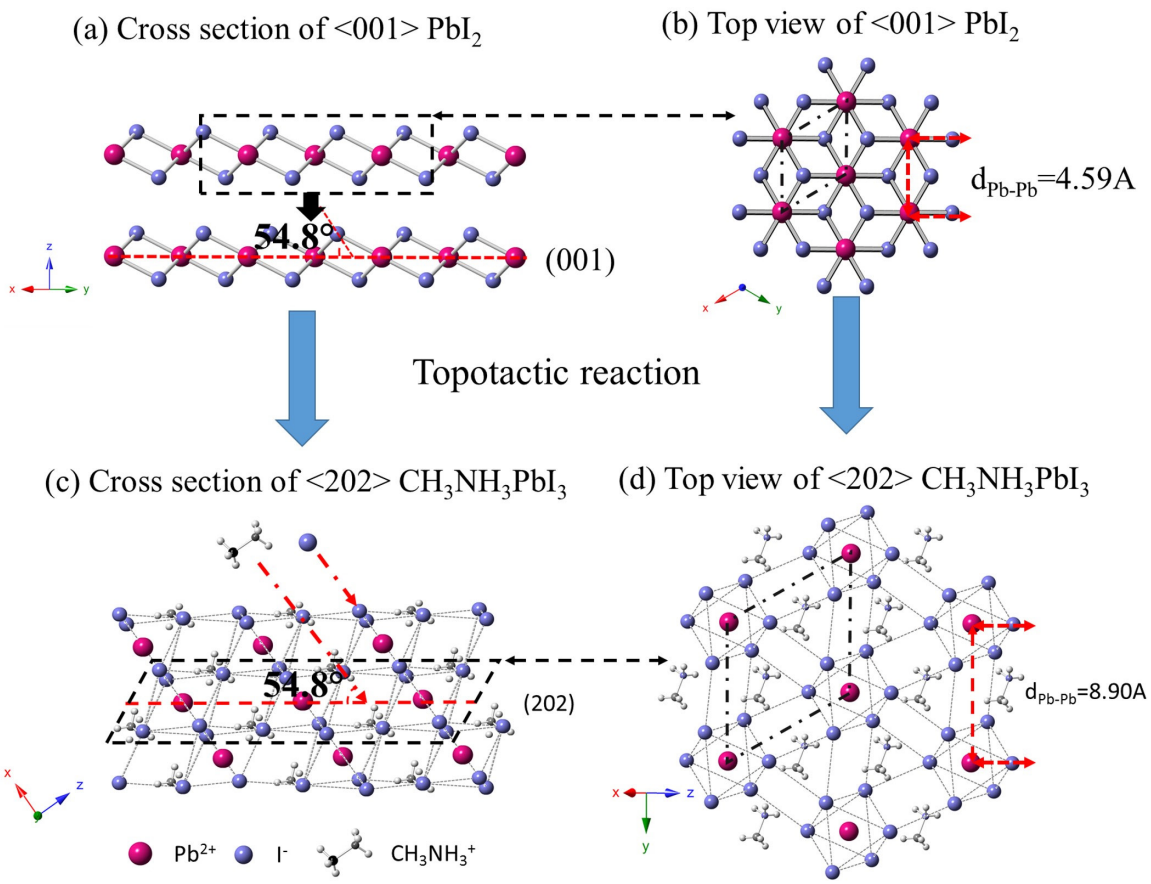


Fig 7.3 (a) Cross-section and (b) top view of <001>-oriented PbI_2. (c) Cross-section and (d) top view of <202>-oriented $CH_3NH_3PbI_3$ formed via a topotactic reaction process.

Although there is a strong preference for the formation of <202>-oriented perovskite during the early stages of the two-step reaction process [6], [18], this orientation has never been previously demonstrated as the dominant orientation after the PbI_2 is fully converted into $CH_3NH_3PbI_3$ [6], [18], [19]. Although Brenner *et al.* observed <202>-oriented perovskite after 16 h of reaction, due to a thicker PbI_2 crystal film, it was not fully converted into $CH_3NH_3PbI_3$, and a significant amount of <110>-oriented $CH_3NH_3PbI_3$ was

also observed in their films [26]. This difference is most likely due to the very narrow temperature window of stability for the <202> phase, as reported in this work.

Fig 7.4 (b) and (c) compare the XRD spectra of a set of perovskite films prepared via our novel seed-assisted combined method at 125°C and the conventional solid-state method. Even though the growth temperature is 10°C lower than in the solid-state process reported in [11], our solid-state samples show a single dominant <110> orientation that is more akin to vapor-phase processes [11], [24]. Moreover, at precisely the same temperature of 125°C, the seed-assisted sample shows a single dominant <202> orientation. It has been argued that the destruction of the initial Pb^{2+} structure by high thermal energy is the main reason for the deconstruction−reconstruction reaction process [6]. As the thermal energy is identical in both cases, this cannot be the reason for the difference in the reaction process observed here. It is likely that a small amount of perovskite mixed in PbI_2 results in a change in the activation energy of the two different reactions, which changes the preferential orientation at the two temperatures.

To understand the differences between <110>- and <202>-oriented samples, three simulated structures are studied in this paper. As shown in Fig 7.4 (a), Structure 1 represents an ideal tetragonal structure based on cell parameters extracted from experimental results. Structure 2 is one of the structures reported in [27], where all MA^+ cations are aligned along the <001> direction parallel to the z-axis. The cell parameters are refitted with experimental values of $a = \mathrm{b} = 8.84\text{Å}$ and $\mathrm{c} = 12.68\text{Å}$ to match our observed XRD pattern. As the MA^+ cations aligned along the z-axis are highly symmetrical in the xy plane, the unit cell shows no octahedral tilting in this plane. Meanwhile, attracted by the positively charged MA^+ cations, the negatively charged I^- anions within the (001)

plane are shifted along the <001> direction, twisting the Pb-I framework along the z-axis. Structure 3 is obtained after the DFT relaxation of Structure 2. After the relaxation, the MA^+ cations rotate ~15° within the yz plane, causing a further change in the Pb-I framework. As a result, the Pb-I bonding rotates ~3° from the z-axis, primarily due to movement of the I^- anions, whereas the Pb^{2+} ions show little change in position. In this case, the projection of the unit cell on the xz plane is identical to that in Structure 2. However, there is a considerable change in the xy and yz planes caused by the rotation of the MA^+ cations.

Fig 7.4 (b) and (c) also show the simulated XRD of all three structures when <110> or <202> is the preferred orientation, respectively. The difference between all three structures for the <110> orientation is slight. However, due to interference between the (202) and (404) peaks, the relative intensity between (202) and (404) is sensitive to the structural disorder of the Pb-I framework. In the ideal tetragonal structure (Structure 1), the predominant peak in the <202>-oriented sample becomes <404>, which differs from the experimental results. For Structure 2, as the shifts in the Pb-I framework affect both (202) and (404), the results show a close agreement with the experimental results despite the structural instability of this model [27]. Interestingly, for the <202>-oriented sample, this shift causes the MA^+ cation to be aligned at an angle of ~55° from the (202) plane, consistent with the preferred growth direction reported by Brenner *et al.* [26] in 2016.

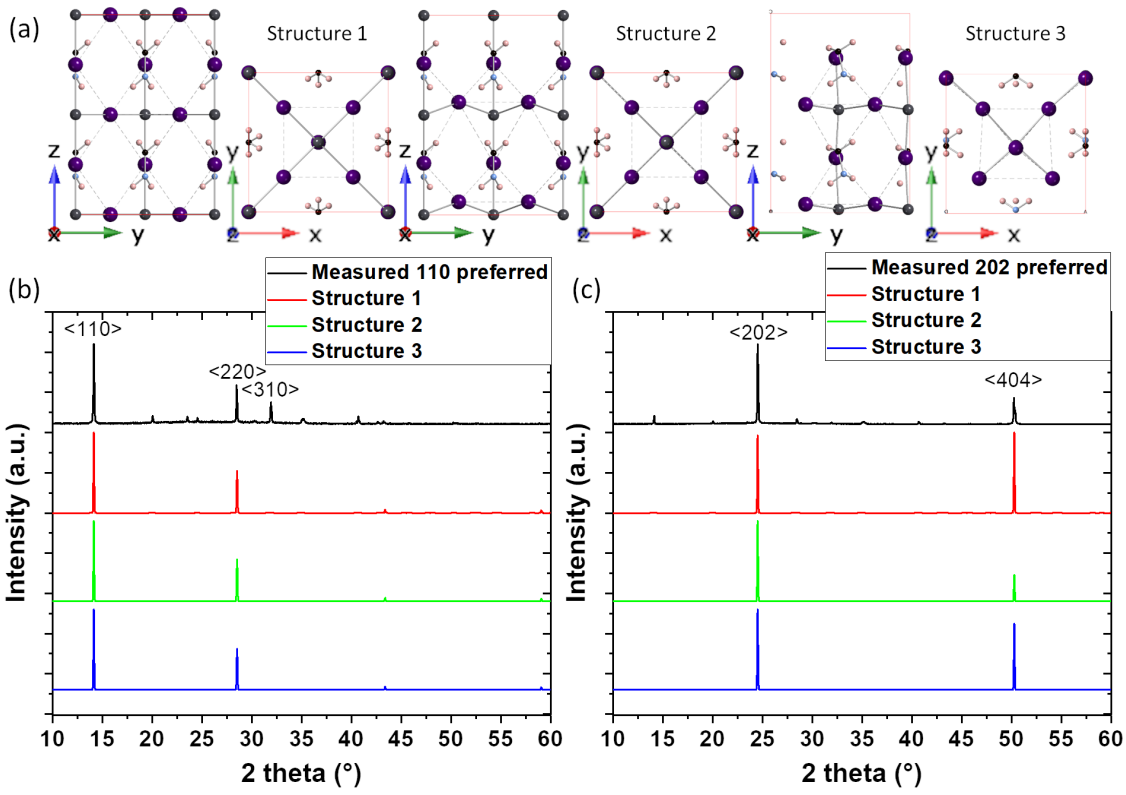


Fig 7.4 (a) Three structures used in this study (viewed from the <100> and <001> directions). Structure 1 is an ideal tetragonal structure, Structure 2 is based on a unit cell that contains all MA^+ cations aligned along <001> [27], and Structure 3 is obtained by performing a DFT relaxation of Structure 2. (b, c) Measured and simulated XRD of the (b) <110>- and (c) <202>-oriented perovskite films from Structures 1–3.

Besides changing the preferred growth orientation, it appears that the intermediate state during the growth process has a considerable impact on the film morphology. As shown in the inset in Fig 7.5, the <202>-oriented sample has a smaller grain size than that of the <110>-oriented sample, with sizes of 0.4 μm and ~1.2 μm, respectively. Because the growth process is the same in both cases except for the seed layer, it seems reasonable to

conclude that the difference in morphology arises from the difference in orientation. It can be argued that the growth of the <202> perovskite is not perpendicular to the substrate, which has an impact on film morphology [6].

The difference in the orientation and grain size seems to result in very different behaviors for photovoltaic solar cells under dark and light environments. Fig 7.5 shows J−V curves for perovskite solar cells with <110>- and <202>-oriented films under a dark environment. As both devices possess the same architecture fabricated under identical conditions, it appears that the small grain size of the <202>-oriented perovskite results in more grain boundaries, leading to a larger dark current.

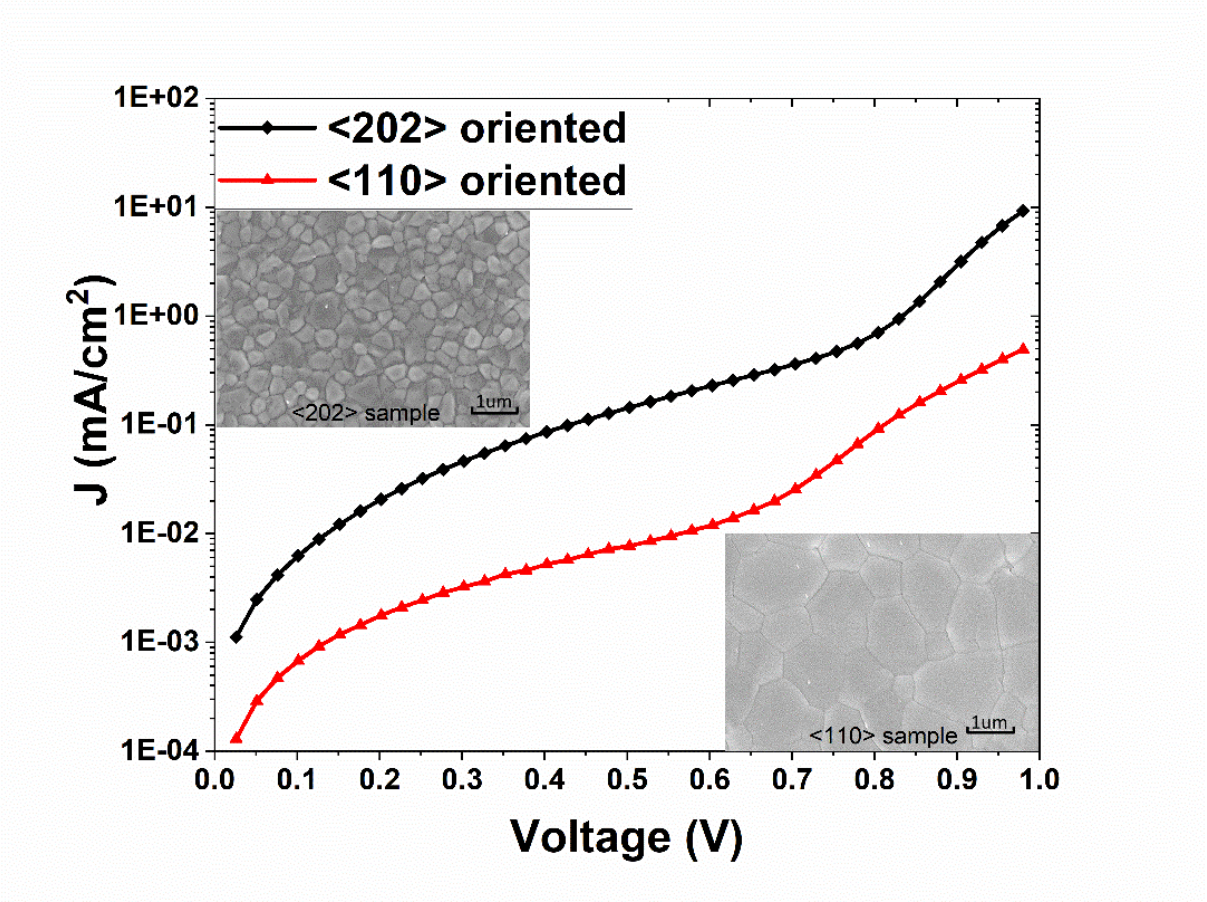


Fig 7.5 J–V curves of perovskite solar cells with <110>- and <202>-oriented perovskite measured in a dark environment. Corresponding SEM images of each sample are shown in the inset, where the <202>- and <110>-oriented perovskites show a grain size of approximately 0.4 µm and 1.2 µm, respectively.

Fig 7.6 illustrates Raman spectra collected from samples with the two orientations. Accurate peak positions were obtained by fitting experimental data with a series of Gaussian curves, as shown in Fig 7.6 (b)–(d). Under 830-nm excitation, strong peaks related to the glass substrate are observed above 350 cm^{-1}. These peaks occur because the 830-nm source is below the band gap of the $CH_3NH_3PbI_3$ layer (1.5 eV) and can probe the entire device structure, unlike the 532-nm excitation, which is absorbed by the

$CH_3NH_3PbI_3$ layer. In subsequent analyses, these glass bands were removed by subtracting a glass reference spectrum collected under the same conditions.

When measured at a given wavelength, the <110> and <202> samples exhibit identical Raman bands with a deviation in position. The bands vary for the different excitation wavelengths. For 830 nm, Raman peaks are observed at 122 cm^{-1}, 160 cm^{-1}, and 252 cm^{-1}. It has been previously proposed that the 122-cm^{-1} Raman mode arises from an overlap of the stretching motions of the Pb−I cage and the secondary peak of the MA liberation motion [28], [29]. For 830-nm excitation, the 160-cm^{-1} Raman mode is most likely due to the main peak of the MA liberation [29], [30]. Due to its strong overlap with other peaks, the position of this mode may previously have been considered less accurate. The Raman mode at ~250 cm^{-1} has been assigned to the torsional mode of the MA cation [31]. For measurements with an excitation wavelength of 532 nm, in addition to the 122-cm^{-1}, 160-cm^{-1}, and 252-cm^{-1} peaks observed for 830-nm and 532-nm excitation, there are four additional peaks at 72 cm^{-1}, 79 cm^{-1}, 91 cm^{-1}, and 109 cm^{-1}. The 532-nm Raman spectra can be attributed to two causes. The 122-cm^{-1}, 160-cm^{-1}, and 252-cm^{-1} Raman modes constitute a band, which is in good agreement with the spectra for 830-nm excitation. The two main peaks at 122 cm^{-1} and 252 cm^{-1} agree well. The main difference is the relatively stronger peak at 160 cm^{-1} for 532-nm excitation, which contains a strong shoulder area between 150 cm^{-1} and 200 cm^{-1}; this peak does not arise for 830-nm excitation. Additional peaks at 72 cm^{-1}, 79 cm^{-1}, 91 cm^{-1}, and 109 cm^{-1}, which are commonly attributed to degradation products such as MAI, PbO_x, and PbI_2 [32]–[34], are absent for 830-nm excitation. The 72-cm^{-1} and 91-cm^{-1} Raman modes agree well with the PbI_2 Raman modes reported at 73 cm^{-1} and 94 cm^{-1} [32], [35]. The 79-cm^{-1} Raman mode

seems closer to the 81-cm^{-1} Raman mode for PbO_x than the mode for PbI_2 [36]. The 109-cm^{-1} mode is often attributed to the MA^+ compound of $CH_3NH_3PbI_3$, whose position can be shifted by the incorporation of H_2O molecules [37]. However, this peak was not detected for 830-nm excitation. Considering that the degradation of $CH_3NH_3PbI_3$ often follows as $CH_3NH_3PbI_3 \rightarrow PbI_2 + CH_3NH_3I$ [38], the presence of PbI_2 is often accompanied by the formation of MAI. Therefore, we attribute this peak to MAI, which also has a Raman mode at 110 cm^{-1} in ambient conditions [39]. The absence of a degradation band for 830 nm suggests that the creation of PbI_2 is due to exposure of the sample to a 532-nm laser, implying that the origin is photodegradation or localized thermal degradation from the absorption of laser power. This result highlights the benefits of applying excitation below the band gap when analyzing perovskite samples. In summary, there are no differences in the Raman spectra observed between <110>- and <202>-oriented samples. This result is likely due to the dynamics of the MA^+ cations in the perovskite, where the orientation has little to no impact on the Raman mode due to the rotational motion of MA^+ ions [31].

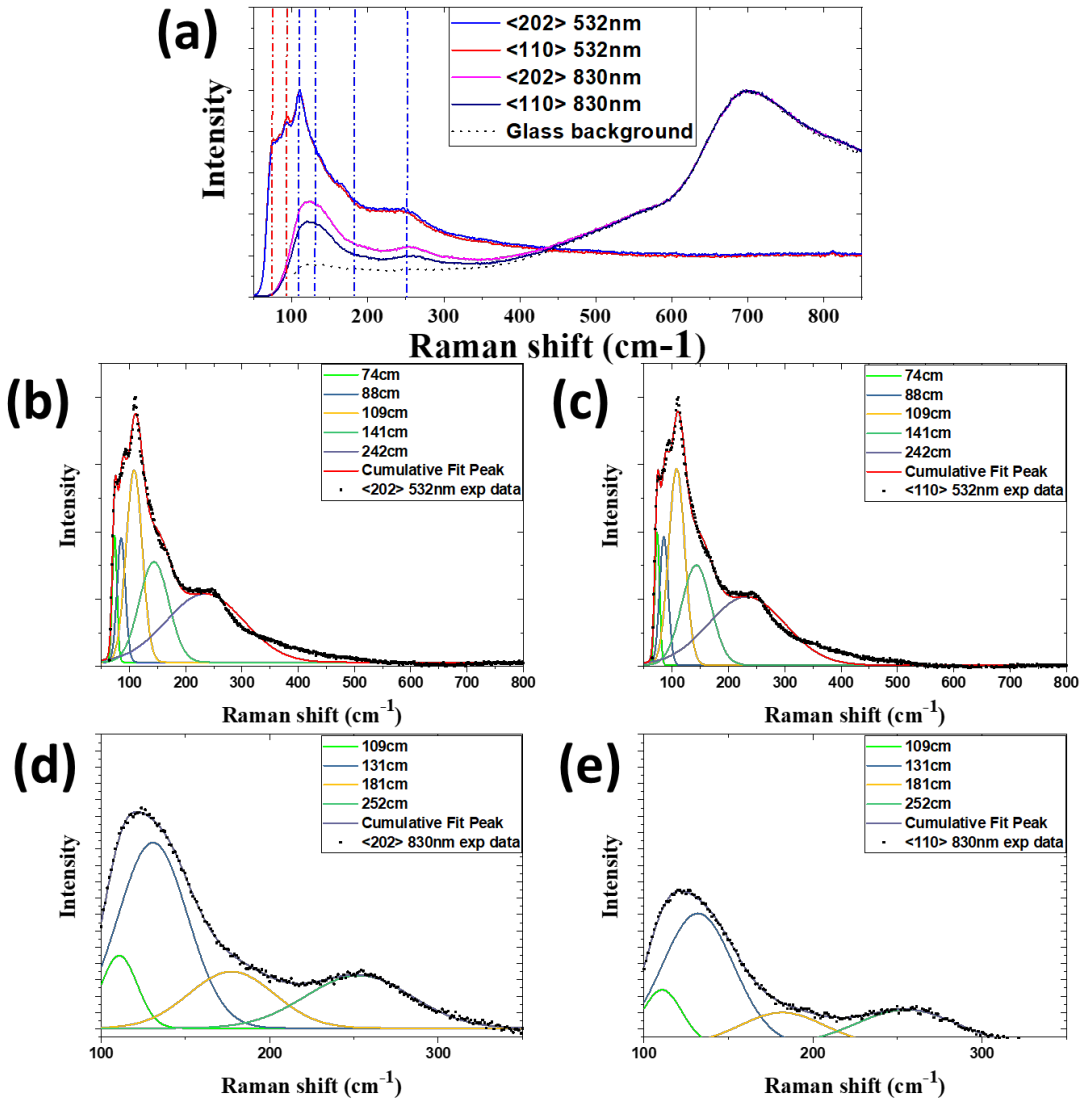


Fig 7.6 (a) Normalized Raman spectra of $CH_3NH_3PbI_3$ perovskite layers measured using excitation at 532 nm and 830 nm. The vertical dashed lines indicate the peak positions of the Raman band, obtained by fitting the spectra with Gaussian curves. (b, c) Raman spectra obtained for 532-nm excitation, fitted with Gaussian curves. (d, e) Raman spectra obtained for 830-nm excitation fitted with Gaussian curves.

Fig 7.7 presents J−V curves for the best PSCs with <110> and <202> orientations under AM1.5G sunlight (100 $mWcm^{-2}$). Benefiting from the improved grain size in comparison to earlier work, which obtained a grain size <1 μm by using the solid-state reaction (V_{oc} = 0.87 V, J_{sc} = 17.9 $mA \cdot cm^{-1}$, FF = 0.643, PCE = 10%) [11], our cell with <110>-oriented perovskite demonstrates an improved J_{sc} of 23.7 $mA \cdot cm^{-1}$, a V_{oc} of 0.95 V, an FF of 0.541, and a PCE of 12.2%, corresponding to an improvement of 32% for J_{sc}, 9% for V_{oc}, and 22% for PCE. Our sample prepared using <202> orientation resulted in V_{oc} = 0.93 V, J_{sc} = 22.5 mA, FF = 70.9%, and PCE=14.9%. Both types of samples produce a comparable V_{oc}, while the <202> orientation repeatedly results in a better PCE attributed to a higher FF and current density, despite having a higher dark current. This result seems contradictory, but it does agree with the relatively small effective electron and hole mass along the <202> direction. With interface engineering [40], post-solvent vapor treatment [41], [42], or H_2O doping [42], typical inverted (p–i–n) PSCs of <110>-oriented perovskite have demonstrated PCE values of more than 20% (V_{oc} = 1.03 V, J_{sc} = 23.51 mA, FF = 0.83, PCE = 20.1%) [42]. By comparison, it seems reasonable to conclude that, despite the smaller grain size, the <202> orientation demonstrates significant potential for producing similar levels of photovoltaic performance.

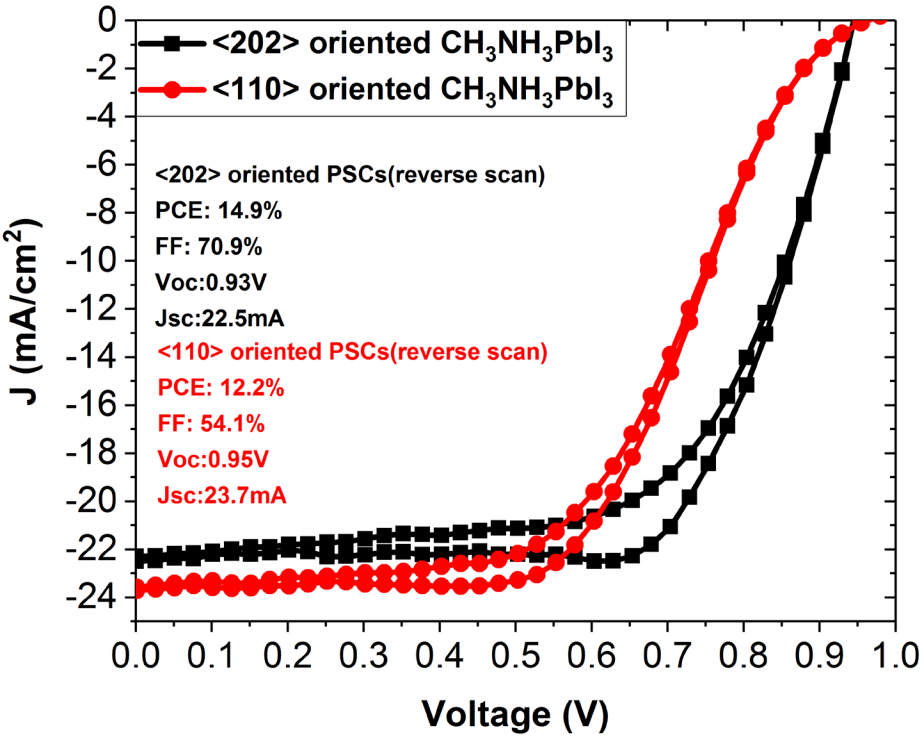


Fig 7.7 J–V curves of perovskite solar cells with <110> and <202> oriented perovskites under AM1.5G simulated sunlight.

7.4 Summary

We have successfully demonstrated a novel seed-oriented combined method, which results in the formation of highly oriented <202> perovskite layers in a narrow temperature range around 125°C. Despite the fact that the grain size of <202> perovskite is significantly smaller than that of the more conventional <110> orientation, leading to a higher dark current, the photovoltaic performance, and particularly the fill factor, seem to be only minimally impacted. Both the <110> and <202> oriented samples are capable of producing a current density exceeding 20 mA/cm2, with the <202> orientation consistently yielding a high fill factor.

Quarti et al.'s examination of a range of structures [27] suggest that the most probable structure is one in which the MA+ cations are aligned along the z-direction, which is the closest fit to the observed lattice constants and XRD results. This alignment would lead to structural distortion, with the MA+ cations aligning parallel to the preferred growth direction of the ensuing topotactic reaction. The primary mechanism behind the observed enhancement in performance appears to be a high mobility due to the lower effective mass [26][27].

7.5 References

[1] T. M. Brenner *et al.*, "Conversion of Single Crystalline PbI 2 to CH 3 NH 3 PbI 3 : Structural Relations and Transformation Dynamics," *Chem. Mater.*, vol. 28, no. 18, pp. 6501–6510, Sep. 2016.

[2] S. T. Williams *et al.*, "Role of Chloride in the Morphological Evolution of Organo-Lead Halide Perovskite Thin Films," *ACS Nano*, vol. 8, no. 10, pp. 10640–10654, Oct. 2014.

[3] M. I. Dar *et al.*, "Investigation Regarding the Role of Chloride in Organic–Inorganic Halide Perovskites Obtained from Chloride Containing Precursors," *Nano Lett.*, vol. 14, no. 12, pp. 6991–6996, Dec. 2014.

[4] A. Dualeh *et al.*, "Effect of Annealing Temperature on Film Morphology of Organic-Inorganic Hybrid Pervoskite Solid-State Solar Cells," *Adv. Funct. Mater.*, vol. 24, no. 21, pp. 3250–3258, Jun. 2014.

[5] S. D. Stranks *et al.*, "Electron-Hole Diffusion Lengths Exceeding 1 Micrometer in an Organometal Trihalide Perovskite Absorber," *Science (80-.).*, vol. 342, no. 6156, pp. 341–344, Oct. 2013.

[6] T. M. Brenner *et al.*, “Conversion of Single Crystalline PbI 2 to CH 3 NH 3 PbI 3 : Structural Relations and Transformation Dynamics,” *Chem. Mater.*, vol. 28, no. 18, pp. 6501–6510, Sep. 2016.

[7] B. J. Foley *et al.*, “Controlling nucleation, growth, and orientation of metal halide perovskite thin films with rationally selected additives,” *J. Mater. Chem. A*, vol. 5, no. 1, pp. 113–123, 2017.

[8] S. Bae *et al.*, “CH 3 NH 3 PbI 3 crystal orientation and photovoltaic performance of planar heterojunction perovskite solar cells,” *Sol. Energy Mater. Sol. Cells*, vol. 160, no. September 2016, pp. 77–84, Feb. 2017.

[9] C.-M. Tsai *et al.*, “Control of preferred orientation with slow crystallization for carbon-based mesoscopic perovskite solar cells attaining efficiency 15%,” *J. Mater. Chem. A*, vol. 5, no. 2, pp. 739–747, 2017.

[10] D. H. Kim *et al.*, “300% Enhancement of Carrier Mobility in Uniaxial-Oriented Perovskite Films Formed by Topotactic-Oriented Attachment,” *Adv. Mater.*, vol. 29, no. 23, p. 1606831, Jun. 2017.

[11] L. Chen *et al.*, “Facile preparation of organometallic perovskite films and high-efficiency solar cells using solid-state chemistry,” *Nano Res.*, vol. 8, no. 1, pp. 263–270, Jan. 2015.

[12] J. P. Perdew *et al.*, “Generalized Gradient Approximation Made Simple,” *Phys. Rev. Lett.*, vol. 77, no. 18, pp. 3865–3868, Oct. 1996.

[13] P. E. Blöchl, “Projector augmented-wave method,” *Phys. Rev. B*, vol. 50, no. 24, pp. 17953–17979, Dec. 1994.

[14] G. Kresse and J. Hafner, “Ab initio molecular dynamics for open-shell transition

metals," *Phys. Rev. B*, vol. 48, no. 17, pp. 13115–13118, Nov. 1993.

[15] G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set," *Phys. Rev. B*, vol. 54, no. 16, pp. 11169–11186, Oct. 1996.

[16] G. Kresse and J. Furthmüller, "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set," *Comput. Mater. Sci.*, vol. 6, no. 1, pp. 15–50, Jul. 1996.

[17] G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method," *Phys. Rev. B*, vol. 59, no. 3, pp. 1758–1775, Jan. 1999.

[18] Y. Zhou *et al.*, "Growth control of compact CH 3 NH 3 PbI 3 thin films via enhanced solid-state precursor reaction for efficient planar perovskite solar cells," *J. Mater. Chem. A*, vol. 3, no. 17, pp. 9249–9256, 2015.

[19] C.-H. Chiang *et al.*, "Planar heterojunction perovskite/PC 71 BM solar cells with enhanced open-circuit voltage via a (2/1)-step spin-coating process," *J. Mater. Chem. A*, vol. 2, no. 38, pp. 15897–15903, 2014.

[20] J. H. Lee *et al.*, "The nature of hydrogen-bonding interaction in the prototypic hybrid halide perovskite, tetragonal CH3NH3PbI3," *Sci. Rep.*, vol. 6, no. 1, p. 21687, Feb. 2016.

[21] Y. Dang *et al.*, "Bulk crystal growth of hybrid perovskite material CH 3 NH 3 PbI 3," *CrystEngComm*, vol. 17, no. 3, pp. 665–670, 2015.

[22] H. Yu *et al.*, "The Role of Chlorine in the Formation Process of 'CH 3 NH 3 PbI 3-x Cl x ' Perovskite," *Adv. Funct. Mater.*, vol. 24, no. 45, p. n/a-n/a, Sep. 2014.

[23] Q. Chen *et al.*, "The optoelectronic role of chlorine in CH3NH3PbI3(Cl)-based

perovskite solar cells," *Nat. Commun.*, vol. 6, no. 1, p. 7269, Dec. 2015.

[24] S. Chen *et al.*, "Characterization of Perovskite Obtained from Two-Step Deposition on Mesoporous Titania," *ACS Appl. Mater. Interfaces*, vol. 7, no. 46, pp. 25770–25776, Nov. 2015.

[25] D. T. Moore *et al.*, "Impact of the organic halide salt on final perovskite composition for photovoltaic applications," *APL Mater.*, vol. 2, no. 8, p. 081802, Aug. 2014.

[26] T. M. Brenner *et al.*, "Conversion of Single Crystalline PbI 2 to CH 3 NH 3 PbI 3 : Structural Relations and Transformation Dynamics," *Chem. Mater.*, vol. 28, no. 18, pp. 6501–6510, Sep. 2016.

[27] C. Quarti *et al.*, "Interplay of Orientational Order and Electronic Structure in Methylammonium Lead Iodide: Implications for Solar Cell Operation," *Chem. Mater.*, vol. 26, no. 22, pp. 6557–6569, Nov. 2014.

[28] B. Park *et al.*, "Resonance Raman and Excitation Energy Dependent Charge Transfer Mechanism in Halide-Substituted Hybrid Perovskite Solar Cells," *ACS Nano*, vol. 9, no. 2, pp. 2088–2101, Feb. 2015.

[29] F. Brivio *et al.*, "Lattice dynamics and vibrational spectra of the orthorhombic, tetragonal, and cubic phases of methylammonium lead iodide," *Phys. Rev. B*, vol. 92, no. 14, p. 144308, Oct. 2015.

[30] M. Park *et al.*, "Critical Role of Methylammonium Librational Motion in Methylammonium Lead Iodide (CH 3 NH 3 PbI 3) Perovskite Photochemistry," *Nano Lett.*, vol. 17, no. 7, pp. 4151–4157, Jul. 2017.

[31] G. Grancini *et al.*, "The Impact of the Crystallization Processes on the Structural and Optical Properties of Hybrid Perovskite Films for Photovoltaics," *J. Phys.*

Chem. Lett., vol. 5, no. 21, pp. 3836–3842, Nov. 2014.

[32] M. Ledinský *et al.*, "Raman Spectroscopy of Organic–Inorganic Halide Perovskites," *J. Phys. Chem. Lett.*, vol. 6, no. 3, pp. 401–406, Feb. 2015.

[33] P. Pistor *et al.*, "Advanced Raman Spectroscopy of Methylammonium Lead Iodide: Development of a Non-destructive Characterisation Methodology," *Sci. Rep.*, vol. 6, no. 1, p. 35973, Dec. 2016.

[34] S. M. Jain *et al.*, "Vapor phase conversion of PbI 2 to CH 3 NH 3 PbI 3 : spectroscopic evidence for formation of an intermediate phase," *J. Mater. Chem. A*, vol. 4, no. 7, pp. 2630–2642, 2016.

[35] Y. Zhou *et al.*, "Challenges in the ambient Raman spectroscopy characterization of methylammonium lead triiodide perovskite thin films," *Front. Optoelectron.*, vol. 9, no. 1, pp. 81–86, Mar. 2016.

[36] M. Baleva and V. Tuncheva, "Optical Characterization of Lead Monoxide Films Grown by Laser-Assisted Deposition," *J. Solid State Chem.*, vol. 110, no. 1, pp. 36–42, May 1994.

[37] R. Segovia *et al.*, "Evolution of Photoluminescence, Raman, and Structure of CH3NH3PbI3 Perovskite Microwires Under Humidity Exposure," *Nanoscale Res. Lett.*, vol. 13, no. 1, p. 79, Dec. 2018.

[38] M. Shahbazi and H. Wang, "Progress in research on the stability of organometal perovskite solar cells," *Sol. Energy*, vol. 123, pp. 74–87, Jan. 2016.

[39] R. Kumar Singh *et al.*, "Stability behavior of chemically synthesized organic electrolyte salts and methylammonium lead halide perovskite light harvester," *Adv. Mater. Lett.*, vol. 8, no. 6, pp. 707–711, 2017.

[40] H. Zhou *et al.*, “Interface engineering of highly efficient perovskite solar cells,” *Science (80-.).*, vol. 345, no. 6196, pp. 542–546, Aug. 2014.

[41] J. Lian *et al.*, “Organic solvent vapor sensitive methylammonium lead trihalide film formation for efficient hybrid perovskite solar cells,” *J. Mater. Chem. A*, vol. 3, no. 17, pp. 9146–9151, 2015.

[42] C.-H. Chiang *et al.*, “The synergistic effect of H 2 O and DMF towards stable and 20% efficiency inverted perovskite solar cells,” *Energy Environ. Sci.*, vol. 10, no. 3, pp. 808–817, 2017.

CHAPTER 8

Conclusions and future works

8.1 Summary

The focus of this thesis revolves around the development and understanding of materials and devices that can be integrated with solar cells for intelligent autonomous sensors. This includes the study of operational dynamics of solid-electrolyte Ta_2O_5/ZnO transistors, and their potential use in neuromorphic applications. Concurrently, we investigate a unique deposition technique that influences the growth orientation of $CH_3NH_3PbI_3$ perovskite that has not been reported before to our knowledge. The effects of this process on film morphology and solar cell efficiency are scrutinized.

Chapter 4 provides a detailed explanation of the fabrication process and operational mechanisms of our Solid-Electrolyte Field-Effect Transistors (SE-FETs). Expanding on the initial model presented in [1], we introduce a gate current model that further enables the device simulations discussed in Chapter 5. This chapter also brings to light the constraints discovered in the interpretation of the device's off-region from the previous model, particularly where a symmetrical drift-diffusion process under both positive and negative biases was assumed. Although this assumption worked well for the drain-source terminal, it did not align with the measured gate current of the device. By adopting a better implantation of a dynamic diffusion model, a better alignment between experimental results and simulation of the gate current is made possible.

In Chapter 5, we highlight the off-state operation of SE-FETs, maintaining the power consumption for writing operations within the nanowatt range. We further delve into the impact of oxide thickness on memory performance and exhibit logic operations using

a single SE-FET device. To bolster our exploration of potential applications, we use simulations to demonstrate the training and execution of an SE-FET-based crossbar array designed for supervised learning tasks.

Chapter 6 presents experimental demonstration of a three-terminal SE-FET-based reservoir network exhibiting enhanced dimensionality. The inherent diffusion-based mechanism within this network enables slow fading memory without the need for supporting array for added delay. Furthermore, the unique three-terminal structure enables simultaneous reading and writing operations, effectively eliminating the necessity for input down-sampling. This structure ensures uninterrupted data processing of continuous input, such as voice and continuous wave. Combination of inherent device properties and suitable network architecture resulted in the highest learning efficiency of 94.44% in a SE-FET based reservoir reported to date.

In Chapter 7, we present a unique seed-oriented combined method, inducing highly oriented <202> perovskite layers within a restricted temperature window. We also delve into the differential formation process and explore the impacts of this new orientation on film morphology and photovoltaic efficiency.

8.2 Future work

Due to the unforeseen event from an accident due to a fire the clean room in 2019 and delay in installing the equipment, planned work could not be achieved, particularly in integration of the intelligent autonomous sensors.

As discussed in Chapter 1, one direction involves the integration of perovskite solar cells with Ta_2O_5/ZnO-based transistors. Given that both devices are based on ITO substrates, the fabrication process necessitates the deposition and creation of Ta_2O_5/ZnO-

based transistors first, followed by the fabrication of perovskite solar cells using a solution-based technique.

In Chapter 7, while we've studied the impact of perovskite orientation on cell performance, the effect on hysteresis and memory remains unexplored. A study on devices featuring a straightforward memristor structure could potentially shed light on orientation-related hysteresis and memory differences. This approach could also isolate the influential factor of the charge transfer layer [3][4] and lead to a more nuanced understanding of this novel orientation. Such a study had been planned to complete the study on this new perovskite orientation and could potentially benefit the field of perovskite based memristor.

8.3 References

[1] A. Kumar *et al.*, "Negative Capacitance beyond Ferroelectric Switches," *ACS Appl. Mater. Interfaces*, vol. 10, no. 23, pp. 19812–19819, Jun. 2018.

[2] C. Du *et al.*, "Reservoir computing using dynamic memristors for temporal information processing," *Nat. Commun.*, vol. 8, no. 1, p. 2204, Dec. 2017.

[3] B. Wu *et al.*, "Charge Accumulation and Hysteresis in Perovskite-Based Solar Cells: An Electro-Optical Analysis," *Adv. Energy Mater.*, vol. 5, no. 19, pp. 1–8, 2015.

[4] H. S. Kim and N. G. Park, "Parameters affecting I-V hysteresis of CH3NH3PbI3 perovskite solar cells: Effects of perovskite crystal size and mesoporous TiO2 layer," *J. Phys. Chem. Lett.*, vol. 5, no. 17, pp. 2927–2934, 2014.