Investigation of Logic Errors in Memristive Crossbars for Computing-in-Memory

Von der Fakultät für Elektrotechnik und Informationstechnik der Rheinisch-Westfälischen Technischen Hochschule Aachen zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Leon Brackmann, M.Sc.

aus Eckernförde

Berichter: Universitätsprofessor Dr.-Ing. Rainer Waser
Universitätsprofessor John Paul Strachan, Ph.D.

Tag der mündlichen Prüfung: 26.05.2025

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.

Curriculum Vitae

26.05.2025	Doctoral Examination for *Doktor der Naturwissenschaften* (Dr.rer.nat.)
2021 - 2025	Research Assistant and Doctoral Candidate Institute of Materials in Electrical Engineering and Information Technology 2 RWTH Aachen University, Aachen, Germany
July 2021	Master of Science
2019 - 2021	Studies of Physics RWTH Aachen University, Aachen, Germany
March 2019	Bachelor of Science
2015 - 2019	Studies of Physics Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany
2007 - 2015	General Higher Education Entrance Qualification (Abitur) Konrad-Adenauer-Gymnasium, Meckenheim, Germany

Abstract

Valence Change Mechanism (VCM) redox-based Resistive Random Access Memories (ReRAM) are possible candidates to replace CMOS-based SRAM in data storage applications. These devices offer non-volatility, scalability, and an improved reliability, making them well-suited for data storage systems. The inclusion of filamentary VCM-ReRAM cells with Complementary Metal-Oxide Semiconductor (CMOS) transistors in 1T-1R crossbar arrays further enhances the device control, enabling the integration of large-scale resistive architectures. Beyond pure data storage, resistive switching memories are suitable for Computing-in-Memory (CIM) applications, including Boolean logic and arithmetic operations. However, resistive logic CIM is still in development as the device variability and integration challenges are limiting a large-scale experimental verification. Therefore system-level simulations are exploited to predict the performance of resistive computing systems.

While energy consumption and latency are straightforward to evaluate, the reliability and error probabilities in resistive CIM systems remain less understood. Due to the inherent device variability, demonstrating a single logic gate is insufficient to guarantee the reliability of large-scale systems involving thousands of operations. This thesis tries to address this gap by comprehensively investigating single-gate logic operations in 1T-1R CMOS/ReRAM crossbars and through a combination of experimental studies and variability-aware simulations.

The work examines two key branches of resistive logic: readout-based and switching-based operations. Readout-based logic relies on the sensing of resistive states as analog quantities and classifying them as digital values. Its reliability is limited by the accuracy of state sensing and the overlap of resistive distributions. Switching-based logic, in contrast, is based on a conditional state transition through an input-dependent voltage divider. Here, the reliability is determined not only by the resistive state variability but also by the stochastic nature of the switching mechanism. Therefore, both logic concepts require individual approaches in order to estimate their error proneness.

By systematically analyzing both concepts, this thesis proposes methodologies to evaluate and improve the reliability of both types and extrapolates the findings to more complex CIM systems. Building on previous single-gate demonstrations, the thesis provides comprehensive experimental measurements of co-integrated 1T-1R crossbars and variability-aware circuit simulations to quantify the error probabilities of single logic gates. Furthermore, it identifies critical operational parameters to improve the reliability. On the example of a full adder, a hybrid approach is discussed which combines readout-based and switching-based logic in order to optimize the reliability.

These methodologies can be adapted to other resistive-based logic gates and may be extended beyond filamentary VCM devices to explore a broader range of resistive technologies.

Kurzfassung

Valenzwechselspeicher (VCM) als Unterkategorie der Redox-basierten resistiven Random Access Memories (ReRAM) sind mögliche Kandidaten, um CMOS-basierte SRAM in Datenspeicheranwendungen zu ersetzen. Diese Bauelemente bieten Nichtflüchtigkeit, Skalierbarkeit und eine erhöhte Zuverlässigkeit, wodurch sie sich gut für Datenspeichersysteme eignen. Die Integration von filamentärschaltenden VCM-ReRAM Zellen mit komplementären Metal-Oxid Halbleiter (CMOS)-Transistoren in 1T-1R-Crossbar-Arrays verbessert die Bauelementkontrolle weiter und ermöglicht die Entwicklung von resistiven Architekturen auf großer Skala. Neben der reinen Datenspeicherung eignen sich resistiv schaltende Speicherelemente auch für die Ausführung von Rechenoperation direkt im Speicher, dem Computing-in-Memory (CIM), einschließlich Boolescher Logik und arithmetischer Operationen. Resistive Logik befindet sich jedoch noch in der Entwicklung, da die Variabilität der Bauelemente und Anforderungen in der Integration eine groß angelegte experimentelle Verifizierung einschränken. Daher werden Simulationen auf Systemebene zur Vorhersage der Leistung von Rechensystemen basierend auf resistiv-schaltenden Bauelementen genutzt.

Während der Energieverbrauch und die Latenz einfach vorherzusagen sind, sind die Zuverlässigkeit und die Fehleranfälligkeit in resistiven CIM-Systemen noch weniger untersucht. Aufgrund der inhärenten Bauelementvariabilität reicht der Nachweis eines einzelnen Logikgatters nicht aus, um die Zuverlässigkeit von größeren Systemen mit Tausenden von Operationen zu garantieren. Die vorliegende Arbeit versucht diese Lücke zu schließen mittels einer umfassenden Untersuchung von Operationen in einzelnen Logikgattern in 1T-1R CMOS/ReRAM-Crossbar-Arrays und durch eine Kombination aus experimentellen Studien und variabilitätsabbildenden Simulationen.

In dieser Arbeit werden die zwei wichtigsten Zweige der resistiven Logik untersucht: lesebasierte und schaltbasierte Operationen. Die lesebasierte Logik beruht auf der Erfassung von Widerstandszuständen als analoge Größen und deren Klassifizierung als digitale Werte. Ihre Zuverlässigkeit ist durch die Genauigkeit der Zustandserfassung und die Überlappung der Widerstandsverteilungen begrenzt. Die schaltbasierte Logik hingegen basiert auf einem bedingten Zustandsübergang aufgrund eines eingangsabhängigen Spannungsteiler. Hier wird die Zuverlässigkeit nicht nur durch die Variabilität der Widerstandszustände, sondern auch durch die stochastische Natur des Schaltmechanismus bestimmt. Daher erfordern beide Logikkonzepte eigene Ansätze, um ihre Fehleranfälligkeit abzuschätzen.

Durch die systematische Analyse beider Konzepte werden in dieser Arbeit Methoden zur Bewertung und Verbesserung der Zuverlässigkeit beider Typen vorgeschlagen und die Ergebnisse auf komplexere Rechensysteme extrapoliert. Aufbauend auf früheren Einzelgatter-Demonstrationen enthält die Arbeit umfassende experimentelle Messungen von ko-integrierten 1T-1R Crossbar Arrays und variabilitätsabbildende Schaltungssimulationen zur Quantifizierung der Fehlerwahrscheinlichkeiten einzelner Logikgatter. Darüber hinaus werden kritische Betriebsparameter zur Verbesserung der Zuverlässigkeit identifiziert. Am Beispiel eines Volladdierers wird ein hybrider Ansatz diskutiert, der auslesebasierte und schaltbasierte Logik kombiniert, um die Zuverlässigkeit zu optimieren.

Die verwendeten Methoden lassen sich auf andere widerstandsbasierte Logikgatter übertragen und auch über filamentäre VCM-Bauelemente hinaus auf ein breiteres Spektrum von resistiv-schaltenden Technologien ausdehnen.

Acknowledgements

This dissertation was written during my time as doctoral researcher at the *Institut für Werkstoffe der Elektrotechnik II* of the RWTH Aachen University. During this time, I have been fortunate to collaborate with a group of exceptional colleagues and contributors whose support was essential in the completion of my dissertation.
First, I would like to thank **Prof. Dr. Rainer Waser** for providing me with the opportunity to pursue my research at his institute and for supervising my thesis. My appreciation also goes to **Prof. John Paul Strachan**, Ph.D. for serving as the co-examiner of my work.
I own many thanks to my group leader **Dr. Dirk Wouters** for his assistance, guidance, and engaging discussions on my research journey. I also greatly appreciated **Dr. Stephan Menzel**'s valuable support regarding the topic of resistive switching devices.
My work benefited from the collaborative efforts within the ROBCOMM-project alongside the *Chair of Dependable Nano Computing* at the Karlsruhe Institute of Technology. Therefore, I am thankful to **Prof. Dr. Mehdi Tahoori** and his team for their collaboration and access to the server infrastructure. In particular, I want to thank **Atousa Jafari** for her support throughout the years.
The experimental segments of my thesis significantly profited from the expertise on analog circuit design shared by **Ricardo Heinen** and **Jonas Zoche** from the *Lehrstuhl für integrierte Analogschaltungen*, as well as contributions from the *Peter-Grünberg-Institut 7*, namely **Dr. Felix Cüppers** and **Danielle Storelli** in fabricating VCM devices, **Dr. Stefan Wiefels** and **Xiaohua Liu** for their role in conducting measurements and **Dr. Sebastian Siegel** for his contribution to the MAD-200 project.
My gratitude extends to **Dr. Ulrich Böttger** and **Martina Heins** for their invaluable administrative support, **Udo Evertz** for his great technical assistance, and Thomas Pössinger for his aid with illustrations. I further acknowledge the support of my student assistants **Mahdi Najjar** and **Jinyun Park**.
Another full-hearted thank you goes to all my colleagues at the *IWE 2* for providing a supportive and friendly work environment. In particular, **Dr. Tyler Hennen** for his collaborations and the Bachelors-group, **Dr. Fenja Berg**, **Dr. Nils Kopperberg**, **Daniel Schön**, **Dr. Nils Sommer** and **Pascal Stasner** for thoughtful discussions and fun activities.
For the enriching office atmosphere and fruitful discussions I would like to thank **Dr. Christopher Bengel**. My deepest appreciation goes to **Tobias Ziegler**, who has already supervised my master's thesis and later became an office mate, research collaborator and friend.
Lastly, I cannot express enough thanks to my family and friends, whose encouragement has supported me through these years.

Contents

1 Introduction **1**

2 Fundamentals **5**
2.1 Classification of Non-Volatile Emerging Memories 5
2.1.1 Bipolar Filamentary VCM-ReRAM 6
2.1.2 Reliability Issues in VCM-ReRAM 9
2.1.2.1 Switching Variability 10
2.1.2.2 Read Noise 12
2.1.2.3 Endurance and Retention 13
2.2 Application of Redox-based Memristive Switching Devices 14
2.2.1 Memristive Crossbars 14
2.2.1.1 Peripheral Circuitry 16
2.2.2 Memristive Crossbars for Computation-in-Memory 18
2.2.3 Classification of Memristive Logic CIM Concepts 19
2.2.4 Requirements for Memristive Logic CIM Operations 20
2.3 Statistical Tools 21
2.3.1 Normal Distribution 21
2.3.2 Log-Normal Distribution 22
2.3.3 Overlap between Probability Density Functions 22
2.3.4 Fitting of Cumulative Distribution Functions 23
2.4 Simulation and Measurement Tools 23
2.4.1 Simulation Tools 23
2.4.1.1 Juelich-Aachen Resistive Switching Tools 23
2.4.2 Measurement Methods 26
2.4.2.1 MAD-200 testchip in $130\,\mathrm{nm}$ 26
2.4.2.2 ELMOS testchip in $350\,\mathrm{nm}$ 26

3 Readout-based Logic Operations **31**
3.1 Fundamentals of Readout-based Logic Operation 31
3.2 Primitive Gates for Readout-based Logic 33
3.2.1 Error Estimation Framework for Readout-Based Logic 33
3.2.2 Single Device Readout 35
3.2.2.1 Characterization of Bitcell Variability 35
3.2.2.2 Influence of the Operation Voltage 40
3.2.2.3 Influence of the Operation Position 46

3.2.2.4 Read Instability in Device Classification . . . 48
3.2.3 Two Input Operations . . . 55
3.2.3.1 Characterization of Two Device Operations . . . 55
3.2.3.2 Influence of the Operation Voltage . . . 60
3.2.3.3 Influence of the Operation Position . . . 67
3.2.4 Multi-bit Evaluation . . . 69
3.2.5 Transistor-specific Logic Operations . . . 77
3.2.5.1 Single Device . . . 77
3.2.5.2 Two Device Operations . . . 82
3.3 Summary of Non-Stateful Readout Logic . . . 86

4 Switching-based Logic Operations 89
4.1 Fundamentals of Switching-based Logic Operation . . . 89
4.2 Primitive Gates for Switching-based Logic . . . 92
4.2.1 Error Estimation Framework for Switching-Based Logic . . . 92
4.2.2 Operative Conditions . . . 95
4.2.3 RESET-based Switching Logic . . . 98
4.2.4 SET-based Stateful Switching Logic . . . 100
4.2.4.1 Characterization of Operation Reliability . . . 102
4.2.4.2 Input State Conservation . . . 109
4.2.4.3 Framework Benchmarking . . . 112
4.2.5 Extension of Two Input Logic . . . 115
4.2.6 Transistor-specific Logic Operations . . . 121
4.3 Summary of Stateful Switching Logic . . . 125

5 Crossbar Operations for Logic & Arithmetic Computing-in-Memory Evaluation 127
5.1 Parallel Multi-Gate Evaluation . . . 127
5.2 Arithmetic Operations on Crossbar Level . . . 137
5.2.1 Error Propagation for Concatenated Operations . . . 137
5.2.2 Full-Adder Implementation . . . 137

6 Conclusion 141
6.1 Summary . . . 141
6.2 Outlook . . . 143

A Appendix 145
A.1 Derivation of Intersection Points for Normal Distributions . . . 145

Bibliography 147

List of Figures 171

List of Tables 175

CHAPTER 1

Introduction

The invention of the first semiconductor transistor in 1947 at Bell Laboratories by John Bardeen, William Shockley, and Walter Brattain [1] stands as one of the most impactful technological achievements in human history. Recently, the rapid development of artificial intelligence (AI) has emerged as the next technological advancement. Large-Language Models (LLMs) and AI assistants now perform remarkable tasks and often surpass human capabilities. This progress might suggest that AI model software has become the cornerstone of technological progress. However, the increasing demand for data centers and their significant environmental and energy costs highlight an undeniable reality: every software optimization is fundamentally dependent on the computing hardware. The Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) remains the most widely used transistor type for building such hardware. Therefore, continuous efforts are driven to enhance its performance such as shrinking the integration size below $3\,\mathrm{nm}$ and the development of 3D vertical architectures [2]. Besides this, a second research focus explores potential MOSFET replacements or extensions. Examples are the power MOSFETs based on novel material systems like GaAs or SiC, optoelectronics and photonics for computational applications, and solid-state memory technologies for data storage. Among all data storage alternatives, the Valence Change Mechanism (VCM) Redox-based Resistive Random Access Memories (ReRAM) are the focus of this thesis. These devices offer a high technology-readiness level for storage applications due to their non-volatility, scalability, and significant improvements in reliability. Several industry activities launched recently in this field for instance Weebit Nano Ltd. [3] or the Taiwanese semiconductor world-market leader TSMC [4] collaborating with the German manufacturer Infineon [5].

Besides their usage in data storage, the unique properties of resistive switching memories make them suitable candidates for Computing-in-Memory (CIM) applications. Particularly the analog switching capability enables significant advantages in the completely new paradigm of neuromorphic or neuro-inspired computing. At the same time, ReRAM devices are also useful for existing concepts, such as Boolean logic operations and arithmetic functions, which paves the way for improving classical computation. The resistive logic Computing-in-Memory is still in a developmental phase. Numerous single-operation concepts have been proposed and experimentally demonstrated in passive resistive or active co-integrated ReRAM/CMOS structures [6–8]. Based on this device-level exploration, high-level simulators and hardware/software co-integration mappings have been proposed to predict and optimize the computing functionalities [9–13]. Yet due to variability issues and the complexity of integrating such systems in actual hardware, the experimental verification of the resistive computing algorithms on a system level remains challenging. Therefore, the simulations and theoretical system-level contributions are critical for identifying potential use cases of resistive computing.

While energy consumption and latency are relatively straightforward metrics to evaluate, the reliability and error rates in resistive CIM systems have not yet been explored in sufficient detail. Established concepts from CMOS, such as the Architectural Vulnerability Factor (AVF), cannot be directly applied to resistive devices due to fundamental differences in their operation [14]. This implies that the successful demonstration of a single logic gate does not guarantee the reliable operation of a complete computational system involving thousands of steps and highly variable devices. This thesis aims to bridge this gap by conducting a statistical investigation of single-gate logic operations in resistive devices. In this regard, it focuses on co-integrated ReRAM/CMOS architectures utilizing 1 transistor and 1 resistive memory cell (1T-1R) per bitcell. The inclusion of a transistor is the current state-of-the-art for integrating large memristive crossbar as it offers several operational advantages, such as improved isolation and device control. The experimental verification is accompanied by variability-aware circuit simulations. This is required to propose a methodology for extrapolating the single-gate results towards the reliability of more complex, concatenated algorithms and systems. By systematically evaluating the reliability of fundamental resistive-based logic elements, this thesis provides insights into the scalability and practical implementation of larger CIM systems.

Two contrary branches of resistive logic CIM are investigated in this thesis: readout-based logic operations and switching-based logic operations. Readout-based logic operations rely on sensing the resistive states of one or multiple devices as an analog quantity and interpreting these states to perform logic operations. In contrast, switching-based logic operations depend on the conditional switching of a resistive device, where the state transition serves as the basis for computation. The fundamental differences in these concepts lead to distinct influences of variability on their performance. In readout-based logic, the reliability is primarily limited by the accuracy of the state sensing, introducing challenges in distinguish between overlapping resistive distributions. On the other hand in switching-based logic, variability plays a critical role in the switching dynamics, including the resistive switching voltage and switching time, which can significantly alter the operation's correctness. These inherent differences highlight the need for specific approaches to evaluate and optimize the reliability in each type of operation.

The thesis is structured to investigate the logic errors in both branches of resistive logic Computing-in-Memory, readout-based and switching-based, before extrapolating the findings from the single logic gates to more complex concatenated operations. Chapter 2 introduces Redox-based Resistive Random Access Memories (ReRAM) as non-volatile memory technologies, focusing on filamentary-type valence-change mechanism devices, their variability issues and their role in Computing-in-Memory. Furthermore, the exploited measurement setups, simulation environments, and statistical tools used in

this thesis are presented. Chapter 3 focuses on readout-based logic gates. Starting with a single 1T-1R readout, the key factors which influence a successful state classification are investigated. These insights are then transferred to logic operations with an increased complexity as the number of inputs scales from two-inputs, three-inputs up to five-input gates. Finally, the role of the transistor not only as a selection device but also as an active logic element is discussed. Subsequently, Chapter 4 addresses switching-based logic gates. After a theoretical discussion of the switching-based paradigm and the challenges posed by state conservation, the investigation narrows to two-input operations based on the resistive SET process. This subcategory is intensively studied experimentally as these operations have been rarely demonstrated in literature before. The chapter also evaluates the use of the transistor as an active element in switching-based logic. In Chapter 5, the insights from the previous single readout-based and switching-based operations are extrapolating toward concatenated higher-level applications. In particular, the specific challenges of achieving efficient parallel computing are discussed for the switching-based operations. As a practical demonstration, the chapter proposes a combination of readout-based and switching-based logic to optimize the performance of a carry-ripple full adder. Finally, Chapter 6 provides a summary of the thesis and outlines potential future research directions for extending the field of resistive Computing-in-Memory.

CHAPTER 2

Fundamentals

The aim of this chapter is to provide the reader with all necessary background information to understand the following research sections. Thus, it contains fundamental concepts as well as the current state of research related to the topic of this thesis.

At first, it gives an overview of non-volatile emerging memory technologies with a more detailed focus on resistive switching devices, their variability challenges and the corresponding Computing-in-Memory concepts which are the main topic of this thesis. The remaining three sections contain a summary of the research methods that were exploited. This includes statistical methods such as probability distributions and calculations, simulation methods such as simulation environments and models, and a description of the measurement setups and measured devices.

2.1 Classification of Non-Volatile Emerging Memories

In information technology, any abstract information or data has to be stored in a physical manner as a hardware memory. This principle is summarized by Rolf Landauer as "Information is a Physical Entity" [15].

Over the years, different physical entities have been exploited to realize the physical storing of information. Yet, in modern-day electronic devices, the binary representation of data as "0" or "1" is dominating this field. Following the establishment of Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) technology, the working memory classes of Random Access Memory (RAM), specifically static RAM (SRAM) and dynamic RAM (DRAM), as well as the storage Flash memories (NAND-Flash, NOR-Flash), have emerged as the most prominent types of memory devices. Despite the dominance of these MOSFET-based technologies, novel emerging memory concepts have gained attention, including nano-ionic resistive switching Redox-based Random Access Memory (ReRAM). A broad hierarchical position and classification of ReRAM in the field of solid-state memories is illustrated in Fig. 2.1.

ReRAM devices store information within their electrical resistance which can be influenced via field-induced local redox processes. This class subdivides depending on the switching process as Thermo-Chemical Memory (TCM), Electro-Chemical Metallization (ECM) memory, or Valence Change Mechanism (VCM) memory. Each category can further be distinguished regarding the switching type (area-depending or filamentary) and the switching polarity (unipolar or bipolar) [18–20].

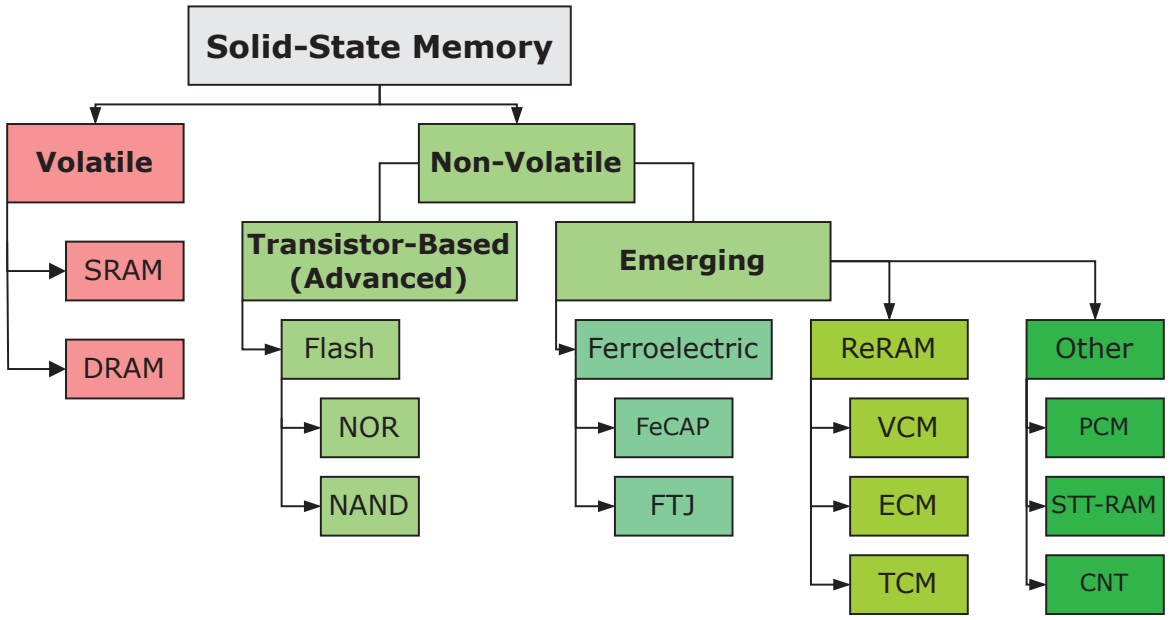


Figure 2.1: Flowchart for the taxonomy of solid-state memories. It should be emphasized that the illustration serves as a broad categorization only. There may be overlaps among certain memory types depicted in the diagram. All abbreviations are referred to in the text. Based on the 2022 International Roadmap of Devices and Systems - Beyond CMOS and Emerging Materials Integration [16] and [17, 18].

Within the sphere of emerging memory technologies, another notable subclass next to resistive RAM is that of ferroelectric memories. This subclass contains devices such as the Ferroelectric Capatitor (FeCAP) and the Ferroelectric Tunnel Junction (FTJ). Complementing these are other innovative storage types, including the magnetic Spin-Torque-Transfer RAM (STT-RAM), Carbon-Nanotube (CNT) transistors, and Phase-Change Memory (PCM). Each of these technologies offers unique characteristics and potential applications that distinguish them from traditional memory solutions.

2.1.1 Bipolar Filamentary VCM-ReRAM

The bipolar Valence Change Mechanism is one of the most prominent switching processes in the ReRAM class. It is observed in Metal-Insulator-Metal (MIM) structures with a Mixed-Ion-Electron Conductor (MIEC) as insulator. Several transition metal-oxides such as HfO_2, ZrO_2, Ta_2O_5 or TiO_2 are known as possible MIECs [18, 21, 22]. In filamentary type switching, the redox processes are locally confined within the metal-oxide. For completeness, the area-dependent switching also has to be mentioned, where field-induced switching reaction occurs on the whole cross-section of oxide. This work however focuses on filamentary switching devices.

The switching mechanism of VCM devices is based on the local ion-migration of oxygen vacancies ($V_O^{\cdot\cdot}$) or interstitial metal cation within the metal-oxide [23]. For the sake of simplicity, the following explanation solely focuses on the oxygen vacancies as local charge donors. In general, a high concentration of oxygen vacancies corresponds to a high electric conductivity, i.e., a small electrical resistance, while areas of low oxygen vacancy concentration show a low electric conductivity and therefore a higher electrical resistance. Under applied electric fields the oxygen vacancies can move

locally within the metal oxide which will lead to an increase or decrease of the MIEC electrical resistance and the MIM resistance consequentially.

In a MIM with a MIEC as insulator, the two metal electrodes are usually selected in a way to differ in their work functions and electron affinities. The electrode with higher work function and lower oxygen affinity is called (electronically) active electrode (AE), the electrode with lower work function and higher oxygen affinity is called ohmic electrode (OE). This refers to the ohmic contact which forms between the OE and the metal oxide while at the interface between metal oxide and AE a Schottky-type contact behavior is observed.

After fabrication, the metal oxide is in a pristine, highly insulating state. An initial electroforming step is necessary in which the oxide is locally reduced and oxygen vacancies are created [24]. The electrical resistance of the device reduces drastically during this electroforming step. Therefore it is necessary to include a current compliance, either by an additional transistor or resistor in series. This prevents the destruction of the device due to high currents. During the electroforming step in filamentary-type ReRAM, a highly n-conductive filament is formed which is confined to a small cross-section of the total cell area. The structure of this filament can be simplified to a highly conductive *plug* region, which reaches from the ohmic electrode towards the active electrode. Between the interface of the filament in the metal-oxide and the active electrode a potential barrier or *disc* region is formed. In this picture, the residual plug region acts as oxygen vacancy reservoir and shows a constant, very low electrical resistance. In contrast, the electrical resistance of the disc region depends highly on the potential barrier height which is determined by the concentration of oxygen vacancies. By applying an electrical voltage, oxygen vacancies are moved into/out of the disc region. This movement reduces or increases the potential barrier height [22]. A lowering of the oxygen vacancy concentration in the disc region results in a high potential barrier and a high electrical resistance. This is called a *RESET* transition, which results in the *OFF state* or high-resistive state (HRS). If vacancies are moved into the disc region, the potential barrier is reduced and so is the electrical resistance. This transition is called *SET*, which leads to the *ON state* or low-resistive state (LRS). The total device resistance of the MIM depends mainly on the MIEC resistance, which again strongly depends on the disc region resistance. Consequentially, the VCM switching process can be simplified to the movement of oxygen vacancies into or out of the disc region and the modulation of the potential barrier height at the AE-MIEC Schottky interface [18, 23]. Since the voltage polarity with respect to the AE controls the migration direction of the oxygen vacancies, the switching is described as bipolar switching. If negative voltage from AE to OE results in a SET transition, the switching is called *counter-eightwise* or *c8w* for short. If the same voltage polarity results in a RESET transition, the switching is called eightwise or *8w* respectively [25]. A *c8w* switching cycle for a bipolar filamentary VCM cell after the initial electroforming process is illustrated in Fig. 2.2. In the depicted device, platinum (Pt) forms the AE, zirconium (Zr) the ohmic electrode and zirconium oxide (ZrO_x) the MIEC metal oxide.

The bipolar filamentary resistive memory stands out as one of the leading contender for beyond-CMOS storage-class memory and the implementation in emerging computing paradigms such as Computing-In-Memory (CIM) [16]. This is emphazised by the rising engagement of conventional semiconductor companies in ReRAM technology, for instance Weebit Nano Ltd. [3] or the Taiwanese semiconductor world-market leader TSMC [4], who also recently announced a collaboration with German technology company Infineon [5]. VCM-ReRAM classifies for the use as solid-state memory as it provides at least two stable states, namely HRS and LRS [26]. Along with the potential for multi-level storage [27] and a stable analog switching [28] it positions itself at promising candidate for beyond-CMOS computing. These states demonstrate stability over time (retention) [29–31] and predictable reproduction across multiple cycles (endurance) [32, 33]. The state transitions (switching)

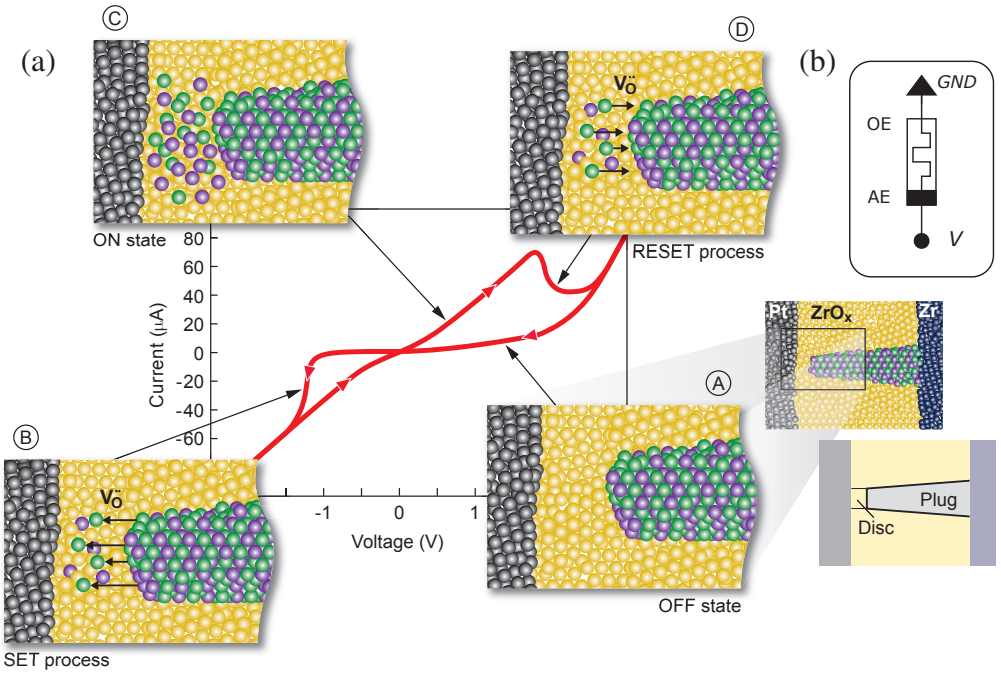


Figure 2.2: (a) Idealized voltage-current (*I-V*) characteristic and atomistic illustration of a counter-eight-wise valence change mechanism switching process in a zirconium (Zr)/zirconium-oxide (ZrO_x)/platinum (Pt) device stack. The zirconium-oxide as Metal-Mixed-Ion-Electron Conductor is represented as yellow spheres, the conductive filament of mobile oxygen vacancies ($V_O^{\cdot\cdot}$) and immobile zirconium cations as green and violet spheres, and the metallic platinum as grey spheres. In the beginning (Ⓐ) the device is in the high-resistive state (HRS) or *OFF state*. Due to the low concentration of oxygen vacancies in the *disc* region between the conductive filament and the platinum electrode, the cell shows a high resistance. If a negative voltage with respect to the active platinum electrode is applied (Ⓑ), oxygen vacancies drift from the *plug*-region into the disc region. This increases the $V_O^{\cdot\cdot}$ in the disc and reduces the cell resistance. After this SET process the cell is in the low-resistive state (LRS) or *ON state* (Ⓒ). By applying a positive voltage w.r.t. the active platinum electrode, a RESET process is induces (Ⓓ). During this process, oxygen vacancies are dragged from the disc region into the conductive plug again. This leads to a low $V_O^{\cdot\cdot}$ concentration in the disc and therefore a high cell resistance again (Ⓐ). Reproduced with permission from [25, p.700]. ©2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim. (b) The circuit symbol used in the latter to represent a filamentary-type bipolar VCM device. Note that subfigure (b) is not part of the original published image.

can be induced by applying sufficiently high electrical fields while a non-destructive state sensing (read-out) is achievable with lower electric fields. The non-linear characteristic of the VCM switching ensures a large voltage window between both operations [34]. Furthermore, material properties such as compatibility with CMOS co-integration and 3D scalability show promising results, although they remain areas of ongoing research with room for improvement [16, 26].

In the following, the terms *memristive devices* or *resistive devices* are used in reference to the resistive switching ReRAM devices, especially the investigated VCM filamentary-type devices. The attribute "memristive" is a combination of the words "memory" and "resistance", as it describes a device which can store an information (memory) in its (electrical) resistance. This terminology is related to the term "Memristor" which was originally created by Leon Chua in order to describe a missing theoretical 4th circuit element next to resistor, capacitor and inductor [35]. Starting from this theoretical formulation, the term was later used to classify the real-world examples of resistive switching devices [36, 37]. Since the accordance of the theoretical postulated "Memristor" with the real-world resistive switching phenomena is questionable [38], the terminology *memristive devices* or *resistive devices* are used throughout this thesis to focus on the experimentally observable devices independent of the theoretical discussion.

2.1.2 Reliability Issues in VCM-ReRAM

Up to this point the fundamental aspects of filamentary VCM devices were covered under idealized assumptions for the filament forming and its behavior during the switching process. Now, the focus is shifted towards the critical examination of non-idealities which are inherent in VCM-type memristive devices. Understanding these effects is substantial for the effective utilization of memristive devices in various applications. The suitability of memristive devices for the utilization in data storage or computing applications depends on their ability to reproduce their electrical states and properties. This includes the accessible range of electrical resistance values, the possibility to precisely switch the device within this range or at least between the two sub-states, HRS and LRS, and the device's ability to maintain its state after switching for an arbitrary duration [39]. In general, this kind of reliability analysis is not solely an issue of memristive devices but a crucial aspect across all domains of electronics [40, 41]. However as the memristive device research is not as mature as competing CMOS technologies, the understanding and mitigation of variability effects is not yet as advanced. Investigating the variability, analyzing its effect on memristive logic circuits and potentially identifying viable solutions to ensure the circuit reliability, is the focus of this thesis.

Variability in filamentary VCM-type ReRAM is assumed to be caused by the inherent filament stochasticity and the stochastic nature of the underlying redox reaction [42–45]. The stochasticity manifests itself at various stages in the memristive device operation. This begins during the initial filament formation, throughout all subsequent switching processes and even in absence of any applied voltage. Two major effects can be distinguished in this regard. Firstly, the field-driven migration of oxygen vacancies into or out of the filament during a switching cycle. While the migration of oxygen vacancies is directed by the external electric field, it remains stochastic. Secondly, there are uncorrelated, random shifts in the ionic filament configuration which are thermally driven by the ambient temperature.

As explained above, the electrical conductivity of the ReRAM device depends on the potential barrier at the Schottky-like contact at the interface between active electrode and oxide layer. Despite the estimation that this potential barrier has a constant value, it shows a complex spatial dependency on the three-dimensional filament geometry and the local charge carrier concentrations. Since these parameters are not constant attributes but exhibit also spatial and temporal variations, both the electrical

resistance of the device and its response to external fields, i.e., the switching process, are subject to variations [43, 44]. These variations include switching variability, read noise, endurance and retention.

Switching variability includes all variations occurring during the field-induced switching process, typically between the two states, HRS and LRS. Closely related is the *endurance* or ability to cycle which denotes the degradation and potential breakdown of the device due to repeated switching cycles. In contrast, *read noise* refers to variations in the device resistance during a quasi-static situation with no or only a low applied voltage without initiating a switching process. While the term read noise is typically explored for short timescales, the long-term state stability is described as *retention*.

2.1.2.1 Switching Variability

The electrical resistance of a VCM-type device after the switching operation is an analog value within a certain continuous range [18, 27]. This analog feature can be exploited as an advantage for performing analog vector-matrix-multiplications based on memristive devices for instance in artificial neural network application [46, 47]. Unfortunately, the analog capabilities of ReRAM devices demonstrate several issues such as state drift or degradation [48, 49]. Therefore it might be beneficial to divide the analog resistance range into only two sub-ranges, LRS and HRS. In this context the switching variability describes the observation that under identical external switching conditions such as applied voltage amplitude, pulse width or ambient temperature, the memristive device will in general not show the same posterior switching electrical resistance. Rather the resulting states, HRS and LRS, are part of a distribution [50].

The switching variability can be further categorized into two major types: cycle-to-cycle (C2C) and device-to-device (D2D). The prior one, C2C variability, describes the observed variations between multiple switching processes of the same memristive device. On the other hand, D2D variability refers to the variations between different memristive devices while being exposed to the same switching conditions. This distinction is useful as the D2D variability encompasses variations which are introduced during the device fabrication and initial filament formation. Given that variations in the device composition (e.g., oxide thickness) or the fabrication quality (e.g., number of material defects) are intrinsic to a specific device and cannot be corrected retroactively, it becomes substantial to minimize these variations during the fabrication process. It has to be noted that although the origin of both variations can be separated theoretically, the statistical evaluation is much more challenging. This is due to the fact that for a measured device all switching variations overlap and therefore dedicated statistical methods have to be used [51].

VCM devices and their switching variability have been subjected to intensive experimental investigations [52–56] as well as comprehensive theoretical analysis [44, 57, 58]. A preferred experimental methodology to study the switching kinetics of ReRAM devices is a current-voltage ($I - V$) loop measurement. These measurements are performed by applying continuous triangular voltage sweeps, which alternate between the SET and RESET polarity [50].

Noticeable shifts towards a lower LRS and a higher HRS indicate that an increase in voltage amplitude coupled with slower sweep rate tends to enhance the switching process, both in the SET [28, 52, 59, 60] and RESET [52, 59, 61] direction. However, it is important to note that the interplay between voltage amplitude and sweep rate shows a strong non-linear behavior [28, 62–64].

For example, in the context of a SET transition it is observed that a linear increment in the voltage amplitude leads to an exponential decrease in the time required for switching into comparable posterior resistance states [60, 62, 65]. Additionally, the SET process appears very abrupt which is related to the thermal runaway of the device filament [34, 66]. Nevertheless, when the voltage amplitudes reduce, the

duration for observing a SET switching grows exponentially [28, 62]. This abruptness and threshold behavior in the voltage amplitude is a critical factor in determining the success of a SET switching. In contrast to this abrupt SET behavior, the RESET switching is observed as a gradual process which allows for programming also intermediate HRS values [61, 67].

While the aforementioned insights into the behavior of memristive devices under triangular sweep conditions are valuable, a more critical consideration for their practical application lies in the understanding of their response to rectangular pulse conditions. In the context of electronic applications, the implementation of pulses with constant width and amplitude is more feasible within the mixed digital/analog circuit periphery. This preference for fixed pulse conditions is dictated by two key components: the circuit clock frequency, which determines the pulse width, Δt, and the available Digital-Analog Converters (DAC), which set the pulse amplitude, V_{pulse}. These elements play a dominant role in defining the operational parameters for memristive devices within electronic systems. Therefore studying the operational parameters for a pulse-based switching is essential for the effective integration and performance of these devices in real-world applications.

First, the voltage amplitudes which are necessary to trigger the SET or RESET transition show a significant variability [52, 55]. This variation present a severe challenge in determining a single, fixed pulse condition as V_{pulse} and Δt that reliably changes into the desired state, either in the SET or RESET direction. As a consequence, the outcome of applying a specific pulse is more effectively expressed in terms of a probability which predicts the resultant distribution of the device's posterior resistance [68–71].

A further complication lies in the observation that a state change in memristive devices is influenced not only by the external pulse conditions but also by the device's prior states. In the simplest case, only the initial device resistance before the switching event has to be considered [72, 73]. However, there are more complex models which extend this dependency on the device's entire switching history [51, 74]. This phenomenon of a device *memory* is linked to the alterations in the conductive filament which shifts its shape with every switching cycle [72, 74].

A feasible opportunity is to adjust the parameters, V_{pulse} and Δt, in order to improve the post-switching resistance distributions. For instance, increasing the SET voltage amplitude can increase the switching probability. Therefore, more devices would result in a LRS resistance. However, this approach encounters two operational challenges: Firstly, increasing the pulse duration results in a reduced operation speed since each individual switching requires more time. This reduction can be a significant drawback for high-throughput applications. Secondly, using higher voltages necessitates the use of suitable peripheral components including appropriate transistors as well as DACs that can handle elevated voltages. This requirement imposes additional constraints on the circuit design, e.g., for the required integration area. Utilizing higher voltages would also lead to an increased energy consumption which is a critical figure of merit for electronic applications. Furthermore, there is a internal limitation concerning the ReRAM device itself. Memristive devices can only tolerate switching voltage up to certain values before a permanent oxide breakdown is risked [75, 76]. This factor necessitates a wise trade-off between selecting the minimum pulse conditions that still reliable induce the desired switching operation.

2.1.2.2 Read Noise

In addition to the above mentioned switching variability, variations in the cell resistance are observed even when voltages below the switching threshold are applied or even under no external bias. Since this voltage range is typically exploited for the cell readout, the terms *read noise* or *read instability* are used to describe this phenomenon. The origin of these variations lies in local fluctuations occurring within the conductive filament [77, 78]. The conductance variations manifest as changes in the measured cell current I_{cell} or the respective cell resistance $I_{\text{cell}} = V_{\text{applied}} / I_{\text{cell}}$. The changes can be quantified as absolute changes Δ or as relative changes $\Delta R / R$. Unlike other electronic noises which are present as continuous signal, the read noise is characterized by discrete, sudden jumps in the conductance [79, 80].

The discrete current jumps are closely linked with the movement of individual oxygen vacancies within the conductive filament. This microscopic process significantly influences the macroscopic electrical device behavior. The oxygen vacancies can transition between adjacent lattice positions, which leads to the discrete nature of the read noise [77, 81]. While certain applications in the field of probabilistic computing appreciate this unpredictable behavior [82, 83], it marks a severe issue for accurate computing applications [27, 84, 85].

In addition, the presence of read instabilities limits the precision of so called *program-verify* approaches [86]. In these program-schemes, small write pulses are applied to a memristive device in order to incrementally increase or decrease the device's resistance. Each write-pulse is followed by a readout to verify the resistance change and determine the next pulse. However, the effectiveness of this method is constrained by the readout accuracy which is determined by the read precision.

It has been observed that as the voltage is increased, the magnitude of current fluctuations ΔI also increases [80, 87, 88]. In addition to the voltage impact, the measured read noise is also influenced by the sampling rate and ambient temperature. The increased read noise frequency at elevated temperatures indicates that the hopping rate of oxygen vacancies in the filament is increased at higher temperatures [77, 87]. Furthermore, recent studies show that the read noise phenomenon is also related to the underlying electronic conduction mechanism within the VCM device [89].

In the context of memristive device variability, the influence of read instabilities manifests primarily through two major effects: Firstly, it adds to the overall switching variability. As a memristive device is programmed by an arbitrary pulse and readout, the resulting resistance value will be a combination of the post-switching state distribution and the state-dependent read distribution. While both effects can be separated in theory, they are closely related in experimental measurements. Furthermore, the read noise complicates the investigation of any switching variability. When the result of a device readout after a switching operation is biased by an unknown read noise, it becomes challenging to isolate and understand the true switching variability.

A pragmatic approach to addressing the read noise challenges is to perform multiple readout measurements of each state and then average the resulting resistance values across all samples. This method provides a more stable and representative value of the actual state resistance. The repeated measurement technique is applicable and beneficial in scenarios where the desired information or application relies on a precise device state and is processed digitally. However, this approach fails when it comes to analog computation where the read instability directly affects the computational accuracy. Performing multiple measurements is not feasible in real-time analog computing or computing-in-memory approaches. Therefore alternative strategies must be employed to mitigate for the effects of read noise such as de-noising or error-tolerant computation algorithms.

2.1.2.3 Endurance and Retention

The aspects of endurance and retention are closely linked to the above described phenomenon of read noise and switching variability in memristive devices. Retention refers to the ability of a memristive device to maintain its resistance state over a certain time period in the absence of an external voltage or only under repeated readout operations. This long-term effect is connected to the short-term read noise [90]. The present oxygen diffusion induces local fluctuations in the oxygen vacancy configuration within the conductive filament close to the active electrode [30, 91]. This will influence the potential barrier and consequentially cause a change in the device electrical resistance state. As diffusive processes are generally temperature-dependent, the retention time decreases with elevated temperatures [30, 91]. The consequence of these fluctuations is a broadening of the measured read current distributions which will follow a log-normal statistic [30]. Similar to the read noise, the long-term state retention is a critical issue, especially in analog computation applications where precise resistance levels are crucial [92]. As a worst-case scenario, the state retention will lead to an overlap of adjacent analog resistances ranges, e.g., HRS and LRS, and effectively a corrupted information. In order for VCM devices to classify as non-volatile devices, the resistive state should remain stable for a period of roughly ten years with no significant retention effects during the time. This duration has been demonstrated recently also for co-integrated memristive devices [31]. Furthermore, emerging research suggests that it is possible to enhance the data retention by precisely calibrating the switching conditions [29] or improving the electroforming operation [93]. The underlying hypothesis is that such adjustments will lead to a more confined and stable conductive filament. This confinement is postulated to mitigate the impact of inherent oxygen vacancy fluctuations. While this topic still needs further research it shows a promising advancement in the optimization of memristive technologies, particularly their applicability in computing and memory applications.

On the other hand, the term endurance describes the capability of memristive devices to precisely reproduce distinct resistive states for an advanced number of programming cycles. In theory it is expected to program a memristive device reliably at any point of its lifespan within the bounds defined by the inherent switching variability. Yet in recent experimental investigation a decline in the programming capabilities of memristive devices was observed over a prolonged device use, i.e., increased number of programming operations applied to one device. This finding suggests that the repeated programming exerts a cumulative effect on the device material that leads to a gradual degradation. In a worst case the resistance window might completely collapse, i.e., the total programmable resistance range shrinks until one effective state is left [31, 48]. Furthermore, it has been observed that memristive devices may reach a state where they become non-responsive to applied switching pulses which is referred to as being *stuck* [31, 33]. This phenomenon represents a critical issue in the operation cycle of a memristive device where it effectively loses its ability to transition between different resistance states. Current reports suggest that the average lifespan of memristive devices endures 10^5 [31, 32] which might be maximized up to 10^7 - 10^{10} cycles for perfectly adjusted programming conditions [94–96]. It is important to mention that these adjustments aim at enhancing the endurance or programming reliability and often involve a trade-off which leads to compromises in other aspects, such as the elevated data retention, a reduced programmable window or an increased energy consumption.

Addressing these trade-offs impede general statements about the usage of memristive technologies and marks a key challenge in the development and application of memristive devices.

2.2 Application of Redox-based Memristive Switching Devices

Since their discovery, redox-based resistive switching devices have been utilized across a diverse range of application. Initially their ability for a non-volatile information storage was leveraged in RAM [97] which led to the term ReRAM as **Re**dox-based resistive **R**andom **A**ccess **M**emory. Since then this nomenclature evolved to refer not only to the specific RAM technology but also to individual devices as ReRAM devices although the application might differ from classic RAM functionalities [98, 99]. One of the key advantages of ReRAM especially when it is considered as a replacement for SRAM or DRAM is the static power reduction. Unlike SRAM or DRAM which require continuous refreshing cycles, ReRAM offers a non-volatile storage and thereby enhances the power efficiency [100–102]. Due to its blend of durability and performance, ReRAM can serve also as a storage class memory which bridges between enduring data retention in Flash storage and rapid data access of DRAM or SRAM memories [103, 104].

The usage of ReRAM devices as pure non-volatile data storage replacement has decreased while the interest in Processing-in-Memory (PIM) rose lately [97]. This shift reflects a broader trend in the computing sector as the focus moves towards the implementation of processing capabilities directly within the memory. This shift can be primarily attributed to the limitations inherent to the classic *von-Neumann* computing architecture. In a simplified picture, this architecture separates data storage (memory) and data processing (computing) [105]. For any executed computational task data has to be transferred from the storage unit to the processing unit. There it is temporarily held in the working memory (RAM) for processing. After this step the results are then transferred back to the memory for permanent storage. This constant shuttling of data between memory and processing units is described as *memory wall* and comes in hand with significant time and energy overheads [106–108]. While the processing speed increases continuously, the data transfer rates are lagging behind which results in an inefficient computational usage [109, 110]. These challenges are further amplified in the context of a rapidly escalating global demand for data storage and computing performance [111]. In the era of artificial intelligence where immense volumes of data are processed and analyzed, this trend is not only continuing but is expected to intensify significantly [112, 113].

The rest of this section first presents memristive crossbars as the prominent architecture used for organizing redox-based resistive switches in memory and processing applications. Afterwards a broad overview over the principles of processing-in-memory and computation-in-memory is given. Based on this the specific concepts relevant to this thesis are explored in-depth in the end.

2.2.1 Memristive Crossbars

A favorable form for organizing memristive devices are memristive *arrays*. In such structures orthogonal metal lines connect the metallic electrodes, OE and AE. At each line intersection or cross-point a memristive device or *bitcell* is located. This creates a matrix-like *crossbar* organization. The configuration offers substantial benefits, including high storage density and straightforward access to all devices via the appropriate selection of a specific metal line pair. In this setup each bitcell contains solely one memristive device. Hence it is referred to as *passive*-crossbar or *1R*-crossbar. Fig. 2.3 illustrates a simplified version of such a passive crossbar. Despite offering the maximum storage density and minimal footprint of $4F^2$, with F being the minimum feature size, the passive crossbar design confronts significant challenges. Notably the issues of sneak path currents and program disturb which arise from the lack of isolation and selectivity in the cells present considerable obstacles [114, 115].

To mitigate these issues, passive selector devices can be added in series to each bitcell. This results

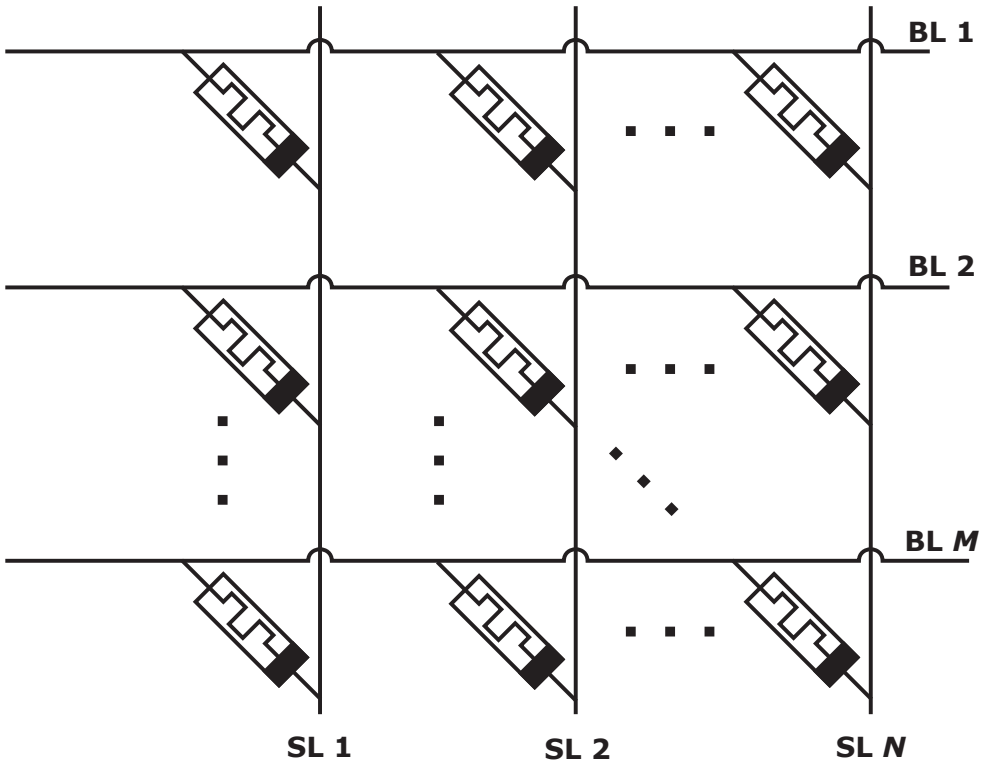


Figure 2.3: Circuit diagram of a passive crossbar array with $M \times N$ memristive devices. Horizontal Bit-Lines (BL) are connecting the OE while vertical Signal-Lines (SL) are connecting the AE.

in so-called *1-Selector 1-Resistor* or *1S-1R* configurations [116, 117]. The two-terminal selector devices can be stacked directly on top of the resistor which reduces the cell footprint. The passive selectors however face complications in fabrication and operation [115, 118]. Consequently the focus of industry and research is shifted towards active crossbar designs. Here, each bitcell consists of a memristive device in series with a MOSFET, leading to a *1-Transistor 1-Resistor* or *1T-1R* arrangement. The 1T-1R design effectively resolves the sneak path issues as it enables an individual device access without affecting adjacent bitcells. Furthermore, the transistor can serve as a current-control mechanism which is important for multi-level programming and analog switching. The approach of co-integrating memristive devices with existing MOSFET technology has been already demonstrated successfully in academia [119, 120] and by the semiconductor industry such as Intel [102], TSMC [4, 121], and Infineon [122].

The 1T-1R arrangement is favorable for its cell isolation and improved noise immunity. However a transistor has four electrical contacts, source, bulk, drain and gate, which requires a higher area footprint and additional circuitry for accessing the transistors. This complexity increases the overall energy consumption. Moreover, the transistor node selection and the specific circuit design introduce their own set of challenges. The transistor node dictates the minimum feature size of the MOSFET devices. In general it is always aimed for a small transistor node as it allows for a higher integration density. The trend of miniaturization is demonstrated by the successful integration of oxide-based RRAM technology on advanced technology nodes of $28\,\mathrm{nm}$-$22\,\mathrm{nm}$ [32, 121, 123] and reaches its current limit at $14\,\mathrm{nm}$ [124]. While downscaling is beneficial for achieving high-density computation and storage, it also necessitates a reduction in the supply voltage to avoid damaging the sensitive transistor structures. For

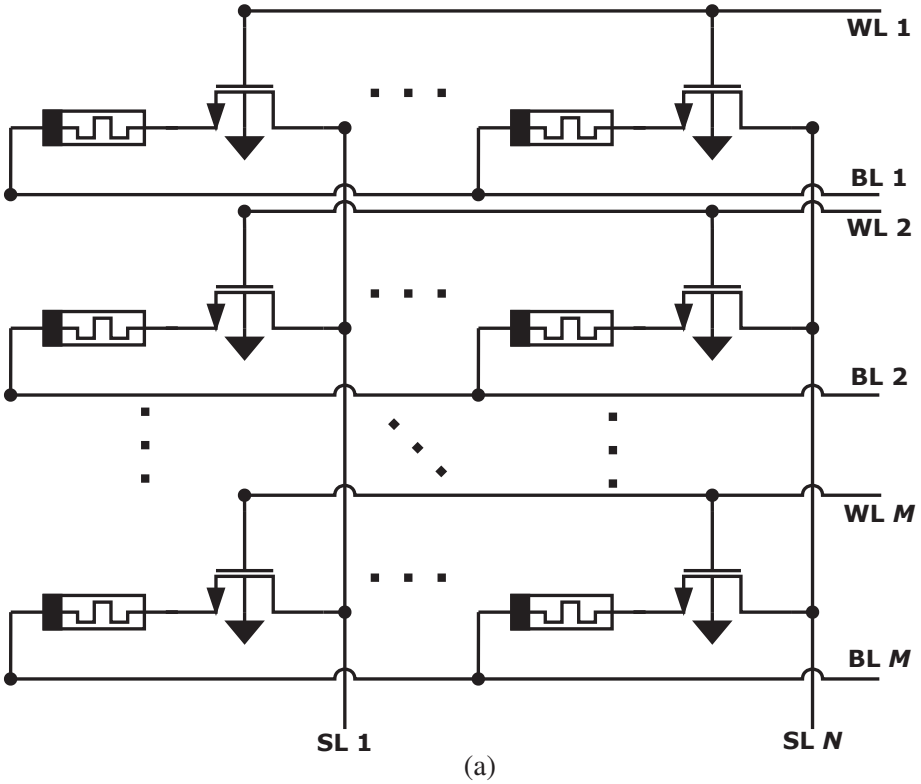
technology nodes smaller than $50\,\mathrm{nm}$, supply voltages typically do not exceed $1\,\mathrm{V}$ [125]. However, the electroforming process in VCM memristive devices often requires significantly higher voltages up to $4\,\mathrm{V}$ [24, 53]. In order to address this discrepancy, several techniques have been proposed such as the application of negative voltages at the transistor bulk in order to reduce the total applied voltages [124] or the implantation of silicon in a HfO_2 layer which substantially reduces the required forming voltages [32].

The co-integration is realized in a two-step back-end-of-line process. First, the MOSFET devices and circuitry are fabricated in a standard Complementary-Metal Oxide Semiconductor (CMOS) process. During this fabrication, several metal layers for interconnect routing are deposited. Between each metal layer an isolating *Inter Layer Dielectric* (ILD) is added as separation. Afterwards, the resistive devices are fabricated on top of the latest ILD layer. Via conductive plugs in this last layer, the resistive devices are connected to the underlying CMOS structures. A typical 1T-1R crossbar is arranged by three layers of interconnecting metal lines, each connecting one of the three bitcell terminals, namely *Word-Lines* (WL) which are connected to the transistor gates, *Signal-Lines* (SL) which are connected to the transistor drains, and *Bit-Lines* (BL) which are connected to one terminal of the memristive device. Four potential line arrangements are possible given this number of interconnecting metal lines. In the simplest arrangement all three interconnect lines are aligned in parallel. This structure forms individual columns of *line-arrays* rather than a matrix-like row-column structure. In a line-array it is not possible to select individual devices as the same signals are shared along one line. For an effective device isolation in a crossbar, one line has to be arranged perpendicularly. In case that the Word-Line is perpendicular, the structure resembles a standard *memory array*. This allows for accessing each bitcell individually as well as a multi-bit device programming. The two other crossbar arrangements are illustrated in Fig. 2.4. Throughout this thesis, the position of a device within a crossbar array can be pinpointed based on its location (i, j) in a matrix-like structure. Here, i refers to the vertical position as rows, with $1 \leq i \leq M$, and j to the horizontal position as column with $1 \leq j \leq N$, in a crossbar of size $M \times N$.

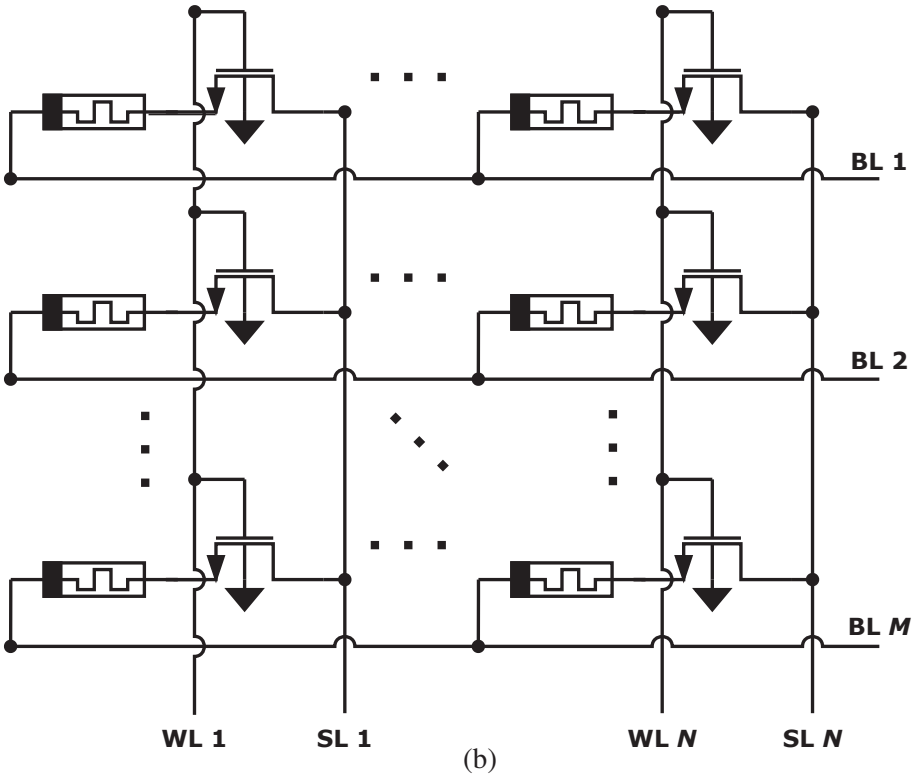
2.2.1.1 Peripheral Circuitry

The active 1T-1R crossbar is an analog component. Therefore it requires the integration of several mixed-signal components as periphery circuits to effectively control the crossbar and ensure a flawless interface with the broader digital computing architecture. *Digital-Analog Converters* (DAC) are essential for converting digital signal of the surrounding computing systems into corresponding analog voltages which can be applied at the interconnect lines of the crossbar. After a computation operation or read-out, *Analog-Digital Converters* (ADC) convert the resulting analog signals, either currents or voltages, back into digital data. This data can then be interpreted and utilized by the surrounding digital computing system. Lastly, *memory controllers* and *registers* are integrated to correctly address, store and retrieve the data during an operation.

Especially for the increasing downscaling of CMOS circuits, the impact of variability or circuit mismatch becomes more severe. In this context the term *mismatch* refers to the phenomenon that two or more transistor devices within the same circuit will behave slightly different even though they are designed to be identical [126]. This discrepancy between theoretical and observed CMOS behavior can arise due to imperfections in the fabrication process, e.g., in the deposited layer thicknesses or dopant concentrations, or variations in the underlying physical processes during the operation, e.g., variations in electric field or temperature [125]. The variations in the manufacturing process can be further distinguished based on the level of their occurrence, either between two individual wafers,



(a)



(b)

Figure 2.4: Possible arrangements of Signal-Lines (SL), Bit-Lines (BL), and Word-Lines (WL) in active 1T-1R crossbars of size $M \times N$, (a) 1T-1R crossbar and (b) Pseudo-crossbar. Devices are referred via their position (i, j) with $1 \leq i \leq M$ and $1 \leq j \leq N$.

across one wafer or within a particular die. The presence of mismatch in the CMOS circuitry can significantly affect the functionality and reliability of co-integrated CMOS/ReRAM crossbars as the mismatch variations might worsen the already occurring variation in the ReRAM devices. This leads to issues such as increased noise, a reduced yield, and overall a limited performance.

The primary concern of this thesis is understanding the reliable utilization of memristive devices in *Computing-in-Memory* applications. For this purpose, an ideal selecting and sensing periphery is assumed as their accuracy falls outside the scope of this thesis. This simplification allows for a more focused examination of the 1T-1R structures themselves. Solely the mismatch variation of the MOSFET transistors in the 1T-1R structures is taken into account due to its direct impact on the bitcell operation and performance while the peripheral components can be optimized separately.

2.2.2 Memristive Crossbars for Computation-in-Memory

Processing-in-Memory has been introduced as a potential solution to the challenges posed by the memory wall in the classic von-Neumann architecture. By breaking up the traditional strict separation of memory and processing unit, PIM aims to reduce the costly data transport [127, 128]. This can be achieved through physically positioning the memory in closest proximity to the processor as so called Near-Memory Computing [129], for instance in 3-dimensional stacking integration [130]. Even more radical is the approach to process the data directly within the memory itself, which fully eliminates the need for data movement [131, 132].

While PIM originally focused on the improvement of charge-based technologies, e.g., SRAM, there has been a growing interest in utilizing resistive memories such as resistive switching devices within the PIM context [97]. It has to be noted that several other terms are used besides Processing-in-Memory such as *Computing-in-Memory*, *In-Memory-Computing*, *Logic-in-Memory* or *Near-Memory-Processing*. To keep a clear terminology throughout the rest of this thesis, the term *memristive Computing-in-Memory* (memristive CIM) is used for the specific ReRAM-based applications beyond pure data storage while the term Processing-in-Memory is used for the more abstract spectrum which includes all technologies.

Memristive Computing-in-Memory can be categorized based on its targeted applications and their computing nature. The most promising areas are currently hardware acceleration for artificial neural networks (ANN) [4, 133], database structures such as in Content-Addressable Memory (CAM) circuits [134, 135] and logic circuits [136, 137]. In the realm of ANNs, memristive devices are used for as hardware accelerators for machine learning operations such as Vector-Matrix Multiplications and Matrix-Matrix Multiplication [133]. This is particularly beneficial for the execution of multiply-and-accumulate operations which are efficiently facilitated through Ohm's law (multiply) and Kirchhoff's current laws (accumulate) in memristive crossbar arrays [138, 139]. Recently, the unique properties of memristive devices have further been harnessed for their potential as artificial synapses to emulate brain functionalities, a field known as *neuromorphic computing*. While neuromorphic computing and hardware acceleration share some ground, neuromorphic computing primarily aims for applications beyond standard machine learning tasks. Memristive neuromorphic computing utilizes the dynamics of memristive devices to replicate the capabilities of biological neural networks [140, 141]. The goal here is not solely to accelerate the raw data processing but to create systems which can adapt and learn similar to a human brain. Normally these examples exploit the memristive device as an *analog* device for data processing.

In contrast to this, *logic Computing-in-Memory* executes Boolean operations using memristive devices as digital elements for various arithmetic and logic functionalities [136, 137]. Therefore it is also referred to as Logic-in-Memory. However, this terminology is unfortunately not sharply defined as

it spans from the implementation of single logic operations on memristive crossbars [142–144] to fully integrated processing units [145–148]. Given the fluid boundaries, the term *logic Computing-in-Memory* (logic CIM) is adopted here to describe the logic computation in memristive crossbars throughout this thesis.

The following section discusses specific details of logic CIM as main aspect of this thesis.

2.2.3 Classification of Memristive Logic CIM Concepts

The approach of executing logic operations directly within the memory is not exclusive to memristive devices as it is present in CMOS-based SRAM computing as well [131, 132, 149]. However, the characteristic features of memristive devices such as their non-volatility and low static power consumption have opened the door to novel computing paradigms and innovative circuit implementations that were not feasible in traditional CMOS technology [104, 150].

The scope of logic Computing-in-Memory ranges from the evaluation of individual Boolean logic gates [6–8, 151, 152] up to complex concatenated logic functions and advanced arithmetic building blocks such as full adders [10, 153] or arithmetic-logic units [154, 155]. The vast majority of these logic CIM concepts relies on the binary data classification of '0'/'1' as the principles of Boolean logic are tailored for this representation. However, the unique multi-level capabilities of memristive devices have sparked the interest for experimental approaches [156], for instance a ternary logic Łukasiewicz adder [157]. In order to effectively identify and compare the diverse field of memristive logic CIM concepts, it is recommended to classify applications and underlying primitive logic functions according to their distinguishing key characteristics.

In the context of memristive crossbar arrays, a broad classification is given by the location where a logic result is computed. This can either be within the array itself or in the array periphery. This leads to two subcategories: *CIM-Array* (CIM-A) and *CIM-Periphery* (CIM-P) [137, 150]. CIM-A concepts produce a logic result directly within the memristive crossbar and thereby integrate processing and memory in the same location. This approach aims to maximize the computation efficiency and data processing speed by reducing the effective memory wall. However, a specialized control over the crossbar is required for executing complex computations. This increases the additional circuit overhead for instance with dedicated memory controllers. Contrary to CIM-A, in CIM-P concepts the logic result is generated within the array periphery which consists of ADCs or de-*multiplexers*. While CIM-P exploits memristive arrays for data storage, the actual logic is executed only in the peripheral circuitry surrounding the array. The crossbar design of CIM-P arrays is therefore similar as in memristive arrays for pure data storage. This reduces the design effort and offers a higher degree of flexibility. However since the logic output is not computed directly in the array, additional steps are necessary to permanently store the result.

Logic CIM concepts can further be distinguished based on their input and output representation and the computing element [158]. The question regarding the input and output revolves around the physical data representation, i.e., its volatility over time. Both, the input and the output, can either be represented by a volatile quantity, e.g., a voltage, or a non-volatile quantity, e.g., a resistance. The terms *stateful* and *non-stateful* are introduced to describe the output data volatility in particular. An operation is stateful if the output is represented as a non-volatile entity. Prominent examples of such memristive logic concepts are *Implication Logic* (IMPLY) by Lehtonen et al. [6], *Memristor-Aided-Logic* (MAGIC) by Kvatinsky et al. [8] or the concept of *Complementary Resistive Switches* (CRS) by Linn et al. [159]. In contrast to this, non-stateful refers to operations where the output is represented by a volatile quantity. This category includes *Scouting Logic* by Xie et al. [151] or *Majority Logic* [160, 161]. The non-volatile

result of a stateful operation is persistent over time and in the absence of electrical power. Contrary, the information in a non-stateful operation is lost if the electrical power is turned off. It has to be noted that this distinction does not include potential data storing operations but only the form which is inherent to the operation. By introducing an additional storing step (write-back) any non-stateful operation result can be converted to a non-volatile quantity.

Another way to classify logic CIM concepts is based on the form of their primitive logic element which acts as repetitive functional block to the computing system. The complexity of this element significantly influences the functionality of the CIM circuit. In the simple case of a crossbar, the primitive logic element is just a single memristive device (1R), two memristive devices as complementary resistive switch, or the serial connection of a transistor and memristive device (1T-1R). These structures offer the advantage of performing versatile applications and they can be integrated in highly dense crossbar arrangements. Furthermore, the fabrication process of these CIM blocks is similar to a standard ReRAM memory array and differs only in the amount of additional CMOS circuitry. Besides these simple examples, there are complex logic design such as the *memristors-as-drivers* (MAD) gates [152], or memristive Content-Addressable Memory (CAM) bitcells [162]. While these elaborated structures offer a reduced and dedicated functionality due to their complexity they allow for a high degree of optimization towards their designated application.

The selection between simple and complex primitive structures in Computing-In-Memory systems is often a trade-off between versatility and specialization. Simple designs such as 1R or 1T-1R are highly flexible and easy to fabricate but may not be optimized for specific tasks. In contrast, more complex elements can be tailored for performing particular functions efficiently but lack the broad applicability and require labor-intense design efforts.

2.2.4 Requirements for the Investigated Memristive Logic CIM Operations

To select appropriate CIM primitives throughout this study, several criteria were established: Firstly, the *compatibility with 1T-1R memristive crossbars*. All investigated concepts must be compatible with the repetitive structure of 1T-1R memristive bitcells as this ensures a high degree of versatility. This aligns with the primary aim of this thesis to explore the general applicability and reliability of logic CIM rather than focusing on a particular specialized concept. Secondly, the selected concepts must include a *state conservation*. This means that the input data is preserved during an operation which allows for its subsequent reuse in further computations. The state conservation criterion does not necessarily binds to the input volatility. For instance, the IMPLY concepts use non-volatile resistances as logic inputs [6]. However it is not state conservative. During the IMPLY operation, the output overwrites one of the input devices, resulting in a partial loss of the input information. In order to reuse this input data, an additional copy step would be required prior to the indented operation. This increases both, the latency and energy consumption, and therefore such concepts are excluded in this study.

The primary distinction among the explored logic CIM concepts lies in the *volatility of the operation output*. This divides the concepts in the two categories of *readout-based* and *switching-based* operations. As the name states, stateful switching-based operations involve the switching of a memristive device which serves as non-volatile output storage. In contrast to this, non-stateful readout-based concepts rely solely on reading out a memristive device state which results in a volatile output current.

2.3 Statistical Tools

A major part of this dissertation covers the investigation of VCM device variability and its effects on logic Computing-in-Memory evaluations. As the underlying phenomena are inherent stochastic processes, the investigation requires suitable statistical tools and methods which are detailed in the following section.

2.3.1 Normal Distribution

The Gaussian or *normal distribution* as continuous probability holds enormous significance in characterizing and analyzing statistical processes of real-valued random variables. In the field of non-volatile memories normal distributions are utilized to describe variability effects in oxide-based ReRAM devices. Several experimental studies reported different statistical behaviors of the high-resistive or low-resistive state distributions as normal- [52, 54, 163], log-normal- [30, 164, 165] or Erlang- [166] distributed. Therefore it is debatable whether all underlying variability phenomena can be accurately described by normal-distributed mechanisms. Nevertheless various studies indicate a normal-distributed behavior [52, 54, 163, 167] at least for the low-resistive states [168, 169]. Furthermore, the well-known and straightforward properties of the normal distribution make it suitable for a resource-efficient modeling of ReRAM devices in larger circuits and applications [84, 170–172].

The *probability density function* (PDF) of the normal distribution is mathematically expressed by the Gaussian function

$$f_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \tag{2.1}$$

This equation describes a single peak with infinitely extended, symmetric tails on both sides. The main parameters for describing the PDF are μ, the center or mean value of the distribution (location of the peak) and its variance, σ^2, where $\mu \pm \sigma$ are the function's inflection points. As the PDF curve represents a normalized distribution, the total area under the curve is 1. It is crucial to consider this normalization in the context of fitting a normal distribution to experimental data since arbitrary data might not be inherently normalized. To account for this, a scaling parameter is introduced which leads to the form

$$\begin{aligned} \tilde{f}_{\mu,\sigma}(x) &= \frac{1}{\displaystyle\int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \, \mathrm{d}x \right)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \\ &\equiv a \times e^{-(x-m)^2/2s} = f_{a,m,s}(x) \end{aligned} \tag{2.2}$$

Visually, the scaling parameter represents the peak's height. Throughout this thesis the term *generalized normal distribution* is used for the expression in Eq. (2.2) to distinguish it from the Gaussian function in Eq. (2.1). An important feature of the PDF is its relation to the *cumulative distribution function* (CDF). The PDF integral over a closed interval from x_a to x_b yields the probability that a randomly selected sample from the distribution exhibits a value between $[x_a, x_b]$. This relation is generalized by the CDF

which is expressed as

$$F_{\mu,\sigma,x_a,x_b}(x) = \int_{x_a}^{b} PDF(x) = \int_{x_a}^{x_b} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \, \mathrm{d}x$$
$$= \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right] \tag{2.3}$$

with erf being the Gaussian error function.

2.3.2 Log-Normal Distribution

The *logarithmic-normal* or log-normal distribution is genuinely linked to the normal distribution. The PDF of a log-normal distribution can be derived from any given normal distribution $f_{\mu,\sigma}(x)$ by applying the natural logarithmic function $ln(x)$ as described by

$$f_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(ln(x)-\mu)^2/2\sigma^2} \tag{2.4}$$

As for the normal distribution, this expression can be generalized by including a scale parameter. This results in

$$\tilde{f}_{\mu,\sigma}(x) = \frac{1}{\int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(ln(x)-\mu)^2/2\sigma^2} \, \mathrm{d}x\right)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(ln(x)-\mu)^2/2\sigma^2}$$
$$\equiv a \times e^{-(ln(x)-m)^2/2s} = f_{a,m,s}(x) \tag{2.5}$$

Unlike the symmetric normal distribution, the log-normal PDF is right-skewed with a skewness of $(e^{\sigma^2}+2)\sqrt{e^{\sigma^2}-1}$. Graphically, this skewness manifests as an extended tail on the right side which indicates a higher probability of observing values greater than the distribution's mean.

Despite this difference, the generalized log-normal CDF can be calculated from its PDF as derived in Eq. (2.3) with respect to the included natural logarithm.

2.3.3 Overlap between Probability Density Functions

A useful measure for the similarity between two probability distributions, either log-normal or normal, is the overlap coefficient. This metric quantifies the intersecting area ratio of both distributions between 0 (no overlap) and 1 (identical distributions) [173]. In order to calculate the shared area of both distributions, the semi-infinite integral of each PDF is calculated with respect to their point of intersection. In the realm of memristive Computing-in-Memory, the overlapping areas are useful to describe the separability error of current distributions assigned to different logic states [174], which will be explained in detail throughout the following chapter.

For all three distribution combinations, log-normal to log-normal, normal to normal and log-normal to normal, there exists a maximum of two real intersection points under the condition that σ of both distributions is not equal, $\sigma_1 \neq \sigma_2$, and all parameters are real number values. The corresponding

equation for a normal to normal intersection is exemplary derived in Appendix A.1. Although analytical formulas for determining the intersection points exist, computing them numerically is often more convenient. These algorithms are user-friendly implemented and available in various higher-level programming languages including Python through the SciPy library [175].

2.3.4 Fitting of Cumulative Distribution Functions

The cumulative distribution functions of both normal and log-normal distributions are closely related in their mathematical expression. Both include the Gaussian error function erf. While this expression is analytically correct, the numerical calculation of erf is computing intensive. The Gaussian erf-function is itself only a special representation of a sigmoid function. This class of functions is categorized as continuous, strictly increasing and bounded, $f : [-\infty; \infty] \mapsto (0, 1)$. For a computing-friendly calculation, the erf terms in the CDF can be replaced by related sigmoid functions without changing the distribution shape.

A standard expression of a sigmoid function is the logistic function, $f(x) = \frac{1}{1+e^{-x}}$. Yet especially for probability calculations, it is important to match not only the distribution shape of the CDF, but furthermore to ensure a good convergence at the distribution tails, $\lim_{x\to-\infty} = 0$ and $\lim_{x\to\infty} = 1$. Among the several sigmoid functions, the Weibull function is selected for this purpose due to its boundary behavior and numerical simplicity [176, 177].

2.4 Simulation and Measurement Tools

2.4.1 Simulation Tools

The comprehensive fabrication and measurement of all discussed circuits and operations is difficult and labor-intensive. Moreover, in order to observe certain physical properties substantial measurement resources are required. To address this challenge, device and circuit simulations are exploited throughout this thesis in order to emulate and forecast the behavior of Computing-in-Memory operations. In memristive simulations, *compact models* play an important role as they provide a simplified but effective representation of the memristive switching behavior and can be integrated into larger electronic circuit simulators. Different model implementations for the resistive switching mechanism exist such as a tunnel-gap modulation [171, 178, 179] or Schottky-barrier tuning [180].

2.4.1.1 Juelich-Aachen Resistive Switching Tools

The Juelich-Aachen Resistive Switching Tools (JART) consist of various compact models for simulating different kinds of oxide-based ReRAM devices. The JART models are developed by the joined simulation group at the Peter-Grünberg Institute PGI-7, Forschungszentrum Juelich and the Institut für Werkstoffe der Elektrotechnik II, RWTH Aachen [181]. One of the models in this class is the JART VCM-v1b compact model. Originally derived based on a previous work from Siemon et al. [180], this model can simulate the electrical behavior and dynamics of filamentary-switching VCM-ReRAM devices by modeling the oxygen vacancy concentration inside the material stack.

The equivalent circuit diagram of the model is depicted in Fig. 2.5. It represents a VCM material stack of active electrode/insulating metal oxide/ohmic electrode as the serial connection of three ohmic contributions, R_{series}, R_{plug} and R_{disc}, and a Schottky diode. In this model, R_{series} combines the electrical

resistance of the ohmic electrode and the line resistances of all interconnects. The electrical resistance of the insulating metal oxide layer after the electroforming process is described as two contributions R_{plug} and R_{disc}. In this picture, the conductive filament is divided into a constant high-conductive *plug* region and a smaller *disc* region. The oxygen vacancy concentration in the *disc* region is variable in order to model the movement of vacancies towards and away from the metal electrode/insulating oxide interface. This concentration change in the model correlates to the resistance modulation of a VCM device under switching.

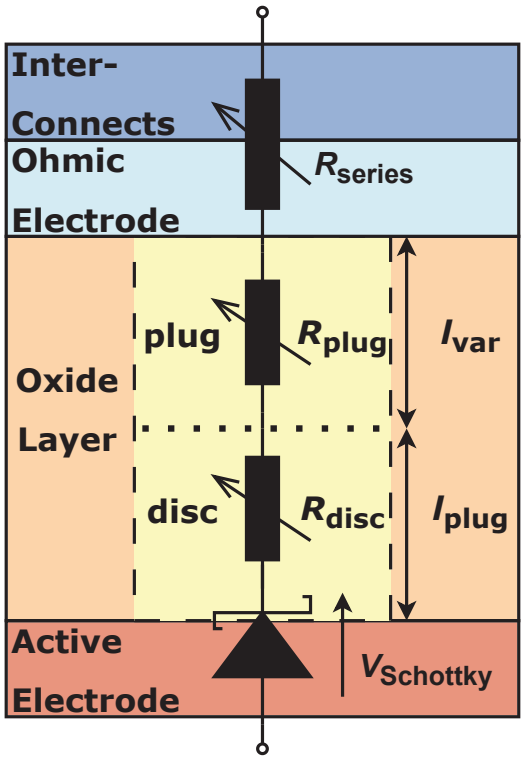


Figure 2.5: The equivalent circuit diagram for the JART VCM v1b model class. The model consists of constant ohmic resistances R_{Series}, a Schottky diode and two variable ohmic resistances, R_{plug} and R_{disc}. This figure is adapted from Bengel et al. [170], used under CC BY 4.0 [182].

The resistances R_{plug} and R_{disc} are calculated based on the filament geometry and the oxygen vacancies as

$$R_{\text{plug/disc}} = \frac{l_{\text{plug/disc}}}{z_{v_{\text{o}}} \times e A \mu_{\text{n}} \times N_{\text{plug/disc}}} \tag{2.6}$$

with e being the elementary charge, $z_{v_{\text{o}}}$ the oxygen vacancy charge number, $A = \pi r^2_{\text{plug/disc}}$ the area of the conductive filament, μ_{n} the electron mobility in the oxide layer and $l_{\text{plug/disc}}$ the length of the plug/disc region. While the oxygen vacancy concentration in the plug region, N_{plug}, is selected as constant high value, the concentration in the disc region, N_{disc}, is exploited as voltage-dependent state variable. The concentration change of N_{disc} is modeled via the ionic transport according to the Genreith-Schriever and de Souza conduction mechanism [183].

The Schottky diode in the model represents an electrostatic barrier at the interface between the active electrode and the insulating metal oxide. The current through this barrier depends both on the applied voltage amplitude and polarity as well as N_{disc}. A detailed explanation of all equations is provided in

Table 2.1: Used simulation parameters for the JART VCM RERAM v1b model.

Parameters	Value
ReRAM Model	JART VCM v1b var
Material Stack	$Pt/HfO_x/TiO_x/Ti$ [184]
Mean filament radius $r_{\text{disc/plug,mean}}$	$45\,\text{nm}$
Mean filament disc length $l_{\text{disc,mean}}$	$0.6\,\text{nm}$
Mean oxygen vacancy concentration HRS $N_{\text{HRS,mean}}$	$8 \times 10^{23}\,\text{m}^{-3}$
Mean oxygen vacancy concentration LRS $N_{\text{LRS,mean}}$	$10 \times 10^{26}\,\text{m}^{-3}$

[181, 184].

The physical model parameters were fitted to experimental measurement data for material stacks with different insulating metal oxides, namely HfO_x/TiO_x, ZrO_x/Ta and TaO_x/W. While the first JART VCM v1b version was originally developed as a deterministic model by Cüppers et al. [28], it was improved over the years to account for non-idealities such as switching variability (C2C and D2D) [170] and read instability [81].

The important parameters for implementing this variability are the minimum and maximum oxygen vacancy concentration in the *disc* region, $N_{\text{disc,min}}$ and $N_{\text{disc,max}}$ where $N_{\text{disc,min}} \leq N_{\text{disc}} \leq N_{\text{disc,max}}$, and the filament geometry in the *disc* region, described by its length l_{disc} and radius r_{disc}. In order to incorporate device-to-device variability within the model, any simulated device receives a randomly generated set of these four parameters ($N_{\text{disc,min,var}}$, $N_{\text{disc,max,var}}$, $l_{\text{disc,var}}$ and $r_{\text{disc,var}}$). This parameter set is sampled from independent Gaussian distributions. To prevent the selection of physically unrealistic parameters, e.g., a negative filament radius, the Gaussian distributions are reasonably truncated. After the variability-aware initialization, with each switching cycle a new set of four parameters is sampled from the underlying distributions. This incorporates a cycle-to-cycle variability to each device. Important for this implementation are the means and standard deviations, μ and σ, for each parameters distribution. These values are adjusted in the model to match the experimentally observed variability scales.

The read instability is implemented by a random jump process based on kinetic Monte-Carlo simulations [81]. In the model, the oxygen vacancy concentration $N_{\text{disc,var}}$ at any time can change by a small value, $\pm\Delta N_{\text{disc,jump}}$. Here, the jump direction is determined by the parameter m which is drawn randomly and the jump frequency is again matched to the experimentally observed read instability.

The electrical circuit simulations in this thesis are conducted in the Virtuoso ADE (Analog Design Environment) Suite from Cadence Design Systems. For the use in this software, the JART VCM-v1b model is implemented in the programming language Verilog-A. Verilog-A is able to generate all necessary random parameter internally through the provided statistical functions and distributions. The transistor devices are accessed by a *Process Design Kit* (PDK) provided by the foundry.

Table 2.2: Transistor simulation parameters for the XFAB 180 nm Process-Design-Kit. [1] The W/L parameter is matched to the available dimensions in the experimentally measured MAD-200 design.

Parameters	Value
Transistor PDK	XFAB XP018 180 nm process
VDD	5 V
Width to Length [1] W/L	4

2.4.2 Measurement Methods

In addition to the discussed employment of circuit simulations, a major facet of this thesis involves measurements on fabricated ReRAM devices. The measurement data plays a crucial role in validating or contradicting the predictions made from simulations regarding the operational parameters or testing applications. In particular this is important for logic Computing-in-Memory concepts as they currently exist at a proof-of-concept stage with no or only a limited experimental demonstration. However, such a rigorous verification is essential for the transitions of logic CIM towards industry-relevant applications. Two different measurement setup and device technologies were exploited to collect all the measurement data in this thesis. Both setups use memristive crossbar structures where filamentary ReRAM devices were co-integrated in commercially available CMOS technologies. The most prominent differences between the implementations are the CMOS technology node and the ReRAM material stack properties.

2.4.2.1 MAD-200 testchip in 130 nm

The majority of experimental measurements shown in this thesis were performed on an integrated ReRAM chip out of the *Memory Advanced Demonstrator* 200 mm (MAD-200) process offered by the Circuits Multi-Projects (CMP) manufacturer [185]. The used test chip was designed by Tobias Ziegler from RWTH Aachen and Dr.-Ing. Sebastian Siegel from Forschungszentrum Juelich. The access transistors and digital select logic were implemented in the HCMOS9A 130 nm CMOS process from STMicroelectronics while the memristive devices were deposited in a back-end-of-line (BEOL) process by CEA-LETI [169] with TiN/HfO_2/Ti as material stack. In the custom layout, each ReRAM is connected in series with a common-source n-channel MOSFET as 1T-1R configuration. Each 1T-1R device is connected to three units for selecting the Word-Line, Signal-Line and Bit-Line. This selection layout is depicted in Fig. 2.6 (a). Two different 1T-1R array structures were implemented on the test chip, a 1T-1R array and a pseudo-crossbar array. The interface between the MAD-200 test chip and the measurement setup is realized by a custom design printed circuit board (PCB). This PCB connects to a Labjack T7 data acquisition board for controlling the digital on-chip circuitry and an external V_{DD} voltage supply. The analog signals are generated by a μC module from aixACCT Systems. This module provides a 64-channel switch matrix for applying signals and simultaneously monitoring the current measurement. An additional PicoScope 5000a PC oscilloscope is exploited for high-resolution voltage measurement at a sampling rate up to 1 GHz.

2.4.2.2 ELMOS testchip in 350 nm

A second test chip for experimental measurements was developed within the NeuroSys cluster. The CMOS layout was designed in a 350 nm process provided by Elmos Semiconductor SE. On top of this,

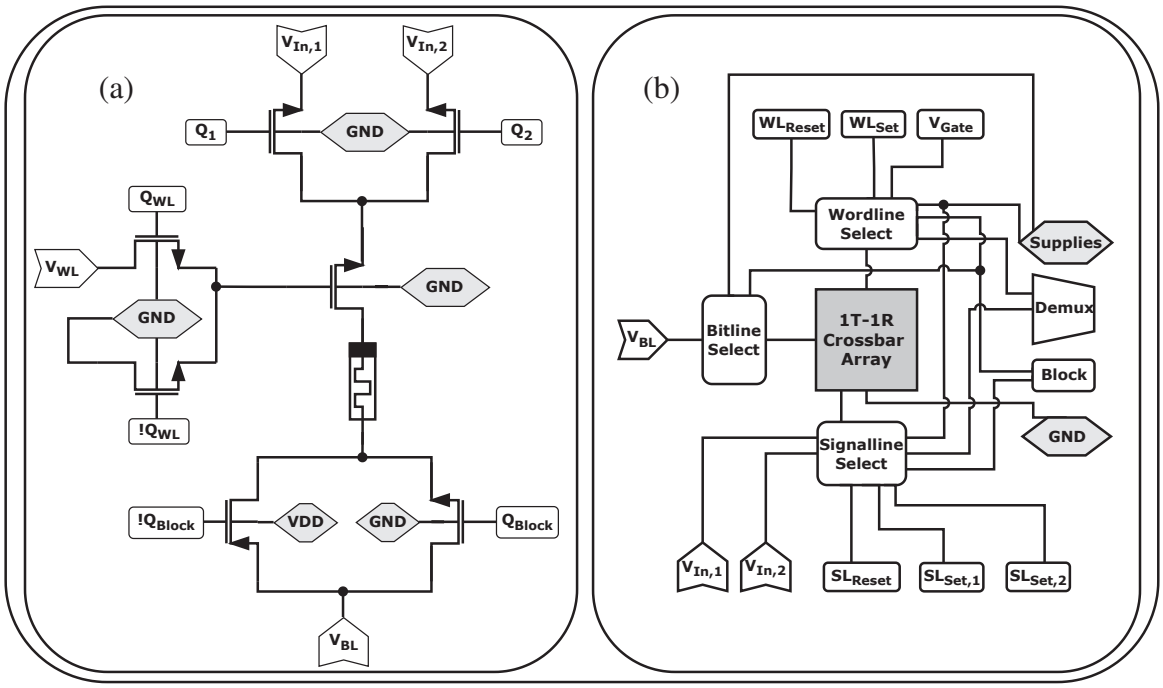


Figure 2.6: (a) Schematic representation of an individual 1T-1R cell and the signal application paths in the MAD-200 design. Each 1T-1R is connected to 3 paths for applying Word-Line, Signal-Line and Bit-Line signals (b) Interfacing of the MAD-200 test chip with the digital selection and peripheral circuitry.

the memristive devices were fabricated using a back-end-of-line process by Dr.-Ing. Felix Cüppers and Danielle Storelli at Peter-Grünberg-Institut 7, Forschungszentrum Juelich. Similar to the architecture of the MAD-200 setup, the Elmos test chip features multiple array structures consisting of co-integrated ReRAM and n-channel MOSFET devices. The unique characteristic of the Elmos test chip is that multiple of the arrays adopt a 2T-1R configuration where two transistors are linked in series to the same ReRAM device. The source terminals of both transistors are connected to a common Signal-Line, while their drains connect to the active electrode of the ReRAM. Each transistor gate however is independently accessible via its own Word-Line. This increases the number of necessary inputs from three (SL, BL, and WL) to four (SL, BL, WL A, and WL B). The schematic view of a 2T-1R cell is depicted in Fig. 2.7 (a). During operations such as SET, RESET or Read-out, only one of the transistors is activated by applying a gate-source voltage, $V_{GS} > 0$, at its Word-Line while the other one is set to ground potential. This configuration effectively simplifies to a 1T-1R setup for each individual operation. The significant benefit of utilizing a 2T-1R approach lies in the ability to design the transistors with individual dimensions and varying their width-to-length (W/L) ratio. W/L is a main parameter which determines the MOSFET channel conductance and drain-source current, I_{DS} [186, PP.143-150]. Recent studies indicates that different W/L ratios can enhance the variability control of 1T-1R operations [187]. While a smaller ratio leads to a more precise control of the SET process, higher W/L ratios are beneficial for reducing the transistor's channel resistance during a Read-out. As consequence, a 2T-1R approach can combine the optimal features of both operations. This offers a superior switching control without the need for a trade-off between the two. However, the major disadvantage of this architecture is the substantial increase in the bitcell footprint due to the inclusion of two transistors.

The Elmos test chip layout, illustrated in Fig. 2.7 (b), incorporates two 2T-1R arrays, one measuring 15

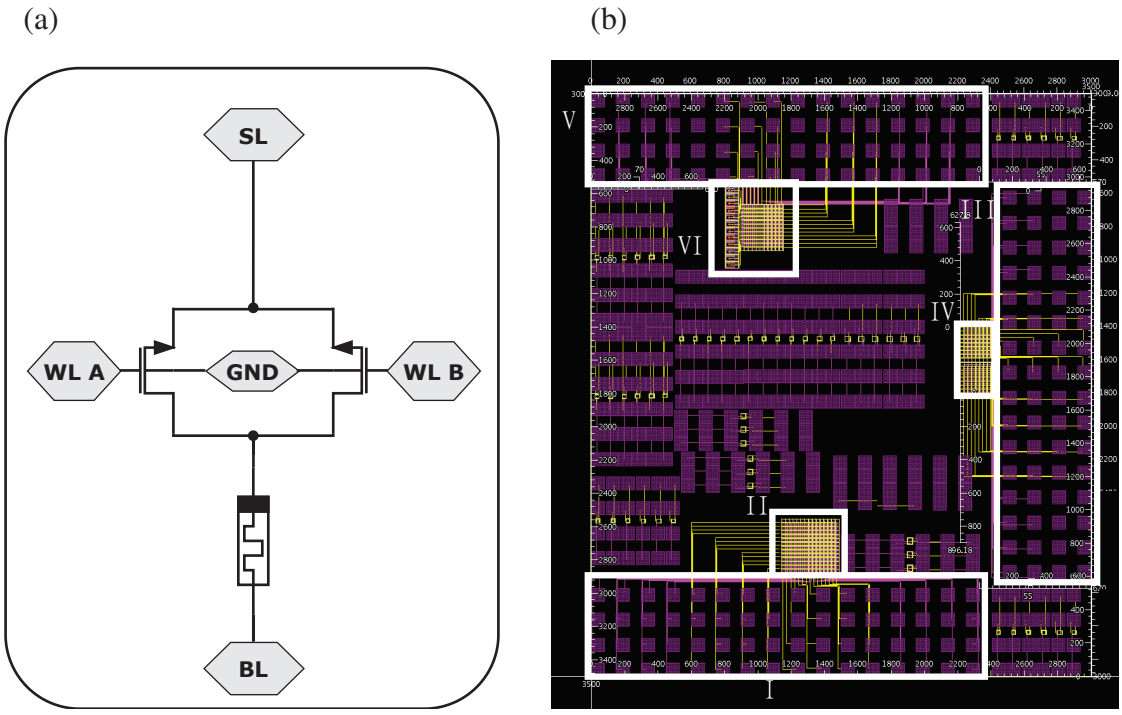


Figure 2.7: (a) 2T-1R bitcell layout utilized in the Elmos test chip. In general, both n-channel MOSFET devices have different width-to-length ratios, W/L. (b) Elmos test chip layout. Highlighted are the three array structures with corresponding probe-pads: 15 x 15 2T-1R (I, II), two 8 x 8 1T-1R (III, IV) and 12 x 12 2T-1R with ohmic resistances (V, VI).

x 15 and the other 12 x 12 (II and VI in Fig. 2.7 (b)). Both arrays are configured in the pseudo-crossbar arrangement. A distinctive trace of the 12 x 12 array is that one of its columns (1 x 12) is fabricated without memristive devices in the back-end-of-line (BEOL) process. Instead these cells are equipped with $100\,\mathrm{k}\Omega^{-1}$ ohmic resistances during the CMOS fabrication stage. These resistances serve as ideal inputs for logic operations which enables a comparison to the memristive devices. Furthermore, the test chip includes two standard 1T-1R arrays (IV in Fig. 2.7 (b)), also arranged in a pseudo-crossbar layout. The three arrays are complemented by several individual test structures designed to evaluate the performance and characteristics of 1T, 1R, 1T-1R, and 2T-1R devices separately.

The test structures on the Elmos test chip are interfaced via dedicated probe pads with the commercially available semi-automatic wafer prober aixMATRIX from aixACCT systems [188]. The system is equipped with 64 independent DAC channels which deliver 64 parallel voltage wave-forms. Furthermore, 32 of these channels are supplied with additional current amplifiers which enable the recording of up to 32 separate current signals at a sampling rate of 100 MHz. Due to the fixed number of 64 measurable channels, each array structure on the test chip is accompanied by 64 probe pads to position the probing needles. These pad structures are highlighted in Fig. 2.7 (b) as numbers I (15 x 15 array), III (two 8 x 8 arrays) and V (12 x 12 array).

Two different material stacks are co-integrated on the Elmos test chip. The first stack consists of Pt/Ta_2O_5/Ta/Pt with metallic platinum electrodes and Ta_2O_5 as insulating oxide layer. The second stack has a TiO_2/ HfO_2 bi-layer as insulating oxide with platinum electrodes, Pt/Ti/TiO_2/HfO_2/Pt.

The high level of customization in the Elmos test chip and the measurement setup provides an enhanced flexibility compared to the MAD-200 process. For instance different metal-oxide can be explored as

resistive switching material. However, this increased flexibility requires time-intensive preparation and verification processes to adjust all necessary fabrication and measurement routines adequately. Therefore, the Elmos setup is primarily used for demonstrating proof-of-concept functionalities which are not implementable in the MAD-200 design. Due to its stability, the MAD-200 setup is utilized for the acquisition of large-scale measurement data for statistical evaluation.

CHAPTER 3

Readout-based Logic Operations

3.1 Fundamentals of Readout-based Logic Operation

The fundamental principle of *readout-based* logic concepts in memristive Computing-in-Memory systems is sensing the state-dependent electrical resistance of one or multiple memristive devices. This can be implemented in a current-based sensing operation. Here, a voltage is applied to a memristive device which induces a current flow. The measured current represents the logic information stored in the device. This whole readout process has to be performed without altering the device's state. Ensuring this *read disturb immunity* is achieved by carefully calibrating the readout voltage pulse. Generally, shorter pulse widths necessitate higher amplitudes for state switching. This is known as the *voltage-time dilemma*. For VCM-ReRAM devices it has been observed that voltage amplitudes below $\approx 0.4\,\mathrm{V}$ do not influence the device state even for extended pulse widths [28, 34, 189]. However this is only valid if the voltage is applied in the RESET direction. For the SET polarity this can already cause a read disturb.

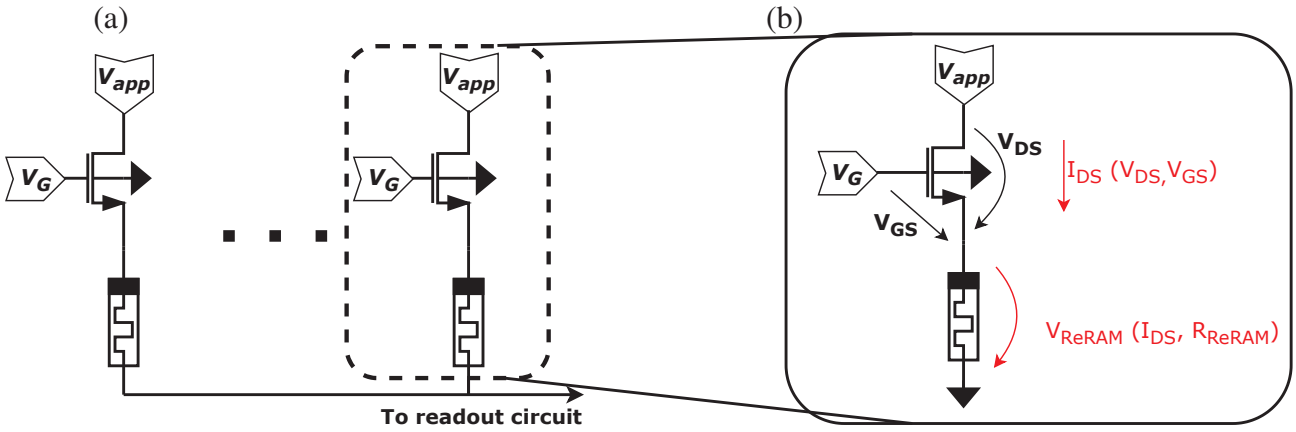


Figure 3.1: Fundamental readout operation in a 1T-1R bitcell. a) Circuit arrangement for the parallel readout of several 1T-1R bitcells. b) Visualization of occurring voltages and currents in a 1T-1R readout.

An important aspect in active 1T-1R structures is the inseparability of the transistor and the memristive device. Therefore, the described sensing process senses the current flowing through the serial combination of the transistor and ReRAM. As previously discussed, the observed electrical

resistance of a VCM-ReRAM, R_{ReRAM}, is the quotient of the voltage across the VCM-ReRAM device, V_{ReRAM}, and the resulting current, I_{ReRAM}, $R_{\mathrm{ReRAM}} = V_{\mathrm{ReRAM}}/I_{\mathrm{ReRAM}}$. This ReRAM resistance is majorly determined by its oxygen vacancy configuration. On the other hand, the drain-source current of a transistor, I_{DS}, is a function of the applied gate-source voltage, V_{GS}, and the drain-source voltage, V_{DS}.

The 1T-1R serial arrangement for a readout operation is illustrated in Fig. 3.1. In this voltage divider configuration the individual voltage drops across both the transistor (V_{DS}) and the V_{ReRAM} will adjust according to the external conditions, namely the applied voltage, V_{app}, the applied gate voltage, V_{G} and the resistive state of the ReRAM device. This interaction will then determine the total current through the series arrangement, I_{DS}. To minimize the transistor's impact on the total measured current and consequently improve the precision in detecting the ReRAM state, it is favorable to operate the MOSFET with a maximum gate-source voltage by applying a high V_{G}. At high V_{GS} the n-channel MOSFET enters its *saturation region*, which is characterized by a low drain-source channel resistance. The total 1T-1R current can be detected and quantified by dedicated analog peripheral components. In the simplest form, the sensing is provided through an *current-mode analog-to-digital converter* (current-mode ADC) which digitizes the analog current signal. Alternatively, a combination of a differential amplifier and a shunt resistor can be used by measuring the voltage drop across the shunt resistor. Both measurement options, current or voltage, effectively translate the analog resistance state of the VCM-ReRAM device into a processable digital information.

The inherent ReRAM variability poses substantial challenges to the reliability of readout-based logic operations. A variation in the resistive states, R_{ReRAM}, will lead to fluctuations in the displayed voltage divider arrangement of V_{ReRAM} and V_{DS} which result in heterogeneous output current. Due to this inconsistency, two 1T-1R cells might produce different readout current even though they represent the same digital value. This scenario is further complicated by the variability that originates from the bitcell transistor as mismatch and process variations in the CMOS fabrication. Therefore, two transistor devices can show a different channel conductivity and I_{DS} even for identical V_{DS} and V_{GS}. A strategic approach to counter these variations is the usage of write-verify algorithms [86]. Through iterative programming and verification, the accuracy of a ReRAM state can be significantly improved. These algorithms can also play a crucial role in compensating the CMOS variations. As the verification step aims to equalize the total current through the 1T-1R arrangement, write-verify algorithms can effectively balance mismatches between the bitcell transistor and the ReRAM. While these techniques can substantially reduce variability in ReRAM devices, they cannot completely eliminate it. One persistent issue is read noise which causes fluctuations in the current within a single device even during the same read operation. To some extend the ADC circuits address this issue as they are designed to map a range of analog values to one digital output. However, the ability of ADCs fails if the current distributions for different digital values begin to overlap. In such a scenario although the ADC may function correctly, it is working with corrupted input data. At this point there is no mechanism within the ADC or beyond to correct the corruption. This issue becomes worse in multi-level operations. As the analog current ranges for different digital values shrink with the increasing number of levels, the likelihood of overlap between the distributions increases. This complicates the ADC's task of accurately mapping the analog input to the correct digital output.

3.2 Primitive Gates for Readout-based Logic

In the following section, first the developed framework for a failure rate estimation is introduced. Afterwards several readout-based logic gates are evaluated with increasing complexity.

3.2.1 Error Estimation Framework for Readout-Based Logic

In the context of memristive readout-based logic a *failure* or *logic error*, F, is defined by a discrepancy between the theoretical output value, S, and the state determined upon the measurement, Q. The theoretical value can be calculated according to the initial operation inputs and the conducted logic function. Based on this error definition, the *error probability* is calculated as

$$\mathbf{P}(F) = \sum_{i=0}^{m} \mathbf{P}(Q_i|S_i) \times \mathbf{P}(S_i); \qquad \text{if } (Q_i \neq S_i) \tag{3.1}$$

with m being the total count of resistive binary inputs which result in $m + 1$ possible unique output combinations [190, 191]. Even though the evaluation result Q is represented as digital value due to the ADC conversion, it is derived from an analog current, I. Therefore, noise fluctuations in the current will not always lead to a logic error. As long as the regions between the analog current ranges which are assigned to digital values $[I_{\mathrm{i,min}}; I_{\mathrm{i,max}}] \rightarrow Q_i$ do not intersect with each other, the state classification of Q remains correct. Describing this analog behavior and evaluating the failure rate presents several challenges for memristive logic gates. Sophisticated methodologies for error modeling or metrics like the architectural vulnerability factor (AVF) [192, 193] that are prevalent in CMOS logic, cannot be directly applied in memristive logic and require adaptation [194–196]. The impact of ReRAM variability on memristive logic computing is widely recognized and has been explored in the past [174, 197]. A common approach is to model the memristive device resistance either as normal or log-normal distributed quantity. This approach has been exploited for Monte-Carlo simulations of bulk wise, 1T-1R readout-based logic gates such as XOR, AND or OR [151, 198–202]. An alternative strategy is to measure integrated 1T-1R structures which involves the fabrication of in-house co-integrated circuits or the purchase of commercially available designs [203–205]. The major drawback of both approaches is the constraint placed by the limited number of data points, either simulation runs or measurement samples, which restricts the failure rate estimation. For STT-RAM devices an alternative theoretical method was proposed which addresses read decision failures through the calculation of voltage-dependent Gaussian current distribution [191]. This technique of calculating the single bit error probability based on readout current distributions has been introduced for the reliability of a single 1T-1R resistive readout operation [174]. Nevertheless, none of the suggested methods is able to fully capture analog failures in memristive readout-based logic.

The statistical framework for readout-based logic operations developed in this thesis has been previously published in [190, 206]. Parts of the following description have been adopted from these publications. The modified framework combines the benefits of various strategies while expanding them towards statistical modeling. It involves four main steps:

1. Merging measurement data with variability-aware simulations to model the gate-specific output currents as statistical distributions.

2. Determining the distribution overlap to obtain gate-specific error probabilities.

3. Including state-separating reference values and non-symmetry error probabilities.
4. Extending the framework beyond two-input logic gates and incorporating failure effects at crossbar level.

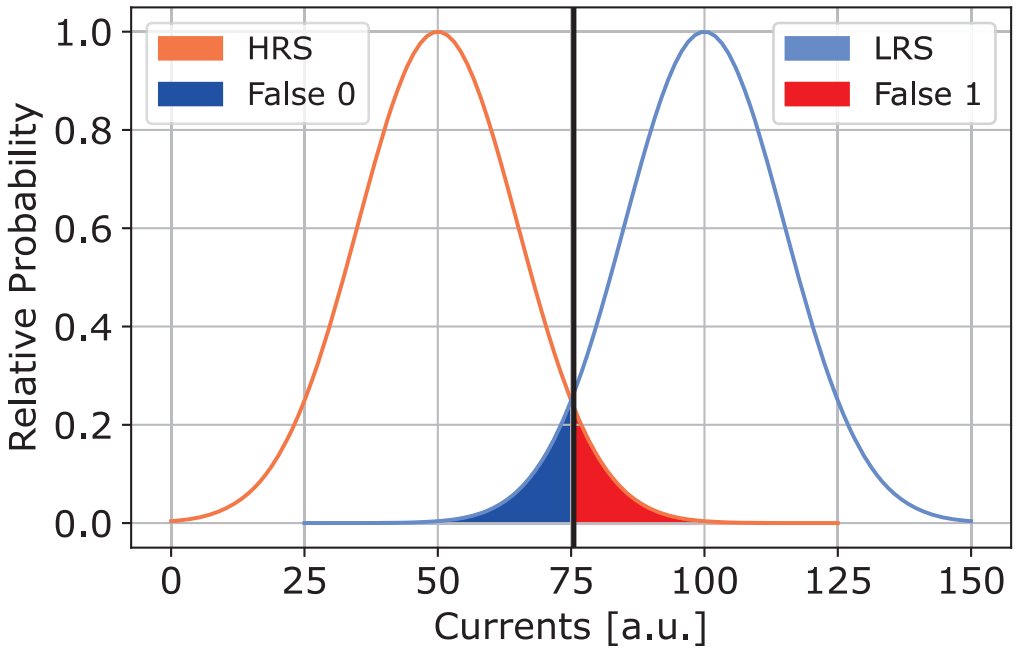


Figure 3.2: Idealized distributions of single device readout currents in the high resistive and low resistive state. The black vertical line marks the intersection threshold. All values above this value are classified as logic 1 while all lower values are sensed as a logic 0.

Particularly the third aspect is of vast importance. As the threshold values, e.g., quantization stages of an ADC, are adjustable, their selection holds a significant degree of freedom. If specific applications favor either false-positive or false-negative errors a proper threshold selection can assist in minimizing the overall error rate. This approach of an adaptive reference for improving the reliability of a sense-amplifier readout has been suggested by Park et al. [207].

The impact of threshold adjustment can be illustrated for a single device readout. In an ideal scenario the output current distributions of HRS and LRS do not overlap. Therefore, the threshold value can be set precisely at the midpoint between both mean values which guarantees a perfect distinction. Yet due to variability effects in real-life measurements there will always be a chance that devices will produce a lower or higher current. This will result in a mis-classification. However if the threshold is increased, it is unlikely that an HRS would be incorrectly identified as LRS (reducing false positive) anymore. On the other hand, this elevates the likelihood that a low LRS might not be detected and is classified as HRS (increasing false negatives). This principle is depicted in Fig. 3.2.

3.2.2 Single Device Readout

A single device sensing and binary state assignment presents the most basic resistive readout application. In the context of performing readout logic, this approach implements two distinct Boolean functions. On the one hand it is the identity function with the assignment of HRS/LRS to 0/1 or, by inverting this result to HRS/LRS as 1/0, a logic NOT. The corresponding truth table is displayed in Table 3.1. However as the ReRAM state representation is arbitrary, both implementations require similar operation schemes and can be interchanged without loss of generality.

Because this methodology represents both a single device characterization and a basic logic function, it will be investigated in detail. The insights gained within this simple operation are of high importance as more complex functions will encounter similar challenges.

Table 3.1: Logic truth table for a one-Input logic IDENTITY / NOT.

A	**ReRAM representation**	**NOT A** $(\overline{A})$
0	HRS	1
1	LRS	0

3.2.2.1 Characterization of Bitcell Variability

As a first step, the program variability of memristive 1T-1R devices is investigated by programming and read-out multiple bit-cells under identical voltage conditions. Yet before the readout variability is investigated, the voltage conditions for programming are tested in order to achieve the best cycling. For this, the source and gate voltage during the SET, V_{SET} and $V_{\mathrm{G,SET}}$, and respectively for the RESET, V_{RESET} and $V_{\mathrm{G,RESET}}$, are varied and the resulting current states evaluated. Fig. 3.3 depicts the post-switching device currents in dependency of each voltage parameter. The data was collected for 64 devices on the MAD-200 chip which were programmed 16 times under each voltage condition.

The data in Fig. 3.3(a) indicates that for applied source voltages above $1\,\mathrm{V}$, the SET process is nearly uniform with a low variability. However by varying the gate voltage during the SET as shown in Fig. 3.3(b), a more gradual switching process can be performed. The high resistive state in Fig. 3.3(c) follows a linear dependency on the applied RESET source voltage while for the RESET gate voltage a saturation effect above $4\,\mathrm{V}$ is visible in Fig. 3.3(d). Based on the data, the final pulse schemes for programming and readout are listed in Table 3.2.

Table 3.2: Voltage settings used for programming 1T-1R devices in the MAD-200 test chip setup.

V_{SET}	$V_{\mathrm{G,SET}}$	V_{RESET}	$V_{\mathrm{G,RESET}}$	t_{pulse}
$2.0\,\mathrm{V}$	$1.0\,\mathrm{V}$	$1.5\,\mathrm{V}$	$5.0\,\mathrm{V}$	$200\,\mathrm{ms}$
	$V_{\mathrm{G,read}}$	V_{read}	$t_{\mathrm{pulse,read}}$	
	$5\,\mathrm{V}$	$0.25\,\mathrm{V}$	$200\,\mathrm{ms}$	

The readout measurements for the 1T-1R devices are conducted by applying a small voltage pulse and measuring the resulting current that flows through the device, as explained above. The outcome of

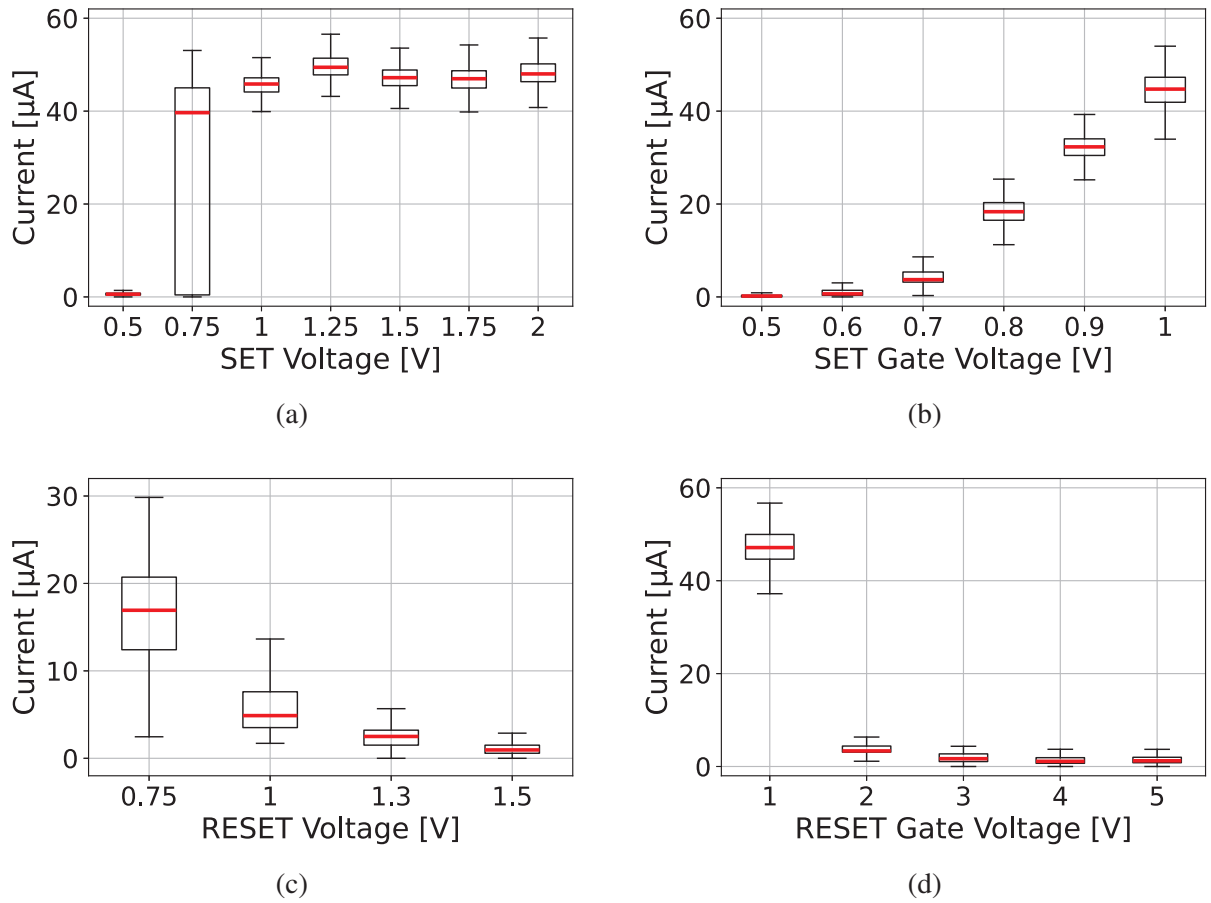


Figure 3.3: Experimental post-operation currents after SET (a) + (b), and RESET (c) + (d), in dependence of the applied voltage conditions for a fixed pulse length, $t_{\text{pulse}} = 200\,\text{ms}$ and readout voltage, $V_{\text{read}} = 0.25\,\text{V}$. The varied voltages are (a) V_{SET} for a constant $V_{\text{G,SET}} = 1.0\,\text{V}$, (b) $V_{\text{G,SET}}$ for a constant $V_{\text{SET}} = 2.0\,\text{V}$, (c) V_{RESET} for a constant $V_{\text{G,RESET}} = 5.0\,\text{V}$, and (d) $V_{\text{G,RESET}}$ for a constant $V_{\text{RESET}} = 1.5\,\text{V}$. Each boxplot contains the data of $1\,024$ resistance states, evaluated from 64 devices for 16 switching cycles.

this process is a time-current signal, $I(t)$. Fig. 3.4 presents 100 exemplary readout time-current traces from devices in the low resistive state (after SET) and high resistive state (after RESET). The traces include the initial rising and falling flanks of the voltage signal as visible on the right and left edge. In order to isolate the quasi-static currents, the signals are sampled only within the range $[5\,\%; 95\,\%]$ of the total pulse duration. After the sampling step, the arithmetic mean of each signal is calculated. These steps transform the original $I(t)$ data from a one-dimensional array into a single value per trace. The computed means of the full recorded data set are displayed in Fig. 3.5.

The measured data set contains 10 consecutive SET-Read-RESET-Read cycles for 512 devices. This results in a total number of $5\,120$ data points that are visualized as a histogram and as CDF plot in Fig. 3.6. This allows for an examination of the underlying distribution characteristics. The linear slope observed in the current values for the low resistive state suggests a normal distribution as the axes are linearly scaled. In contrast, the linear slope seen in the currents for the high resistive state on a semi-logarithmic plot indicates a log-normal distribution. This is visually supported by the histogram shapes.

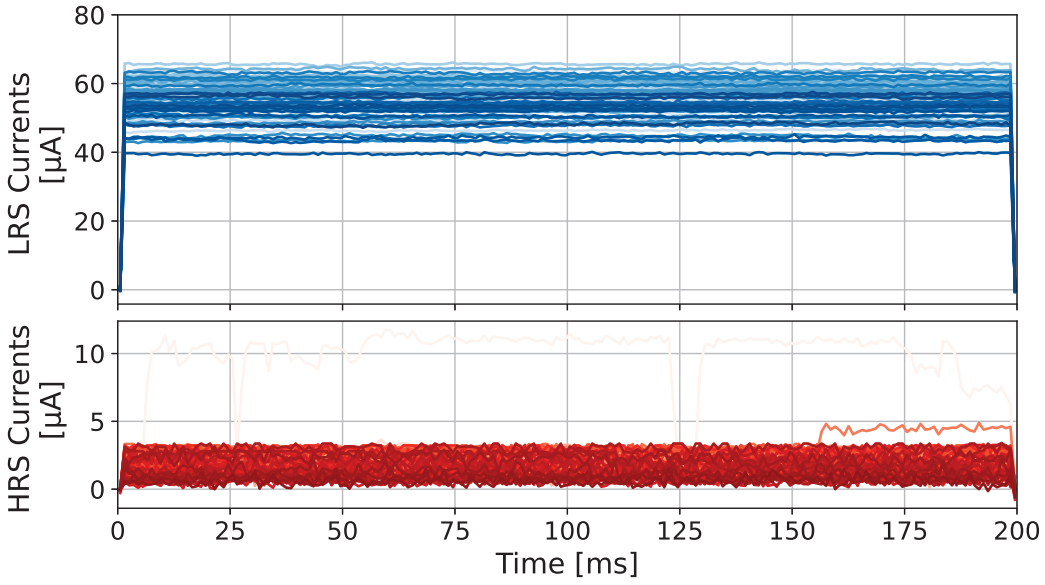


Figure 3.4: Experimental example of the current evaluation for a 1T-1R readout. The plot shows time-current signals from 10 cycles of SET-Read-RESET-Read, recorded from 10 different 1T-1R devices. The voltage pulses are applied according to Table 3.2.

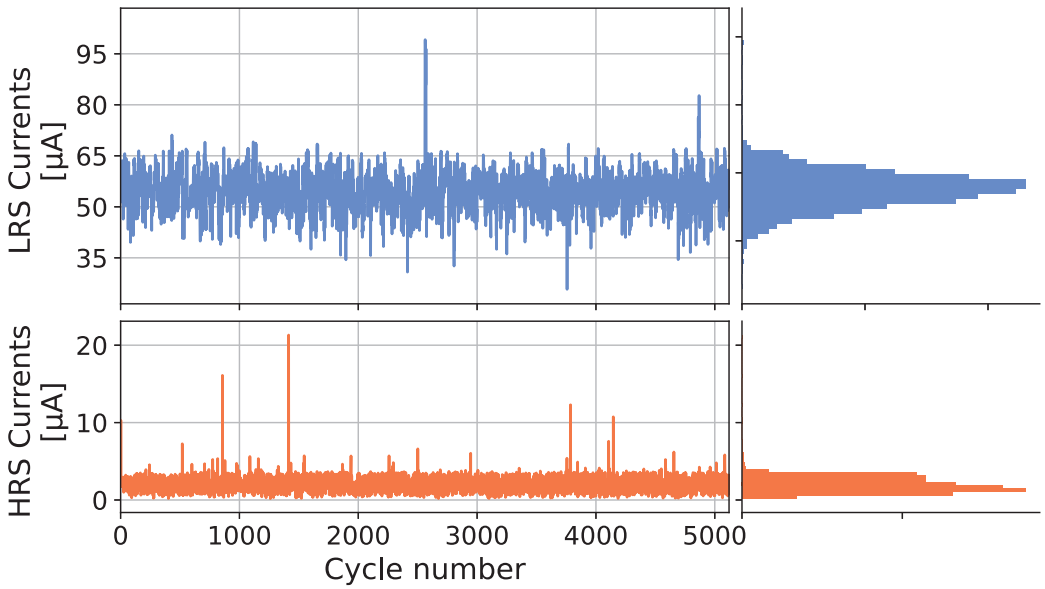


Figure 3.5: Experimental current means of LRS and HRS readout data with corresponding distribution histograms. The plot shows the arithmetic mean of the time-current signals for 10 cycles of SET-Read-RESET-Read, recorded from 512 different 1T-1R devices.

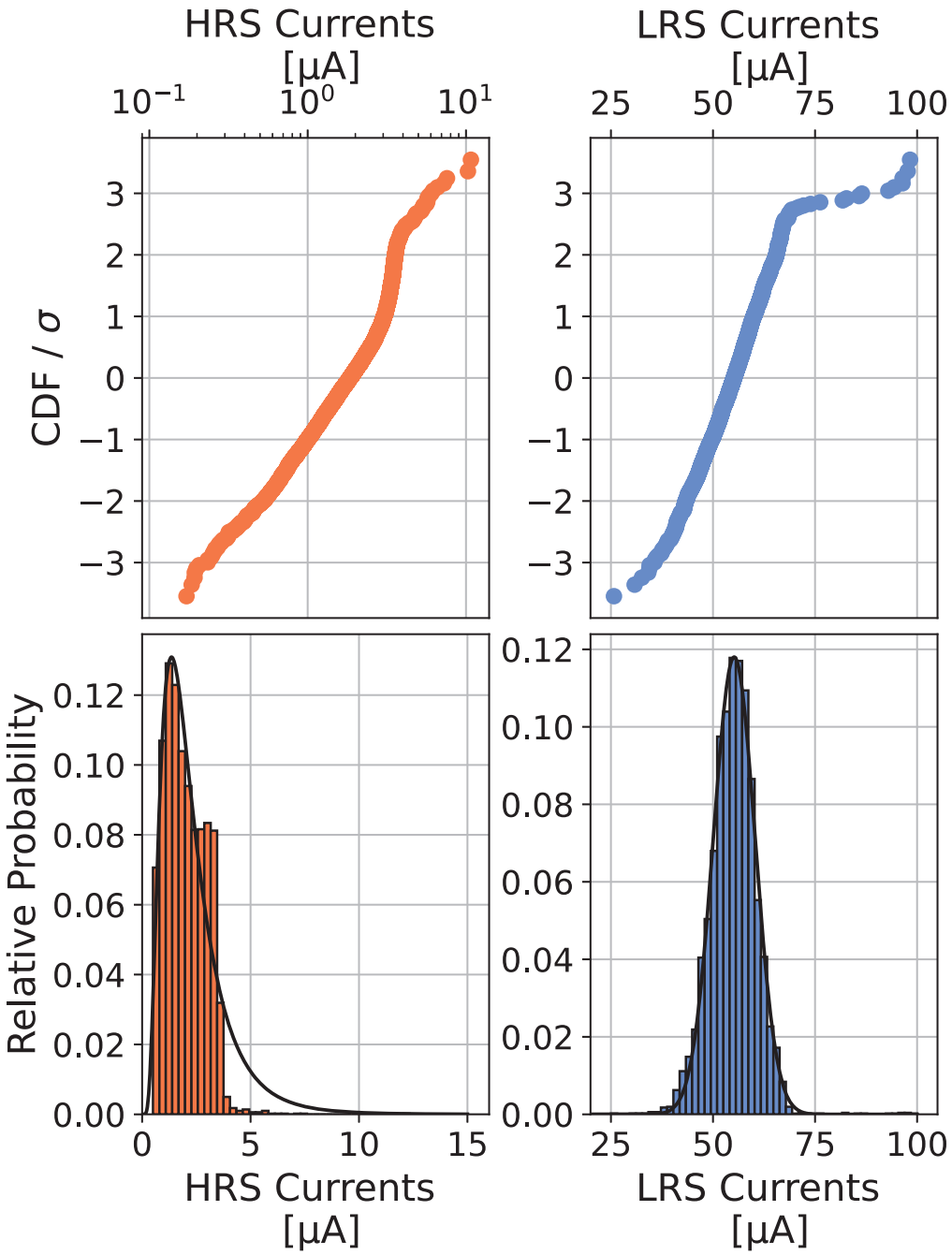


Figure 3.6: Histogram and CDF plot for experimental HRS and LRS single 1T-1R readout currents. Displayed are 5 120 data points, recorded from 512 devices and 10 consecutive read cycles. The LRS CDF is linearly scaled as the underlying data is expected to be normal-distributed. The CDF of the HRS data is depicted in a semi-logarithmic plot to better visualize the data spread its underlying log-normal distribution.

Based on the numerically fitted distribution parameters in Fig. 3.6 and the framework in Section 3.2.1, the threshold-dependent error probabilities for both, HRS and LRS, are calculated. For a device in the high resistive state an error occurs if the measured current lies above the separating threshold. Vice versa, if a device in the low resistive current returns a current lower than the threshold it will be wrongly classified as HRS. Fig. 3.7 depicts the threshold dependency of the error probability as well as the optimal threshold value which distributes the error probability symmetrically.

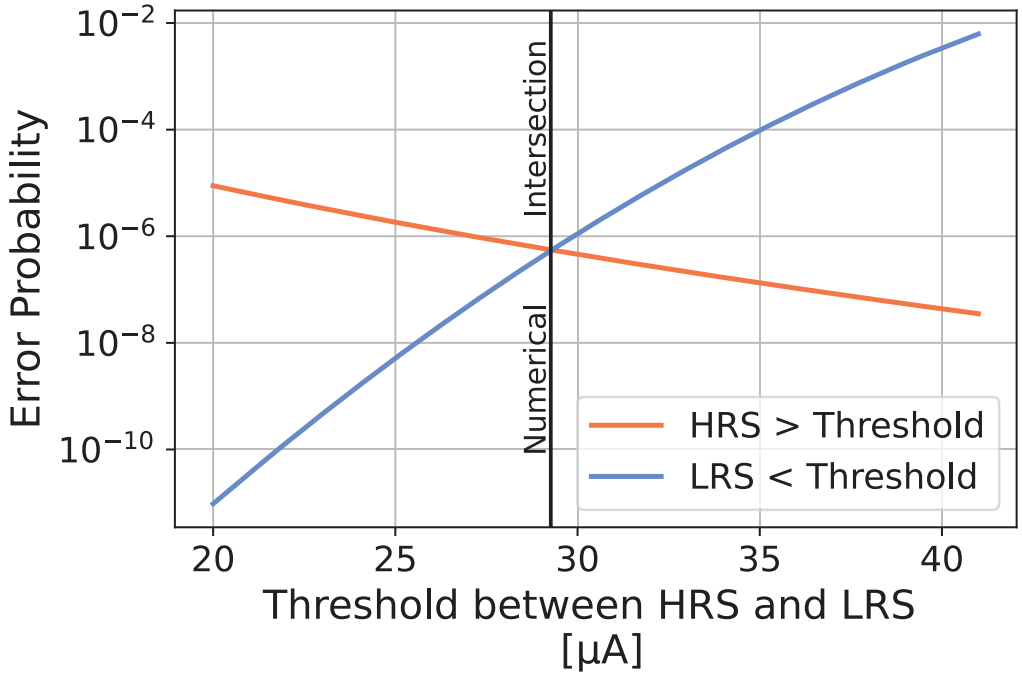


Figure 3.7: Calculated threshold-dependent error probability for a single device 1T-1R readout. Depicted is the probability for a measured LRS current to be below the adjusted threshold, respectively a HRS current to be above the threshold. The vertical black line denotes the numerically calculated intersection of the fitted log-normal (HRS) and normal (LRS) distribution.

It can be seen that without any optimization, the probability for logic errors in a single device readout is in the order of 1×10^{-6}, i.e., 1 wrong evaluation out of a million operations. This number acts as figure of merit while in the following sections several aspects of the readout operation and their influence on the error probability are investigated in detail. For all following analyses, if not stated otherwise, a similar evaluation procedure is conducted.

A similar investigation is conducted using simulated 1T-1R current data based on variability-aware modeling as described in Section 2.4.1. For the transistor models precise process and mismatch data is available from the PDK while the JART VCM v1b ReRAM model has to be adjusted. Three main factors are considered for each parameter: the distribution mean and standard deviation, and its truncation. The last point is essential in order to eliminate high-sigma outliers which might disrupt the simulation. This ensures that all generated values are within a physically reasonable range.

The simulation consists of $1\,000$ runs in which a device is initialized in HRS, followed by a sequence of SET-Readout-RESET-Readout. The resulting current data is processed in a similar manner as used for the experimental data. The goal is to match the simulation results to the experimental measurements.

Therefore, the simulations are conducted with different sets of variability parameters to create the best match.

Table 3.3: Parameters of the simulated 1T-1R current distribution for various initial variability ranges.

Distribution		Parameters HRS	Parameters LRS
Standard Deviation σ	Truncation		
$5\,\%$	$2\,\sigma$	$(1.73 \pm 0.20)\,\mu\mathrm{A}$	$(51.0 \pm 4.6)\,\mu\mathrm{A}$
$5\,\%$	$3\,\sigma$	$(1.74 \pm 0.19)\,\mu\mathrm{A}$	$(51.1 \pm 5.1)\,\mu\mathrm{A}$
$10\,\%$	$2\,\sigma$	$(1.68 \pm 0.48)\,\mu\mathrm{A}$	$(52.1 \pm 7.6)\,\mu\mathrm{A}$
$10\,\%$	$3\,\sigma$	$(1.67 \pm 0.51)\,\mu\mathrm{A}$	$(54.0 \pm 6.7)\,\mu\mathrm{A}$
Experimental Data		$(1.36 \pm 0.60)\,\mu\mathrm{A}$	$(55.2 \pm 5.1)\,\mu\mathrm{A}$

The simulation results are listed in Table 3.3 together with fitted experimental distribution parameters from Fig. 3.6. The simulated LRS currents closely align with the experimental values while the simulated HRS currents are slightly higher than the experimentally observed currents. Yet they fall within one standard deviation of the experimental measurements. A perfect numerical match could only be achieved if the model and experiment are identical in aspects such as process node and material stack. For the subsequent simulations the combination of $\sigma = 10\,\%$ and a truncation of 3σ is selected. The corresponding histogram and CDF is shown in Fig. 3.8. These distributions yield to an intersection value of $26.31\,\mu\mathrm{A}$ compared to $29.27\,\mu\mathrm{A}$ for the experimental data. The error probability at the intersection is 2.6×10^{-5} (simulation) and 1×10^{-6} (experiment), therefore the simulations overestimate the variability compared to what is observed experimentally. This can actually be beneficial from a design and safety perspective as the simulations serve as a worst-case analysis.

3.2.2.2 Influence of the Operation Voltage

As previously mentioned, in the 1T-1R arrangement the total resistance path consists of both, one transistor and one ReRAM device. Therefore, the total current I_{sense} is inseparably influenced by the interplay between ReRAM state and the transistor behavior. The transistor behavior is significantly depending on the applied operation voltages, V_{read} and V_{G}. These voltages will determine the drain-source current of the transistor, $I_{\mathrm{D,S}} \propto V_{\mathrm{G,S}} = V_{\mathrm{G}} - V_{\mathrm{read}}$, which effectively correlates to the channel resistance.

Fig. 3.9(b) and Fig. 3.9(a) display the output current dependency for HRS and LRS on V_{read} and V_{G}. The data set contains 32 switching cycles for 32 devices, in total $1\,024$ resistances states. For each state, the current is readout 25 times in a matrix measurement for 5 values of V_{read} and 5 values of V_{G}. During the readout, all voltages are stepped through in an increasing mode in order to reduce any read disturb.

The data highlights a critical impact of the transistor operation mode on the accuracy of the 1T-1R readout. Specifically while measuring an LRS current, a reduced gate voltage, $V_{\mathrm{G}} \leq 3\,\mathrm{V}$, results in higher transistor drain-source channel resistances and thereby limits the 1T-1R readout current. Moreover, for both states the readout current increases proportionally with the applied sensing voltage. This indicates that a higher V_{read} is beneficial for enhancing the absolute current sensing margin. However, the read voltage, V_{read}, must be sufficiently low to ensure state conservation. The origin of this dependence can be understood by approximating the total 1T-1R path series resistance as an ohmic

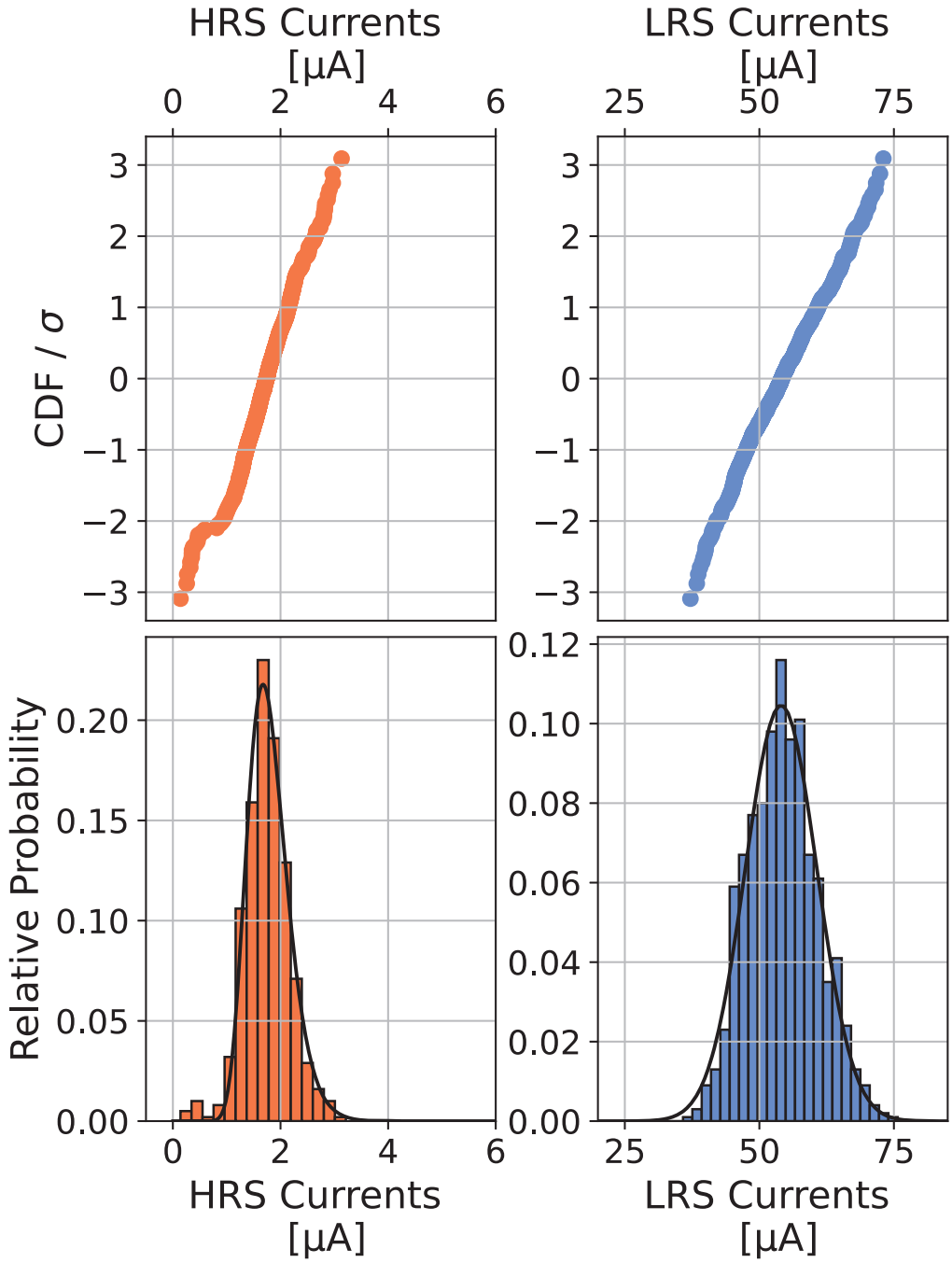
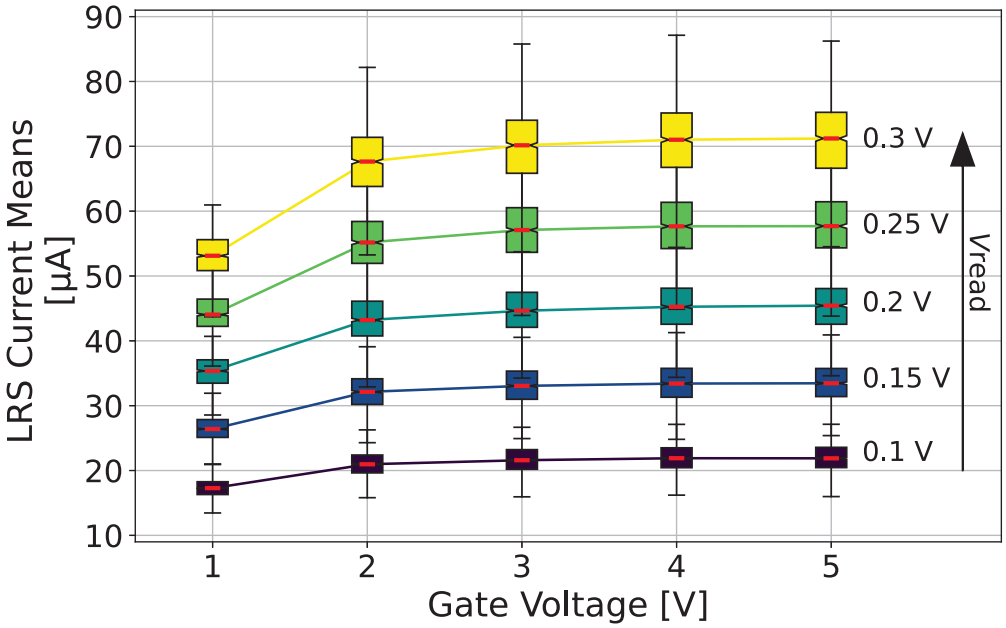
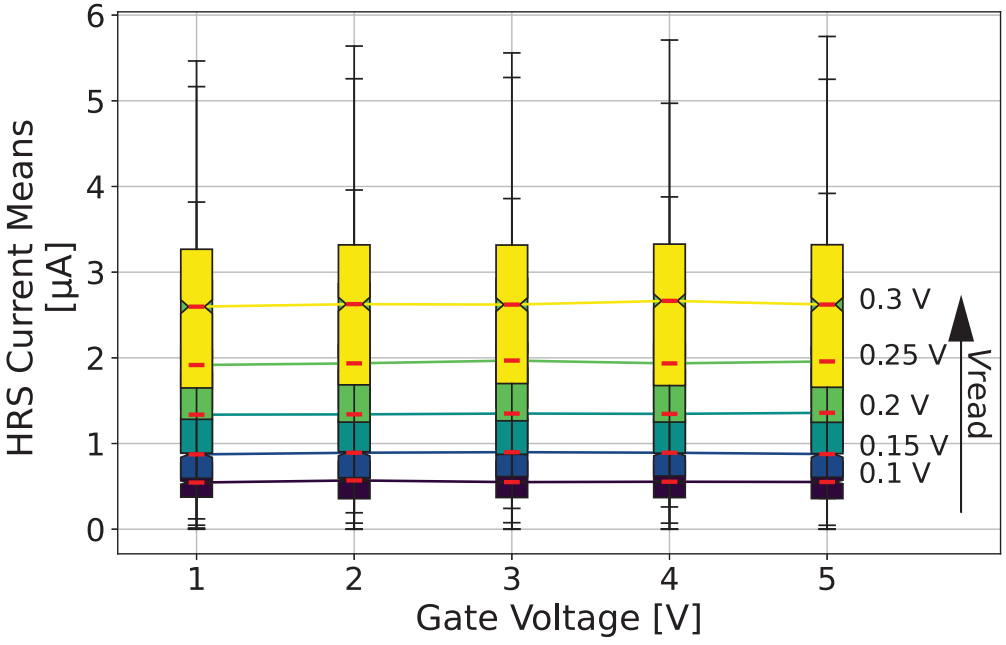


Figure 3.8: Histogram and CDF plot for simulated HRS and LRS single 1T-1R readout currents. Displayed are 1 000 simulated data points with the variability parameters as described in Table 2.1 and Table 2.2. For all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed. Both, the LRS CDF and HRS CDF, are linearly scaled. In contrast to the experimental HRS data in Fig. 3.6, the simulated HRS data shows a reduced spread and trend towards a normal distribution.



(a)



(b)

Figure 3.9: Experimental dependence of the readout currents for (a) LRS and (b) HRS on the applied read voltage and transistor gate voltage, V_{read} and V_{G}. Each boxplot contains the data of 1 024 resistance states, evaluated from 32 switching cycles for 32 devices.

resistance and applying Ohm's Law, $I_{\text{sense}} = V_{\text{read}}/R_{\text{1T-1R}}$. Under the assumption that the ReRAM state remains unaltered and the transistor operates in saturation region, an increase in the readout voltage, V_{read}, will correspond to a linear increase in the sensed current, I_{sense}.

This assumption is however limited by the intrinsic *non-linear* or non-ohmic characteristic of VCM devices. Under an applied bias voltage, V, the current through a ReRAM is voltage-dependent, $I = I(V)$. This non-linearity is caused by the underlying current transport mechanism [208]. The consequence of this aspect is that doubling the readout voltage will not result in a two times larger sensing current [209]. This effect is demonstrated for a simulated 1T-1R device readout with increasing V_{read} and constant V_{G} and varied initial oxygen vacancy concentrations, N_{Init}, which effectively set the ReRAM state. The simulated results for the equivalent 1T-1R resistances, calculated as $V_{\text{read}}/I_{\text{sense}}$, are displayed in Fig. 3.10(a) and Fig. 3.10(b). It is observed that for the JART VCM v1b model, the non-linear voltage-current dependence increases for lower N_{Init} values, e.g., a higher initial ReRAM resistance. Furthermore, the non-linearity shows a reciprocal behavior as the sensed device resistance reduces for increasing voltages. This behavior has also been observed by Kim et al. [210].

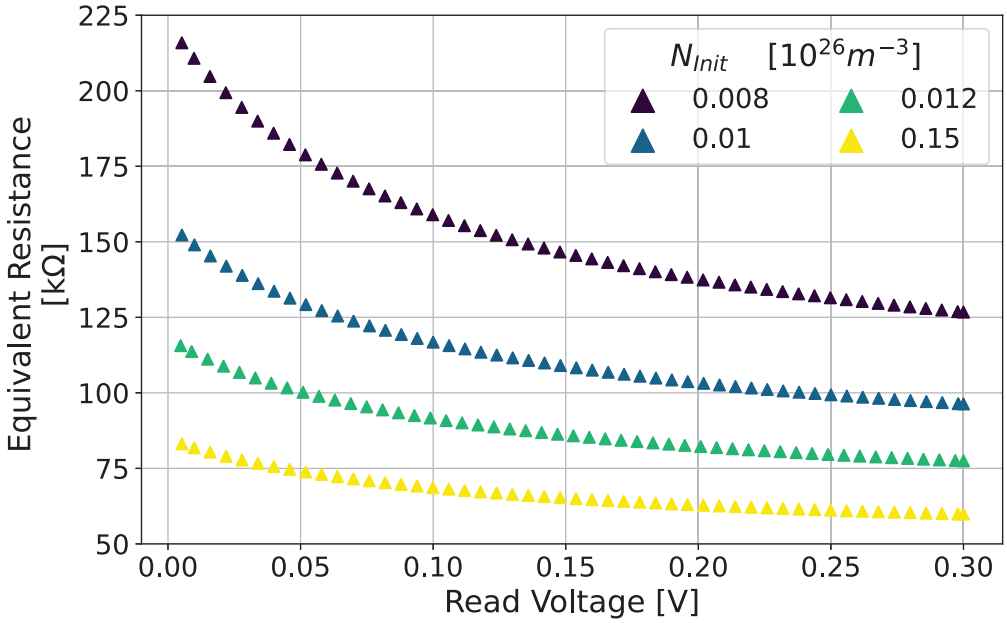
The non-linearity is further investigated experimentally based on the measured data in Fig. 3.9(b) and Fig. 3.9(a). As seen in Fig. 3.11(b), the low resistive state behaves linearly up to a voltage of $V_{\text{read}} = 0.3\,\text{V}$. However, the high resistive state in Fig. 3.11(a) shows a pronounced non-linearity as the equivalent device resistance reduces with increasing read voltages. This behavior matches the expected characteristics from the simulation as well as the literature results.

In order to estimate the total influence of V_{read} on the error probability, the current distributions for 3 different read voltages at $V_{\text{G}} = 5\,\text{V}$ are analyzed. Fig. 3.12 depicts the corresponding histograms with fitted distributions. It is observable that both, HRS and LRS, shift towards higher absolute current values with increasing voltages. At the same time, the absolute difference of the distribution means, ΔI_{Mean}, increases, too. For the simplification of a linear resistance, scaling the readout voltage, V_{read}, by a factor , k, results in a current scaling of approximately $k \times \Delta I_{\text{Mean}}$. The non-linearity might cause this trend to skew towards higher HRS means and lower current differences. However, the increase in readout voltage still achieves a sufficient current scaling.

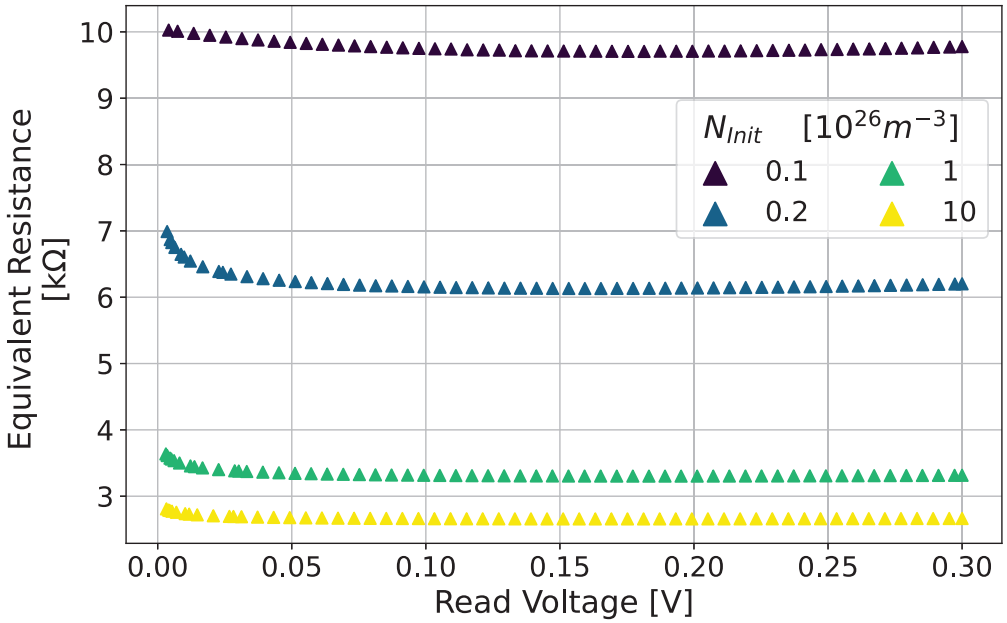
To quantify the impact of the operation voltage, Table 3.4 presents the calculated error probabilities. The results support the heuristic observation that a higher readout voltage V_{read} correlates to an improved current separability. In the previous section an error probability of 1×10^{-6} was estimated for a data sets recorded at $V_{\text{read}} = 0.25\,\text{V}$. The voltage-dependent results in Table 3.4 align with this number and emphasize the critical impact of the readout voltage on a correct state classification.

Table 3.4: Calculated voltage-dependent error probability for the experimental single device readout. The calculation is performed based on the data displayed in Fig. 3.12.

V_{read} [V]	Intersection [µA]	Probability 'False 1'	Probability 'False 0'
0.1	13.37	4.42×10^{-4}	7.49×10^{-4}
0.2	25.96	2.84×10^{-5}	9.72×10^{-5}
0.3	31.11	3.72×10^{-10}	2.01×10^{-10}

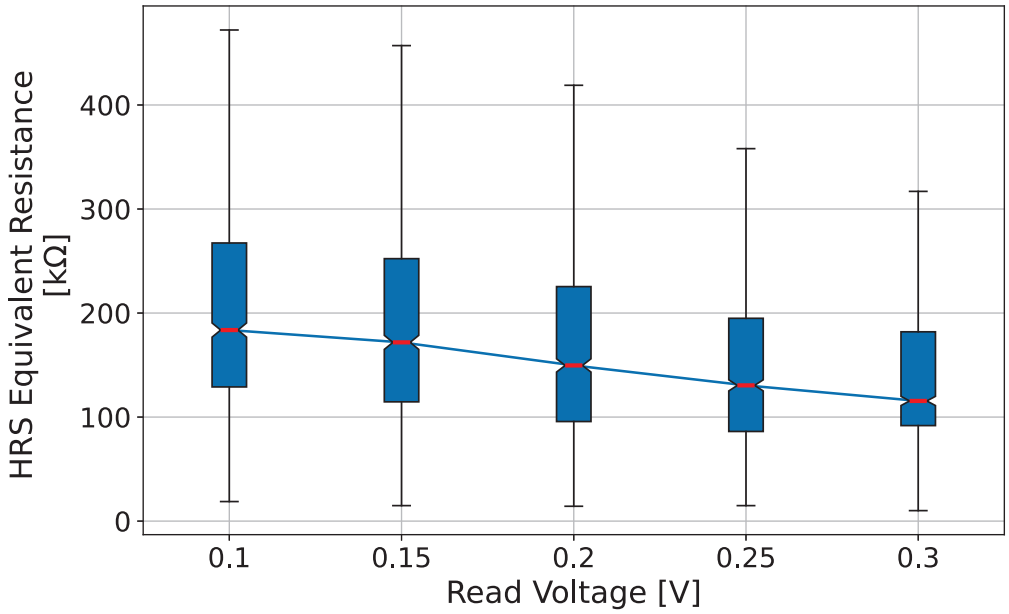


(a)

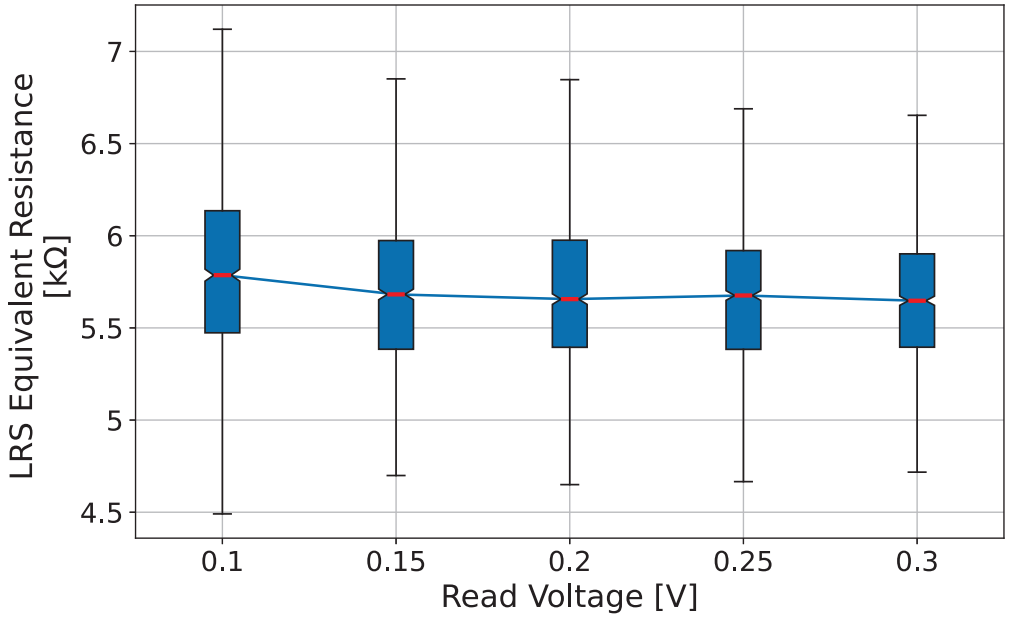


(b)

Figure 3.10: Simulation results of the equivalent resistances, $R_{\text{1T-1R}}$, calculated as $V_{\text{read}}/I_{\text{sense}}$ for initial oxygen vacancy concentrations, N_{init} in the (a) HRS and (b) LRS range for increasing readout voltages, V_{read}, and a constant $V_{\text{G}} = 5\,\text{V}$. For the simulations, all parameters in Table 2.1 and Table 2.2 were used.



(a)



(b)

Figure 3.11: Experimental non-linearity in the measured equivalent resistance, $R_{\text{1T-1R}}$, calculated as $V_{\text{read}}/I_{\text{sense}}$ for (a) HRS and (b) LRS at increasing readout voltages, V_{read}, and a constant $V_{\text{G}} = 5\,\text{V}$. The equivalent resistances are calculated from the data depicted in Fig. 3.9(a) and Fig. 3.9(b). Each boxplot contains the data of 1 024 resistance states, evaluated from 32 switching cycles for 32 devices.

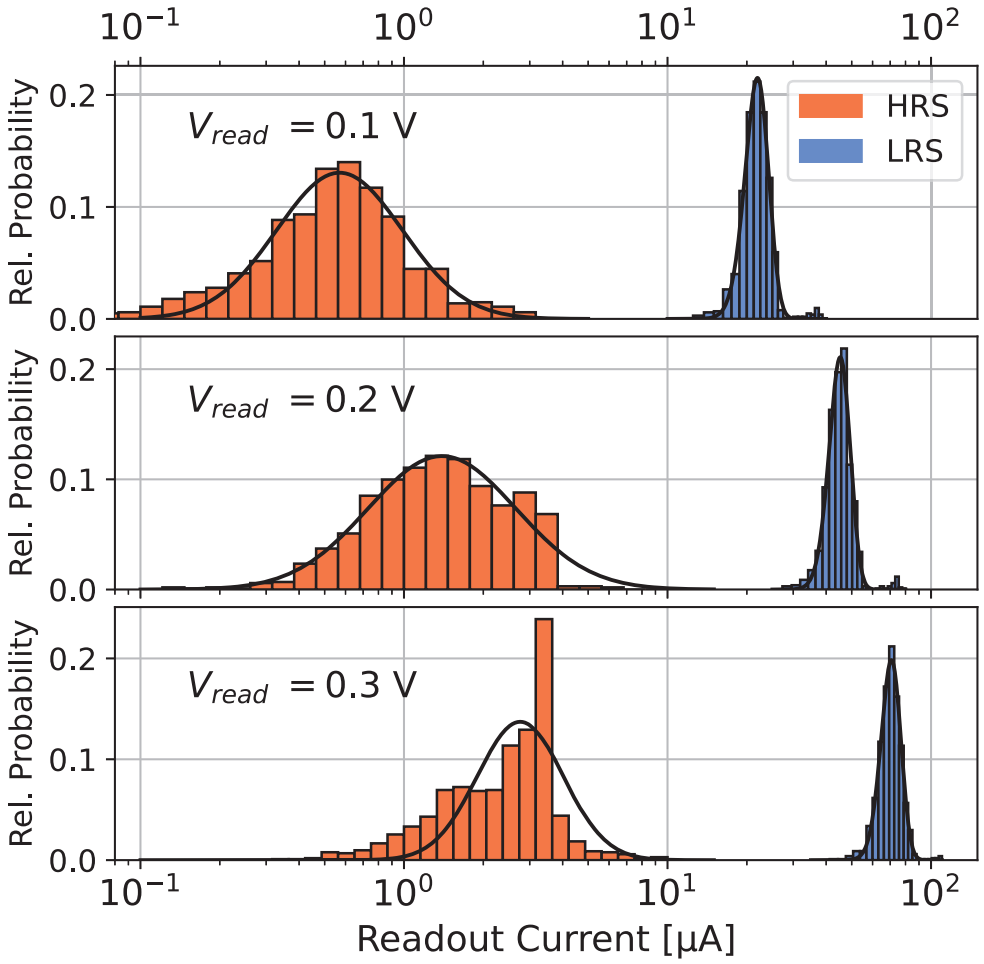


Figure 3.12: Experimental readout HRS/LRS current distributions for increasing V_{read} values at constant $V_{\text{G}} = 5\,\text{V}$. The X-axis is scaled in a logarithmic fashion to better visualize the data spread. Each histogram contains 1 024 data points evaluated from 32 switching cycles for 32 devices.

3.2.2.3 Influence of the Operation Position

The investigated voltage dependency is linked to the physical positioning of the bitcells within a memristive crossbar. For any operation the voltages are set from the outside to a specific value. Even if the output at the source is assumed to be ideal, it is essential to consider the effects of the crossbar wiring and parasitic elements. Due to these resistances and the caused *IR*-drops along the metal interconnects, the voltages applied at each device will depend on its specific position within the crossbar [211]. As seen above in Section 3.2.2.2, the sensed output current is highly voltage-dependent. Therefore if the operation voltages, namely V_{read} or V_{G}, decrease due to parasitic effects, it might have an impact the operation reliability [212].

The parasitic effect is initially investigated with a simulation of a 1T-1R crossbar according to the parameters in Table 2.1 and Table 2.2. The simulations consist of consecutive single device LRS readouts across all 256 Bit-Lines for a specific pair of Signal-Line and Word-Line while incorporating distributed parasitic RC elements per 1T-1R cell. Simulating such a reduced crossbar size (256 x 1) drastically reduces the computation costs compared to the full crossbar of 256 x 256. To include varying SLs/WLs in the simulations, a lumped RC model was exploited [213] with a constant coupling

capacitance of $10\,\mathrm{aF}$ and a resistance per unit cell between $0.1\,\Omega$ and $10\,\Omega$ as suggested by similar studies [214].

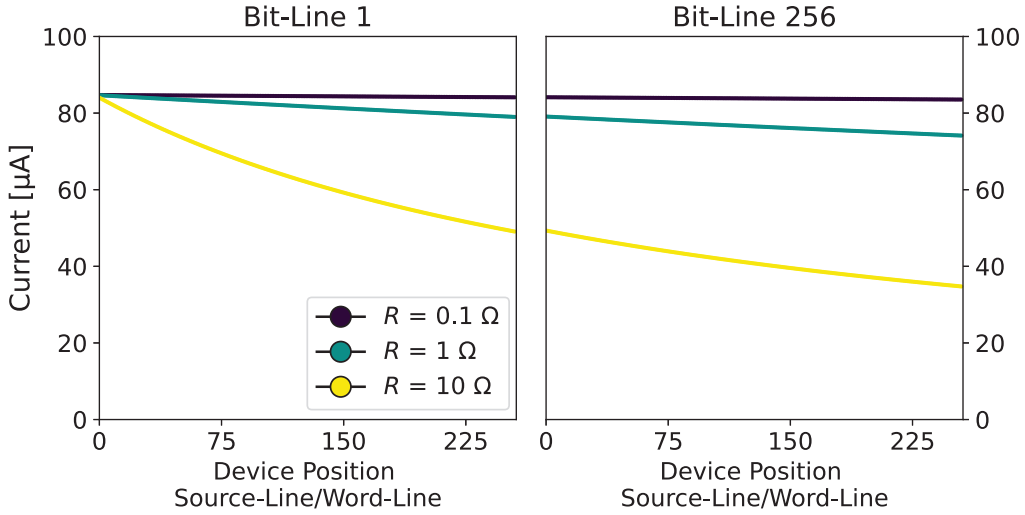


Figure 3.13: Simulated currents for single LRS device readout operations with included parasitic RC elements. Depicted are the simulations results for the first (BL 1) and last (BL 256) rows for all 256 Signal-Lines/Word-Lines (SLs/WLs) in a 256 x 256 1T-1R crossbar. The readout current is sensed on the far left (SL/WL 1). For all simulations a coupling capacitance of $C = 10\,\mathrm{aF}$ was used.

The simulation results are depicted in Fig. 3.13. For low parasitic resistances up to $0.1\,\Omega$ per unit cell, the readout current is constant throughout the whole 256 x 256 crossbar. However a higher line resistance will lead to a reduction of the measured current of up to $50\,\%$. Several strategies might be followed in order to compensate this effect towards reliable operations. First, stabilizing the sensing voltage depending of the operation position. By including buffer devices along the interconnecting metal lines, the voltage drop between the individual bitcells is compensated [190]. However, this strategy requires complex analog periphery.

In addition, also the reference current threshold within the state classification can be adjusted. As previously discussed, this threshold is exploited for converting the analog current into digital values and can be adjusted arbitrarily. If the absolute sensing current scales down due to line resistances, the reference threshold can be lowered accordingly. This enables a classification of low resistive states even for the reduced currents seen in Fig. 3.13. However as described above, the error probability increases with lower absolute readout currents as the sensing margin effectively shrinks.

Finally, a straight-forward measure is to limit the crossbar dimension to certain values, e.g., 256 x 256 in order to bound the total parasitics [211, 215]. While this strategy does not require complex coping mechanism, it increases the total chip footpath. One goal during the chip design is to minimize the required periphery and interconnect area per stored data, either by shrinking the periphery or by increasing the number of bitcells which share a peripheral device. Here, an additional trade-off occurs between increasing the crossbar dimensions for maximizing the area efficiency on the one hand, and maintaining the operation reliability by limiting the negative influence of parasitic elements on the other hand.

Fig. 3.14 contains the measured readout currents for high resistive and low resistive devices along 256 Signal-Lines/Word-Lines in the 1T-1R arrangement. For a better visualization, the devices are grouped into 8 packs of 32 devices and each device is cycled 10 times. The observed currents show no

clear trend of degradation besides statistical fluctuations. This indicates that the line resistances in the MAD-200 design are comparably low and the crossbar dimension (256 x 32) can be exploited for logic evaluations. The line width in the measured 1T-1R crossbar array ranged between $w_{\mathrm{WL}} = 0.36\,\mu\mathrm{m}$ for the interconnects at the transistor gates (Word-Lines), $w_{\mathrm{SL}} = 1.09\,\mu\mathrm{m}$ at the transistor source (Signal-Lines), and $w_{\mathrm{BL}} = 0.72\,\mu\mathrm{m}$ at the ohmic ReRAM electrode (Bit-Lines). The specific metal line resistance was approximately $\rho = 4 \times 10^{-8}\,\Omega\,\mathrm{m}$, the metal layer thickness $t \approx 500\,\mathrm{nm}$ and the maximum interconnect length, $L \approx 1\,\mathrm{mm}$ for 512 devices in total. The total parasitic line resistance can be calculated from this as $R_{\mathrm{line}} = \frac{\rho L}{w_{\mathrm{line}} t}$. This yields to the values $R_{\mathrm{WL}} \approx 222\,\Omega$ and $R_{\mathrm{SL}} \approx 73\,\Omega$. For 512 devices per Signal-Line and Word-Line, this corresponds to a resistance per unit cell of $R_{\mathrm{Unitcell,WL/SL}} \leq 1\,\Omega$. The Bit-Line resistance, $R_{\mathrm{BL}} \approx 111\,\Omega$, is only shared by 32 devices which results in a unit resistance of $R_{\mathrm{Unitcell,BL}} \approx 3.5\,\Omega$.

These values are comparable to other reports of parasitic line resistances in resistive 1T-1R crossbar arrays such as $6.64\,\Omega$ / $14.64\,\Omega$ per unit cell [214] and $3.36\,\Omega$ per unit cell [216]. Furthermore, the simulations results in Fig. 3.13 indicated that for line resistances, $R_{\mathrm{Unitcell}} \leq 1\,\Omega$, the parasitic voltage drop will cause nearly no current reduction. This explain the non-observed position dependency in Fig. 3.14.

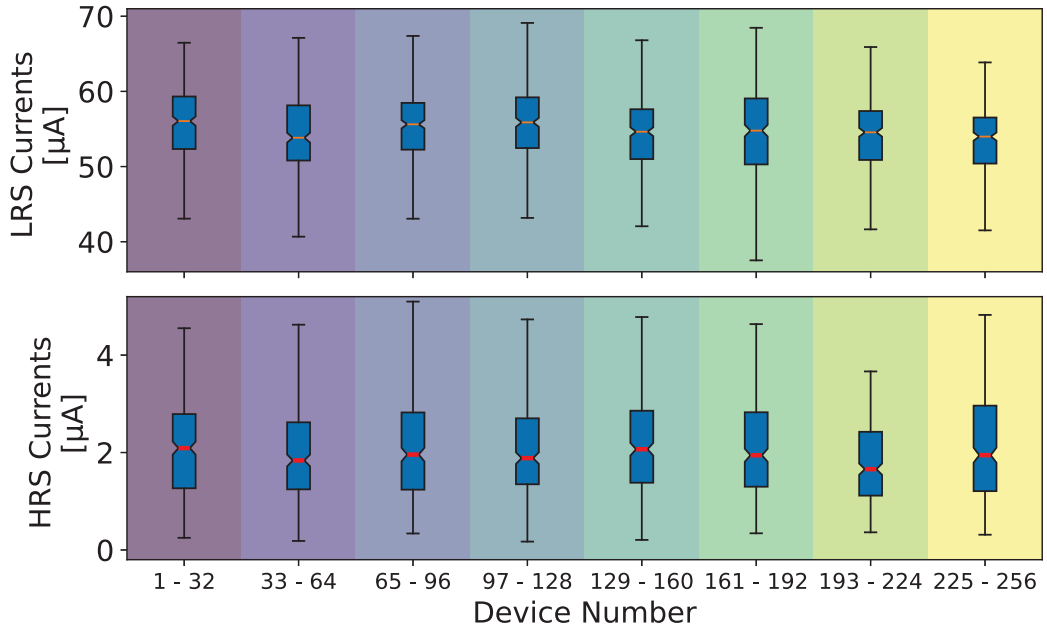


Figure 3.14: Experimental readout currents of LRS (top) and HRS (bottom) devices in a 1T-1R crossbar for the positions SL/WL 1-256 at BL 5. The readout current at each BL is measured on the far left (SL/WL 1). Each box contains 320 data points of 32 devices which are cycled 10 times.

3.2.2.4 Read Instability in Device Classification

The next step is to investigate the *read-to-read variability*, or read instability, and to compare the influence of this read variability to the previously discussed switching variability, D2D and C2C. For this purpose, a data set of 12 devices is used. Each device is programmed 12 times and after each programming, the 1T-1R cell is read out 100 times at $V_{\mathrm{read}} = 0.2\,\mathrm{V}$ and $V_{\mathrm{G}} = 5\,\mathrm{V}$.

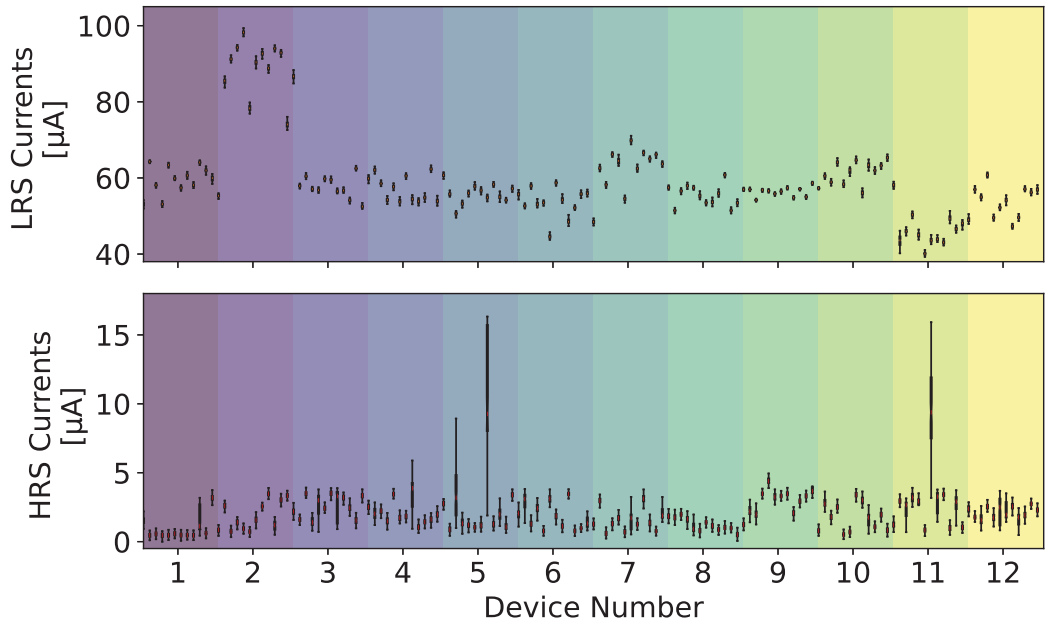


Figure 3.15: Boxplots of experimental readout currents for LRS (top) and HRS (bottom) under read variability. Each boxplot contains 100 readout currents, measured across 12 switching cycles and 12 devices.

The measured currents are depicted as boxplots in Fig. 3.15, organized by device and cycle. Multiple contributions are visible in the data. Firstly, in the low resistive state the cycle-to-cycle variability, illustrated by the difference between the individual boxplots, shows a larger absolute value as the read-to-read instability, i.e., the range of each boxplot. This observation matches with previous studies which already demonstrated that read noise in resistive devices is less significant compared to the variability induced by switching [88, 217]. Furthermore although the average current across multiple programmed cycles for each specific device remains relatively consistent, there is a notable spread across different devices. Particularly in the devices 2 and 11, a noticeable trend of clustering is observed. This feature can be explained by the theory that cycle-to-cycle variability is primarily influenced by the initial electro-forming conditions and other defect mechanism [55, 67]. Depending under which condition the conductive filament is initially formed, its low resistive behavior under subsequent cycling tends to be constrained or predetermined to some extent [55, 218].

In contrast to this, the high resistive state for the devices 2 and 11 displays no such clustering effect. However, individual states demonstrate a larger current spread with a magnitude over $10\,\mu\mathrm{A}$. This refers for instances to the cycles 3 and 8 for device no. 5 and cycle 7 for device no. 11.

To identify the origin of this phenomenon, all sequential readout currents for these devices and cycles are plotted in Fig. 3.16. The read variability of all 3 devices and cycles does not seem to follow a random pattern, but rather it demonstrates discrete current plateaus with abrupt jumps in between. Especially for the two cycles of device no. 5, these jumps occur in both, increasing and decreasing, direction. Within these plateau, for example in cycle 8 of device no. 5 between operations 15 and 52, the readout current shows stable behavior without large variability. This observation correlates to previous findings where discrete jumps in the HRS readout current have been observed [77, 81]. Wiefels et al. [81] explained this behavior by the occurrence of jumps of oxygen vacancies that appear randomly between the filament region with a high oxygen deficiency and the less oxygen-deficient filament gap. In addition to this larger-scale noise caused by discrete oxygen vacancy hopping, another

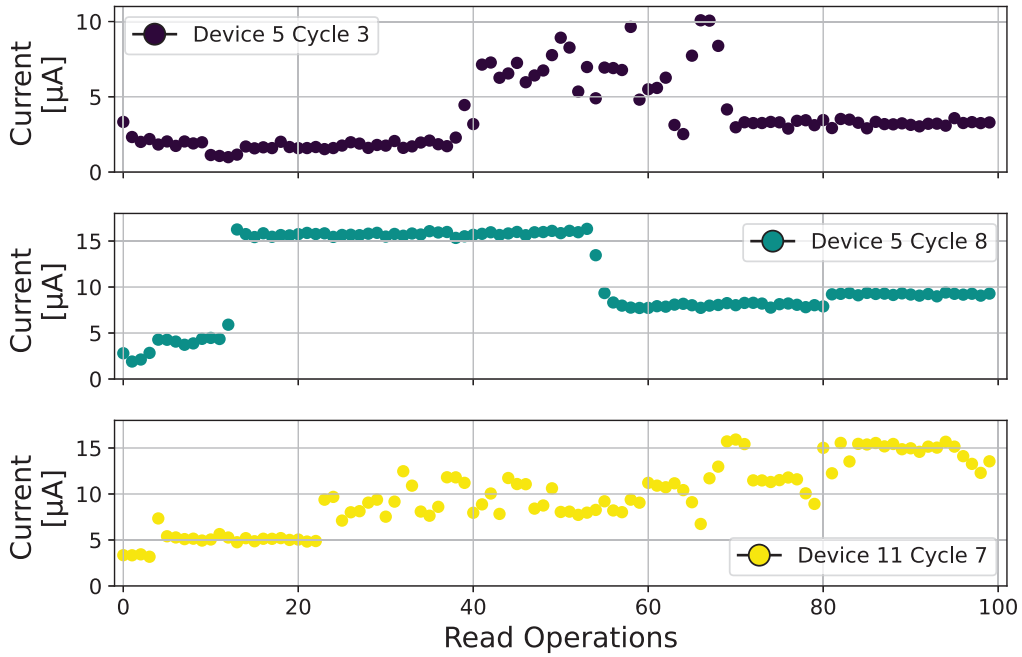


Figure 3.16: Experimental HRS current data of 100 subsequent readout operations for 3 selected devices and the programming cycles that exhibited the highest absolute current spread, as identified in Fig. 3.15.

noise contribution can be observed in the data of device no. 5 cycle 3. However, it is not clear if those fluctuations arise from the ReRAM or if they are caused by the transistor or the external measurement equipment, for instance as Johnson-Nyquist noise. For the data of device no. 5 cycle 8 it is seen that the increased current is only a temporal feature. Although there is no permanent distribution shift, even these temporary increased currents could compromise operational reliability. Therefore, the design must also account for these temporary current jumps to ensure a proper reliability.

The data for cycle 7 of device no. 11 shows a slightly different behavior. As before, discrete current jumps and the formation of plateaus are seen. However in this case, the jumps appear unidirectional towards increasing currents and with a higher additional variability. If the readout operation is performed in SET direction, i.e., V_{read} is applied at the ohmic electrode, a device in HRS state might switch cumulatively towards the low resistive state under subsequent applied readout pulses even if each individual pulse amplitude is lower than the original switching threshold [189, 219]. This would correlate to an unidirectional current increase. However, the readout in this setup is performed in the RESET direction, therefore this explanation does not hold. A similar finding of increasing readout currents under a RESET polarity was reported by Degraeve et al. in 2012 [220]. The group explained the behavior with an “hour glass model”, where the disturb is not caused by a filament growth but a dynamic transition into an equilibrium due to the provided energy.

To further investigate the impact of the readout operations on the stored information, the cumulative distribution function (CDF) of the current distributions for each device and cycle are calculated. The CDFs for the first, 50-th, and 100-th readout operation are illustrated in Fig. 3.17(a) and Fig. 3.17(b). It can be seen that the low resistive current CDF remains consistent across consecutive readouts which indicates a stable programmed state. The previously mentioned device 2 in Fig. 3.15 causes here a noticeable tail toward higher readout currents, specifically above 70 µA. In contrast, the high-current tail of the HRS distribution exhibits a shift towards higher readout currents over multiple readouts. This

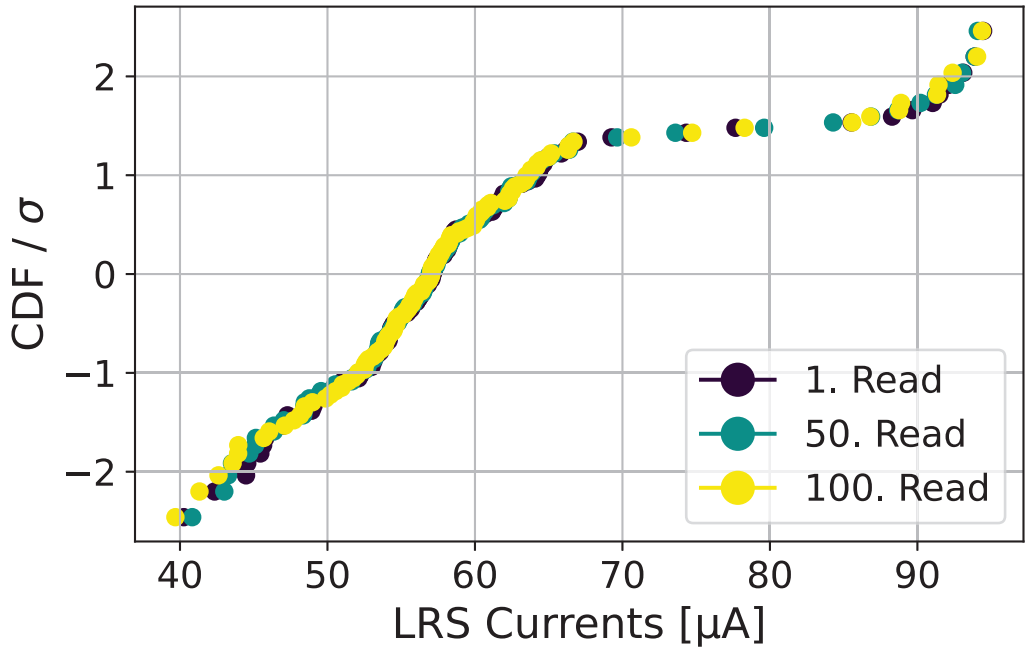
does not affect the distribution below $3.5\,\mu A$, but the currents between 6-10 μA. This behavior matches to the individual devices mentioned above.

Another read disturb measurement is conducted with the Elmos setup. For this data set, the readout operation is performed in both, RESET and SET direction. For the RESET readout, a voltage pulse is applied at the Signal-Line (SL) which is connected to the transistor source. The transistor drain connects to the active electrode of the ReRAM. The ohmic electrode of the ReRAM is connected via the Bit Line (BL) to ground potential. This assignment is called *RESET* direction as a positive voltage pulse at the transistor source and AE will reset the device. In a similar manner, the *SET* direction readout is implemented via a voltage pulse at the BL and ground potential at the SL. For both cases, the transistor gate is biased with a voltage pulse of $V_G = 6\,V$ to ensure that the transistor operates in saturation. In each direction, one device is cycled 10 times in HRS and LRS, and then readout with an equivalent pulse of 500 x $100\,\mu s$.

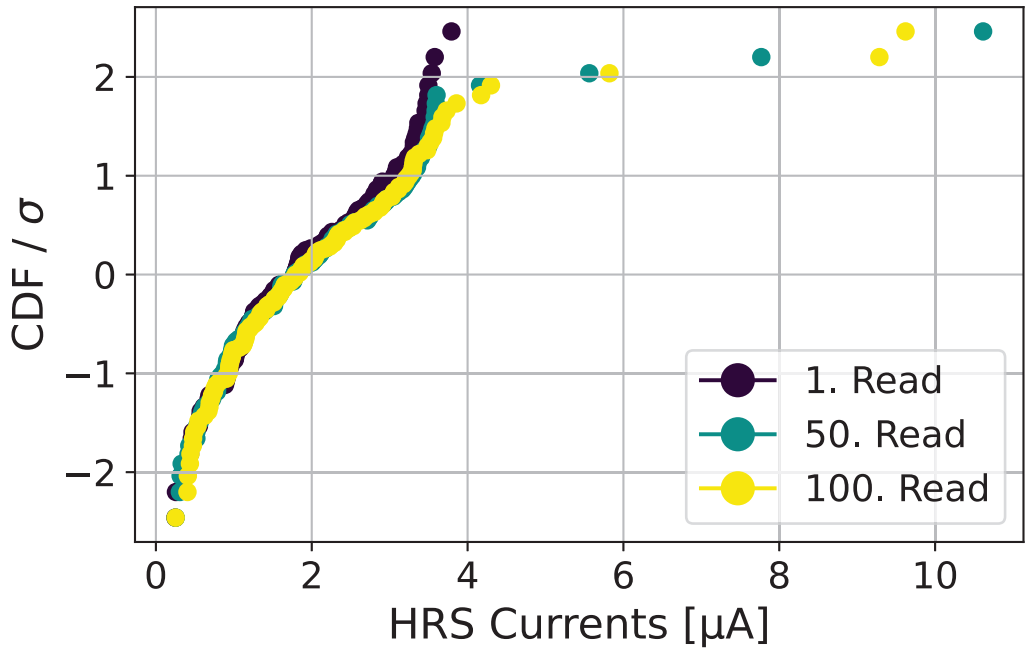
The measured current traces are displayed in Fig. 3.18. The data shows no drift of the readout currents, neither towards higher values in the SET direction or lower current in the RESET direction. First, the measured LRS currents are in both directions absolutely lower as in the MAD-200 data with the same $V_{read} = 0.25\,V$ in Fig. 3.9(a) while the HRS current ranges are similar as compared to Fig. 3.9(b). This difference could be related to the differences between both setups. For example, both designs exploit different CMOS technologies and differ in the actual transistor dimensions. It is possible that the gate voltage of $V_G = 6\,V$ is still not high enough to operate the Elmos transistor in full saturation. Furthermore, the MAD-200 chip consists of HfO_x-based ReRAM while TaO_x-based ReRAM were integrated on the Elmos chip. While the underlying filamentary switching mechanism is similar in both materials, they differ in their resistive properties such as memory window or retention [221, 222]. Azzaz et al. [222] reported a higher LRS resistance in TaO_x compared to HfO_x which supports the observation.

Additionally, in the Elmos data set there is a distinct asymmetry between the SET and RESET direction currents, with the SET direction consistently demonstrating lower current values. The currents differ approximately by a factor of 2. This kind of asymmetry in the $I - V$ characteristics of filamentary ReRAM is already described in the literature [223]. In the 1T-1R arrangement, an additional asymmetry might be introduced by the *body-effect* in the MOSFET device. This effect describes a widening of the substrate depletion region for a negative bulk-source voltage. This will result in an increased threshold voltage [186, P.148]. For constant bias voltages, V_{read} and V_G the transistor drain-source current will differ in devices with and without present body effect. However, if the active electrode is connected to the transistor, a body effect should only occur in the RESET direction and a lower current in this direction is expected [187, 224]. As seen in the Elmos data, the SET direction shows a lower readout current, therefore the difference is likely to be caused by the ReRAM asymmetry. It has already been shown that the absolute difference in currents between HRS and LRS directly impacts the reliable state distinction. Due to the higher LRS currents, the RESET direction might be more favorable.

In order to further investigate this asymmetry, a similar measurement is conducted for the MAD-200. Fig. 3.19 depicted the measured currents for 250 consecutive readout operations in 8 different devices for both polarities. In contrast to the Elmos data, no clear differences between the readout directions is visible which leads to the conclusion that the 1T-1R devices in the MAD-200 setup exhibit no pronounced asymmetry while the Elmos devices are asymmetric. A possible explanation for this difference is the exploit material stack in each setup: HfO_2-based devices in MAD-200 and TaO_x-based devices in Elmos. Furthermore, no progressive read disturb is observed in either direction in the MAD-200 data. Based on the findings, the readout methodology for the Elmos setup is selected as $V_{read} = 0.2\,V$ applied in the RESET direction to ensure reproducible current results. For the MAD-200



(a)



(b)

Figure 3.17: CDF plots for the experimental readout current data of (a) LRS and (b) HRS as presented in Fig. 3.15. Each distribution contains the 144 currents from the n-th readout across 12 devices and 12 cycles.

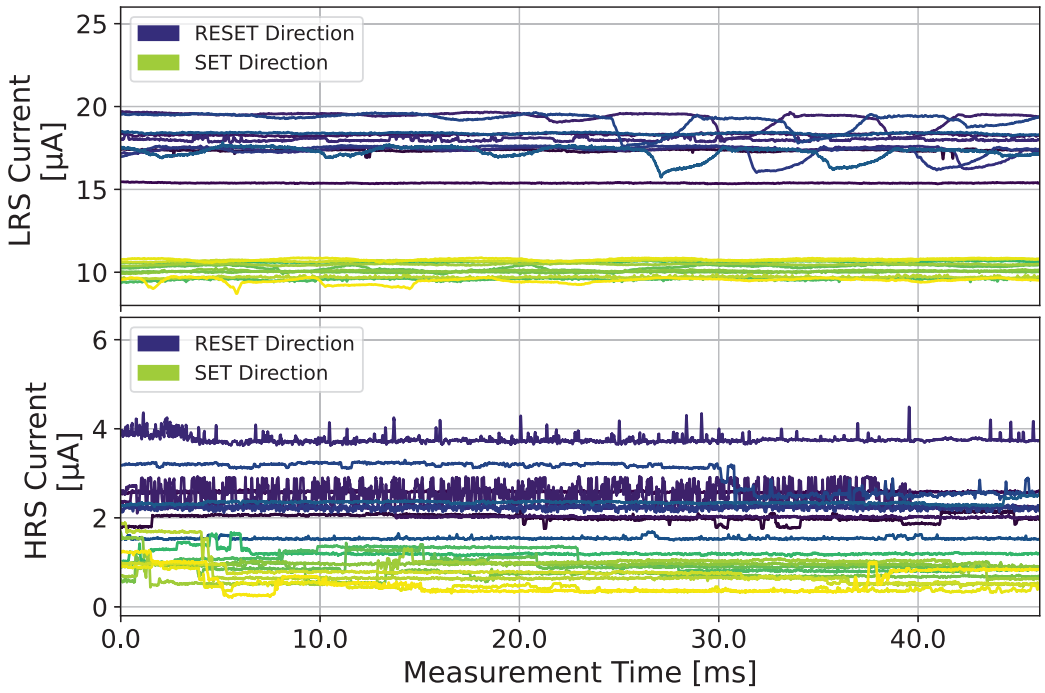


Figure 3.18: Experimental LRS (top) and HRS (bottom) time-current data for readout measurement in SET and RESET direction with the Elmos setup. Each state was cycled 10 times and readout with a pulse duration of 500 x 100 µs, $V_G = 6\,V$ and $V_{read} = 0.2\,V$.

setup, the original methodology, $V_{read} = 0.25\,V$ applied in the RESET direction, is still sufficient as it will not cause any read disturb.



Figure 3.19: Experimental LRS (top) and HRS (bottom) current data of 250 consecutive readout operations for 8 different devices in SET and RESET direction. The readout is performed with $V_{\mathrm{G}} = 5\,\mathrm{V}$ and $V_{\mathrm{read}} = 0.25\,\mathrm{V}$.

3.2.3 Two Input Operations

The single readout operation can be extended to a Two Input readout-based logic function by activating two neighboring memristive devices and measuring their combined readout current. The activation is conducted by applying a V_{read} and V_{G} at both devices at the same time. In this context, *neighboring* refers to the fact that both activated cells share at least a common Signal-Line (SL) or Bit-Line (BL). The current at the common line then will be the sum of both individual 1T-1R currents. In case of the binary individual 1T-1R states, HRS and LRS, 4 different state combinations are possible: HRS-HRS (00), HRS-LRS (01), LRS-HRS (10) and LRS-LRS (11). While this assignment is exploited in the following, it is an arbitrary way to assign binary value to the analog resistive states and LRS-LRS could also be denoted as '00'.

If a single sense amplifier is used, two primitive logic gates are possible: OR, by adjusting the reference current above "00" and below "01/10", and a logic AND, by setting it between "01/10" and "11". If the sensing is inverted, also the operations NAND and NOR are possible. By adding a second sense-amplifier, the individual results of AND (NAND) and OR (NOR) can be combined which yields to an XOR (XNOR). The corresponding truth tables are displayed in Table 3.5. This concept for logic evaluation is known in literature as *Scouting logic* based on a publication of Xie et al. [151] from 2017. However in the same year, a similar approach was demonstrated by a separated group with integrated HfO_x 1T-1R structures [203]. Therefore in the following the neutral term *Two-Input non-stateful logic* is used.

Table 3.5: Logic truth table for the Two-Input logic functions OR, AND and XOR

A	**B**	**ReRAM representation**	A OR B ($A + B$)	A AND B ($A * B$)	A XOR B ($A \oplus B$)
0	0	HRS-HRS	0	0	0
1	0	LRS-HRS	1	0	1
0	1	HRS-LRS	1	0	1
1	1	LRS-LRS	1	1	0

3.2.3.1 Characterization of Two Device Operation Accuracies

The readout currents for the four input combinations, HRS-HRS (00), HRS-LRS (01), LRS-HRS (10) and LRS-LRS (11) are measured and simulated to receive an initial impression of the Two-Input operations. The measurement results are displayed in Fig. 3.20. Here, each of the four input combinations is programmed separately in 256 pairs of 2 1T-1R devices in the MAD-200 setup.

The first conclusion is that the input combinations 10 (LRS-HRS) and 01 (HRS-LRS) lead to similar current values and nearly identical distributions. This can explained by the commutative property of the current sum, $I_{\text{total}} = I_1 + I_2$. If a position-dependence is neglected at this point, the expression will yield to $I_{\text{HRS}} + I_{\text{LRS}} = I_{\text{LRS}} + I_{\text{HRS}}$. The slight deviation of both distributions either results from the mentioned position dependency or the device variability. The inseparability might seem obvious, yet it is still important for the executable logic functions. A Two-Input boolean function transforms two inputs, A and B, to one output value, C, $f : A, B \rightarrow C$. If the value of C does not depend on the order of A and B,

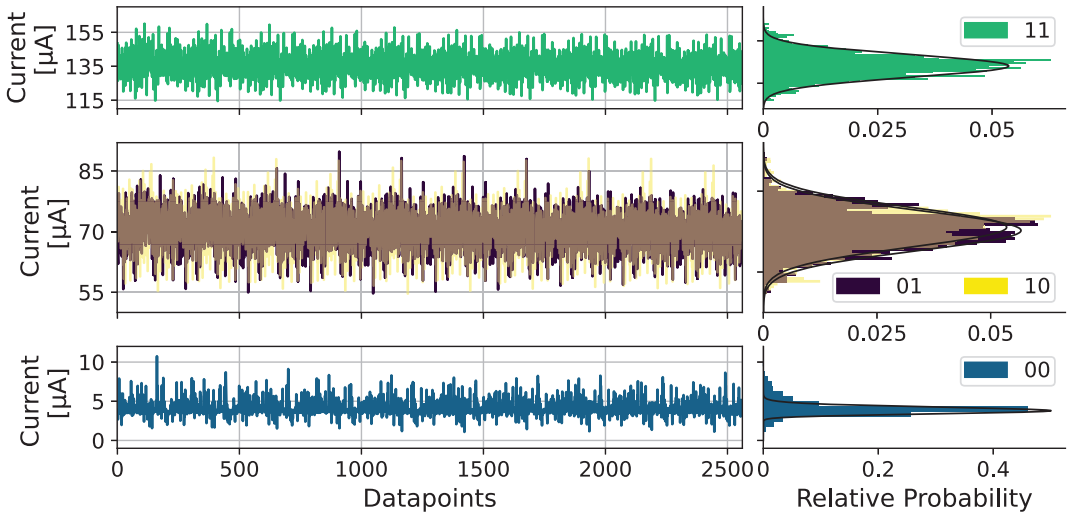


Figure 3.20: Experimental current means of a two 1T-1R device parallel readout with corresponding distribution histograms for the input combinations HRS-HRS (00) (bottom), HRS-LRS (01) and LRS-HRS (10) (middle) and LRS-LRS (11) (top). Each input combination was programmed into 256 device pairs in the MAD-200 setup. After programming, each pair is readout 10 times with a single read of 200 ms at $V_{\mathrm{read}} = 0.25\,\mathrm{V}$ and $V_{\mathrm{G}} = 5\,\mathrm{V}$.

i.e., A and B are interchangeable, the function is called symmetric. As 01 and 10 are inseparable, only symmetric logic operations are executable within this readout-based operation framework. These are for instance AND, OR, XOR. However asymmetric logic gates, such as Implication (IMP) or Converse (CON), are not possible.

The simulations are performed again with the setup described in Table 2.1 and Table 2.2. At this stage, no parasitic elements are included in the simulations, therefore the distributions for 01 and 10 are identical and combined to one distribution. The resulting currents are displayed in Fig. 3.21. Compared to the experimental results, the simulated distributions are notably broader. This overestimation of variability has already been discussed for the single device readouts. Here again the simulations act as worst-case analysis.

Based on the fitted distributions, the corresponding intersection values and error probabilities for the logic AND, OR and XOR are calculated and displayed in Table 3.6 for the measurements and Table 3.7 for the simulations. Multiple state combinations have to be considered for each logic function. As example, a logic OR will result in a false '1' if the 00-distribution is larger than 01, 10 or 11. Vice versa in a logic AND, a false '0' occurs if the 11-distribution is lower than 00, 01 or 10.

Two interesting features are visible in Table 3.6. Firstly, the distributions of 01 and 10 yield to comparable intersections and error probabilities. This is to be expected based on their nearly inseparable distributions. Secondly, for both AND and OR the probability of a false classification is highest if one of the mixed input combinations is involved. In Fig. 3.20 it is visible that the distributions of 10/01 are nearly in the middle between both edge cases, 00 and 11, and for both, AND and OR, one of these edges has to be separated from the other distributions. While it is still possible that 00 and 11 distributions overlap, it is more likely that they intersect with 01/10 due to the proximity.

The most important finding in the error probabilities is the significant asymmetry between the logic operation AND and OR. In both, experiment and simulation, the error probability for AND operations is several orders of magnitude higher than for OR. To further investigate the cause of this asymmetry,

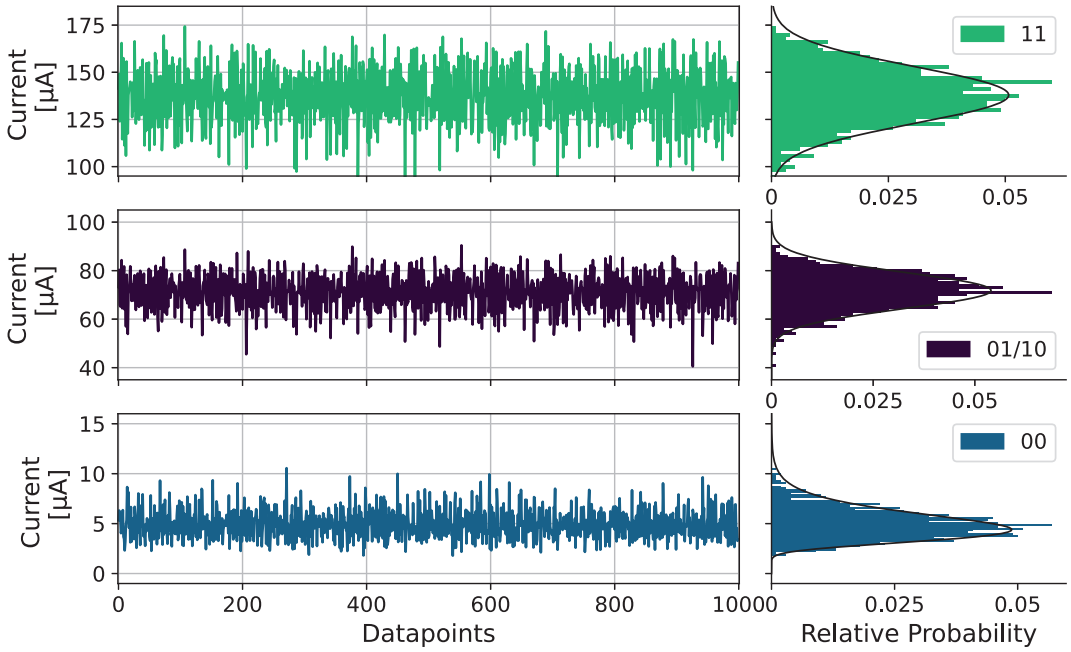


Figure 3.21: Current means of a simulated two 1T-1R device parallel readout with corresponding distribution histograms for the input combinations HRS-HRS (00) (bottom), HRS-LRS (01)/LRS-HRS (10) (middle) and LRS-LRS (11) (top). As the input combinations HRS-LRS (01) and LRS-HRS (10) are identical in the simulations, only one of them, HRS-LRS (01), is simulated. Displayed are 1 000 simulated data points with the variability parameters as described in Table 2.1 and Table 2.2. For all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed.

Table 3.6: Calculated error probabilities for the experimental Two Input non-stateful logic OR, AND and XOR. The calculation is performed based on the data displayed in Fig. 3.20.

Distributions	Intersection [µA]	Probability 'False 1'	Probability 'False 0'
OR			
00 - 01	13.99	7.19×10^{-31}	1.52×10^{-30}
00 - 10	14.24	1.14×10^{-31}	3.08×10^{-31}
00 - 11	25.07	1.10×10^{-62}	1.43×10^{-62}
AND			
00 - 11	25.07	1.1×10^{-62}	1.7×10^{-62}
01 - 11	98.03	1.34×10^{-8}	9.75×10^{-9}
10 - 11	98.42	2.00×10^{-8}	1.37×10^{-8}
XOR			
00 - (01 and 10) - 11		9.8×10^{-9}	9.3×10^{-9}

Table 3.7: Calculated error probabilities for the simulated Two Input non-stateful logic OR, AND and XOR. The calculation is performed based on the data displayed in Fig. 3.21. As previously mentioned the simulations for input combinations 01 and 10 are identical in the simulation, resulting in the same error probabilities being displayed here.

Distributions	Intersection [μA]	Probability 'False 1'	Probability 'False 0'
OR			
00 - 01	28.0	1.48×10^{-8}	1.92×10^{-9}
00 - 10	28.0	1.48×10^{-8}	1.92×10^{-9}
00 - 11	36.9	7.02×10^{-11}	7.22×10^{-12}
AND			
00 - 11	36.9	7.02×10^{-11}	7.22×10^{-12}
01 - 11	93.9	2.15×10^{-3}	1.62×10^{-3}
10 - 11	93.9	2.15×10^{-3}	1.62×10^{-3}
XOR			
00 - (01 and 10) - 11		1.5×10^{-3}	1.62×10^{-3}

Table 3.8: Maximum error probabilities for the experimental and simulated non-stateful Two Input operations OR, AND and XOR according to the previous tables Table 3.6 and Table 3.7.

Measurement			
Logic Function	Threshold [μA]	Probability 'False 1'	Probability 'False 0'
OR	14	6.68×10^{-31}	1.87×10^{-30}
AND	98	3.24×10^{-8}	9.5×10^{-9}
XOR	14 / 98	9.5×10^{-9}	2.25×10^{-8}
Simulation			
Logic Function	Threshold [μA]	Probability 'False 1'	Probability 'False 0'
OR	28	1.48×10^{-8}	1.92×10^{-9}
AND	94	2.05×10^{-3}	1.67×10^{-3}
XOR	28 / 94	1.66×10^{-3}	1.43×10^{-3}

the absolute current spacing is investigated in the following.

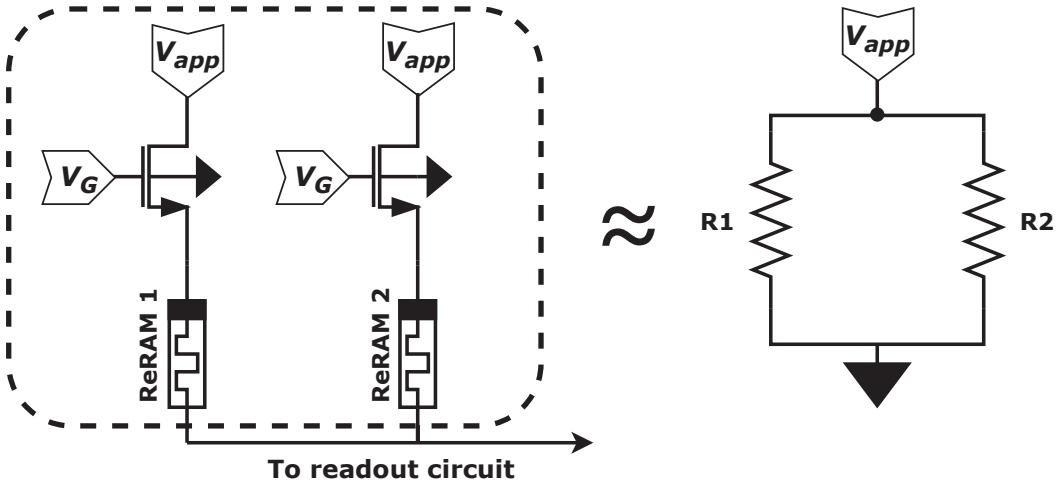


Figure 3.22: The equivalent circuit diagram for a non-stateful two-device readout. Each 1T-1R bitcell is estimated by an equivalent ohmic resistance under the assumption of constant V_{read} and V_{G}.

Fig. 3.22 shows the simplified equivalent circuit of a Two-Input non-stateful readout which consists of two parallel resistances. Each resistance is formed by the serial combination of a MOSFET and ReRAM as explained above. This simplification is valid under the assumption of a constant $V_{\text{G}} > V_{\text{Threshold}}$ and a constant V_{read}. The resistances are either in low-resistance state, R_{LRS}, or in high resistive state, $r \times R_{\text{LRS}}$ with r being the resistance ratio $r = R_{\text{HRS}}/R_{\text{LRS}}$. Based on this circuit, the equivalent readout current is calculated for the three binary cases of HRS-HRS, LRS-HRS / HRS-LRS, and LRS-LRS. Table 3.9 depicts the calculated currents for the three cases in dependence of V_{read}, R_{LRS} and r.

Case	Current
HRS−HRS	$V_{\text{read}} \times \frac{2}{r} \times \frac{1}{R_{\text{LRS}}}$
HRS−LRS LRS−HRS	$V_{\text{read}} \times \frac{1+r}{r} \times \frac{1}{R_{\text{LRS}}}$
LRS−LRS	$2 \times V_{\text{read}} \times \frac{1}{R_{\text{LRS}}}$

Table 3.9: Theoretical readout current for the equivalent circuit in Fig. 3.22. The circuit consists of two parallel resistances with an applied voltage, V_{read}. Each resistance is either in low resistive state, $R_{\text{LRS}}/$ or in high resistive state, $r \times R_{\text{LRS}}/$ with $r = R_{\text{HRS}}/R_{\text{LRS}}/$.

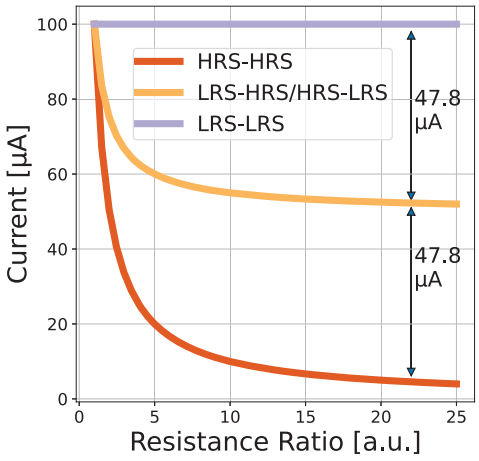


Figure 3.23: Ideal calculated readout currents for a two-device readout based on the equivalent circuit in Fig. 3.22 and Table 3.9. For the calculations the values LRS = 5 kΩ, HRS = $r\times$ LRS and V_{read} = 0.25 V are assumed.

Fig. 3.23 illustrates that with ideal ohmic resistances the spacing of the readout currents between the three cases is symmetric. A similar observation is found for the current distribution from the simulated

data in Fig. 3.21 and the measured data in Fig. 3.20. The top part of Fig. 3.24 depicts the mean value of each distribution. While the mean values are not perfectly symmetric with a deviation around $3\,\%$ this does not account for the significant asymmetry observed in the error probability, as indicated by the framework.

To approach the cause of this asymmetry, not only the distribution means, μ, but also their standard deviations, σ, are considered. This is crucial as the mean value alone does not contain any information about the current spread. Therefore, the lower part of Fig. 3.24 depicts the $\mu \pm 3-\sigma$ ranges of each distribution. It can be seen that for the experimental measurements now the spacing from LRS-HRS / HRS-LRS to HRS-HRS is approximately twice as large as for LRS-LRS to LRS-HRS / HRS-LRS. As previously mentioned, the simulations cover a higher variability as observed experimentally. Here, the separation collapses as the $3-\sigma$ ranges of LRS-HRS / HRS-LRS and LRS-LRS are already overlapping.

The increased spread for the currents of LRS-HRS, HRS-LRS and LRS-LRS compared to the HRS-HRS currents can be explained by the internal ReRAM variability. The single-device investigation in Fig. 3.6 showed that the relative variability, σ/μ of the high-resisitive state is higher than for the low resistive state, $(\sigma/\mu)_{\mathrm{HRS}} = 0.4; (\sigma/\mu)_{\mathrm{LRS}} = 0.1$. However, the low resistive state has a higher impact on the absolute readout current. Therefore, any variability in the low resistive state will lead to a significant increase in the readout current variability. This effect can be illustrated with ohmic resistances. If a $100\,\mathrm{k}\Omega$ resistor demonstrates a $20\,\%$ variation, it leads to absolute currents between $2.083\,\mu\mathrm{A}$ and $3.125\,\mu\mathrm{A}$ at a readout voltage $V_{\mathrm{read}} = 0.25\,\mathrm{V}$. The same variation for a $5\,\mathrm{k}\Omega$ resistor results in currents between $41.66\,\mu\mathrm{A}$ and $62.5\,\mu\mathrm{A}$. If several low resistive states are included, i.e., for the combination LRS-LRS, this trend accumulates to an even higher absolute variability.

The conclusion from this analysis is that without any further optimization, applications which are based on the demonstrated two-device non-stateful logic should favor the use of logic OR gates. As here the distinction is made between HRS-HRS (0) and LRS-HRS/HRS-LRS and LRS-LRS (1), the operation is more reliable than a logic AND, i.e. distinguishing between HRS-HRS/LRS-HRS and HRS-LRS (0)and LRS-LRS (1). Furthermore as the non-stateful XOR is implemented as the combination of AND and OR, it inherits the variability issues that affect the individual operations. For scenarios which require the execution of a logic AND, a potential strategy is to replace the AND operation with a concatenation of logic OR and NOT gates. As the logic NOR gate as concatenation of OR and NOT is universal, all boolean logic gates can be represented based on these two operations. This substitution can help to reduce error probabilities but also leads to the need for more computing cycles. This comes in hand with an increased overall energy consumption. Besides this, the concatenation of single-gates suffers additionally from the issue that individual error probabilities propagate. Therefore, the execution of one logic AND might outperform the concatenation of several logic OR gates, even if its single error probability is higher. The most straight-forward way to achieve more balanced error probabilities is to reduce the LRS variability for instance by program-verify algorithms.

3.2.3.2 Influence of the Operation Voltage

Based upon the results in Section 3.2.2.2, the voltage dependence of the two-device operations is investigated. The experimental procedure consists of the initial programming of two devices into one of four possible combinations. In order to minimize any parasitical influence, two adjacent devices were selected in close proximity to the line access. The programming is followed by a systematic sweep of V_{G} and V_{read}. Each voltage combination was readout 250 times. Between the individual voltage settings, the devices were not re-cycled, only when a new voltage combination was programmed. The results are depicted in Fig. 3.25.

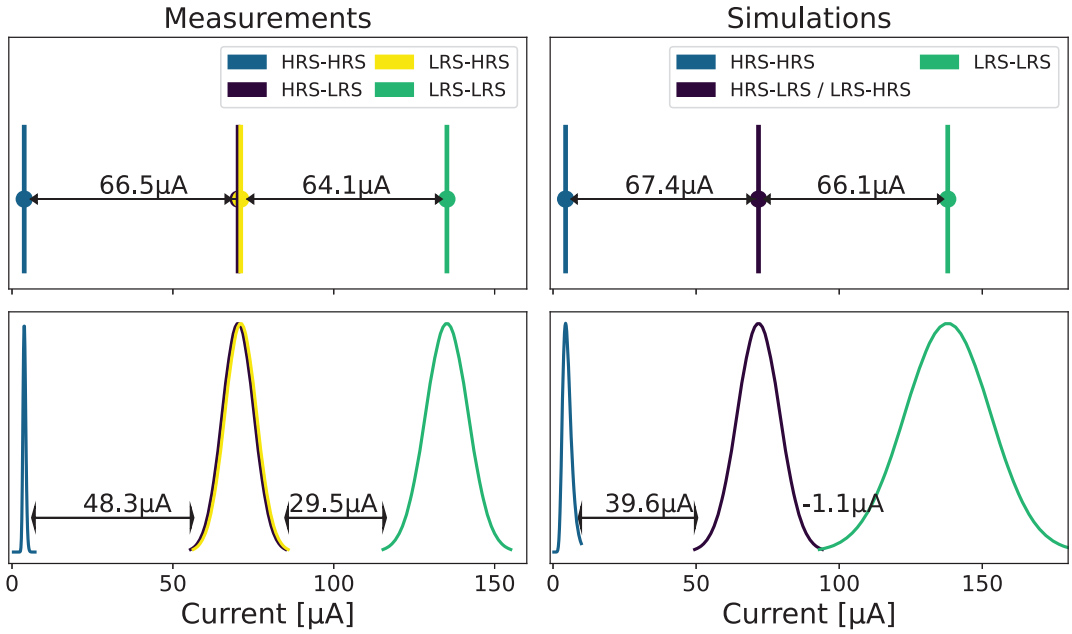


Figure 3.24: Readout current spacing of a two-device non-stateful operation for the measured (left) and simulated (right) current distributions, taken from Fig. 3.20 and Fig. 3.21. On the top, the respective distribution mean μ is depicted; on the bottom, the range of 3 standard deviations, σ as $[\mu - 3\sigma; \mu + 3\sigma]$. Note that the y-axis denotes an arbitrary scaling.

For the 00 output it is interesting that the readout current increases only slightly between $V_{\text{read}} = 0.2\,\text{V}$ and $V_{\text{read}} = 0.3\,\text{V}$. The high-resisitive state faces a strong non-linearity with an declining voltage-resistance relationship as shown in Fig. 3.11(a). Therefore the increase in currents does not scale linear with the voltage. Yet, the observed currents for $V_{\text{read}} = 0.2\,\text{V}$ and $V_{\text{read}} = 0.3\,\text{V}$ are so closely aligned that is questionable whether only the non-linearity can cause this. The gate voltage on the other hand has no significant impact on the readout current in the HRS-HRS configuration. This was already observed for the single device HRS readout in Fig. 3.9(b). The 1T-1R current is determined by the serial combination of ReRAM resistance and the transistor channel resistance. The higher one of both will mainly determine the current limitation. Therefore if the ReRAM devices are in HRS, a significantly lower V_{G} is required in order turn on the transistor. e.g., reduce its channel resistance below R_{HRS}.

For the states 01, 10, and 11, several interesting observations are noted. Firstly, the absolute current values for 01 and 10 show a noticeable difference although a similar behavior is expected. This indicates a strong device-to-device or cycle-to-cycle variability between both programmed combinations. Unfortunately due to the lack of measurement data, this feature cannot be resolved based on this measurement sample. Secondly, in contrast to the 00 combination, the inputs 01, 10 and 11 demonstrate a significant gate voltage dependency. In the range of $V_{\text{G}} \leq 2\,\text{V}$, the current increases significantly which aligns with the single-device behavior in Fig. 3.9(a). Afterwards, the currents saturate for $V_{\text{G}} \geq 3\,\text{V}$ as the transistor operates in saturation mode. This plateau is most pronounced for the $V_{\text{read}} = 0.2\,\text{V}$ data set.

For the two samples $V_{\text{read}} = 0.3\,\text{V}$ and $0.4\,\text{V}$, another noteworthy phenomenon occurs. In both data sets, the readout currents decrease substantially for gate voltages of $V_{\text{G}} \geq 3\,\text{V}$. This contradicts the behavior of the single-device in Fig. 3.9(a). Here, the currents at $V_{\text{read}} = 0.3\,\text{V}$ increased with higher

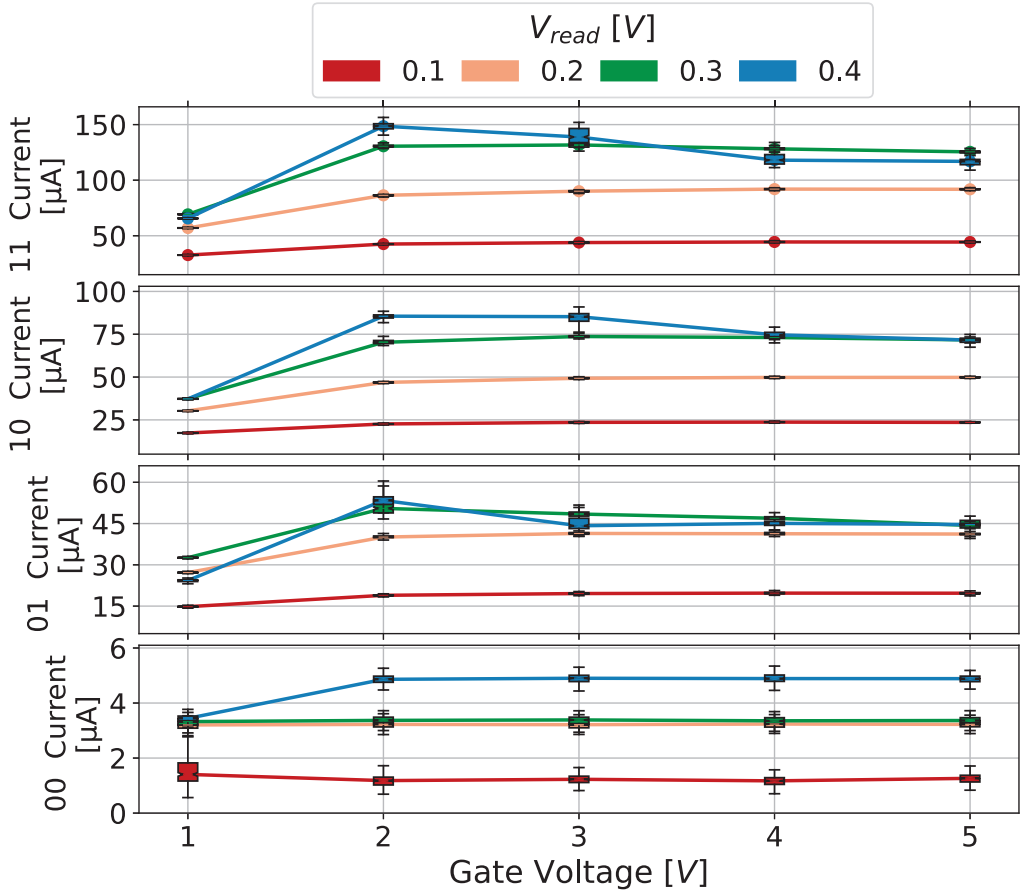


Figure 3.25: Experimental dependence of the readout currents for a two-device operation on the applied read voltage, V_{read}, and transistor gate voltage, V_{G}. Measured are the input combinations HRS-HRS (00) (bottom), HRS-LRS (01) (second from bottom), LRS-HRS (10) (second from top) and LRS-LRS (11) (top). Each boxplot contains the data of 1 programmed device combination which is readout 250 times.

gate voltages. Previously, it had been shown that a larger current margin is beneficial for reducing error probabilities. Therefore, decreasing currents at higher voltages mark a severe issue.

The observed current reduction could be linked to the methodology of the measurement in which the devices are not re-programmed after each test. During an individual testing procedure the voltage pulses are applied in an ascending sequence, starting at $V_{\text{read}} = 0.1\,\text{V}$ and $V_{\text{G}} = 1\,\text{V}$. This pulse is applied 250 times, then V_{G} is incremented up to $5\,\text{V}$. This sequence is repeated for the other readout voltages. As a consequence, each device undergoes a total of $5 \times 4 \times 250 = 5\,000$ readout pulses throughout the test and 250 pulses with the maximum combination of $V_{\text{read}} = 0.4\,\text{V}$ and $V_{\text{G}} = 5\,\text{V}$. As the readout operations are conducted in the RESET direction, devices in the low resistive state might experience a resistance drift towards lower current values over the course of the measurement. To investigate this hypothesis, a subset of the previous data is plotted in the temporal order of the measurements in Fig. 3.26. Included are solely the combinations 01, 10 and 11, starting at $V_{\text{read}} = 0.2\,\text{V}$ and $V_{\text{G}} = 1\,\text{V}$.

For all readout voltages, a discrete current step between $V_{\text{G}} = 1\,\text{V}$ and $V_{\text{G}} = 2\,\text{V}$ is visible. Another minor step occurs at the transition from $V_{\text{G}} = 2\,\text{V}$ to $V_{\text{G}} = 3\,\text{V}$, but only for the combinations 10 and 11. This behavior corresponds to the discussed feature of the transistor channel resistance. With each gate voltage step from $1\,\text{V}$ to $2\,\text{V}$ to $3\,\text{V}$, the channel resistance will reduce and with this its impact on the

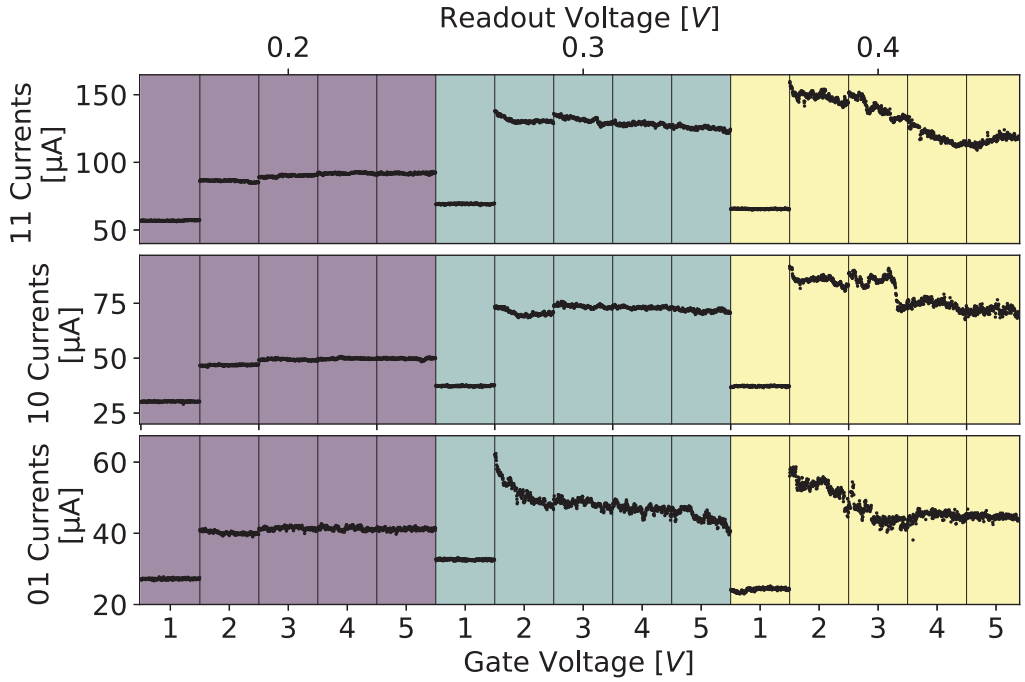


Figure 3.26: Experimental temporal evolution of the readout currents during the two-device voltage dependency measurement depicted in Fig. 3.25. The labeled pulse combinations (V_{read},V_{G}) are applied sequentially from left to right, each tested 250 times. As the measured HRS-HRS (00) data showed no disturbance in Fig. 3.25, it is left out here for better visualization.

current limitation decreases. At gate voltages of $V_{\text{G}} \geq 3\,\text{V}$ the transistor operates in full saturation and all combinations demonstrate stable output currents at least for $V_{\text{read}} = 0.2\,\text{V}$.

At $V_{\text{read}} = 0.3\,\text{V}$ the gradual current decrease is already visible, most prominently in the 01 current data. As the readout voltage is increased further, the reduction become more severe also for these combinations. This decrease shows no uniform behavior. While in 01 and 11 it appears to be gradual over the course of several dozens readouts, the combination 10 demonstrates an abrupt current change. These spontaneous current jumps resemble the observe read instability for the high resistive state. However, in this model the LRS is describe as stable [81]. Therefore both, the gradual decrease and the abrupt jump, might indicate a physical change in ReRAM state.

In the case of an actual resistance change, this modified state will appear continuously across the read voltage transitions. Such a transition is difficult to extract from the current data as the current also changes due to the increased read voltage. A better way to identify this is to calculate the equivalent resistance as $V_{\text{read}}/I_{\text{read}}$ which is depicted in Fig. 3.27. In this representation the gradual increase in the 1T-1R resistance is observable, most prominently for the combination 01. Still there are several artifacts of unclear origin within the data. Firstly, the 11 resistances decreases again towards the end of the measurement. Secondly, the data samples deviate in their voltage-dependency. While the combinations 10 and 11 share a similar response at $V_{\text{read}} = 0.3\,\text{V}$, 01 behaves very different. However at $V_{\text{read}} = 0.4\,\text{V}$, an abrupt current jump is only observed for the combination 10. A possible explanation for this is the low sample size and the missing programming variability in particular. As only one pair of devices is programmed, no device-to-device or transistor variability can be resolved with this data. However, a shift in the readout currents is observable with higher readout voltages, $V_{\text{read}} \geq 0.3\,\text{V}$ especially in

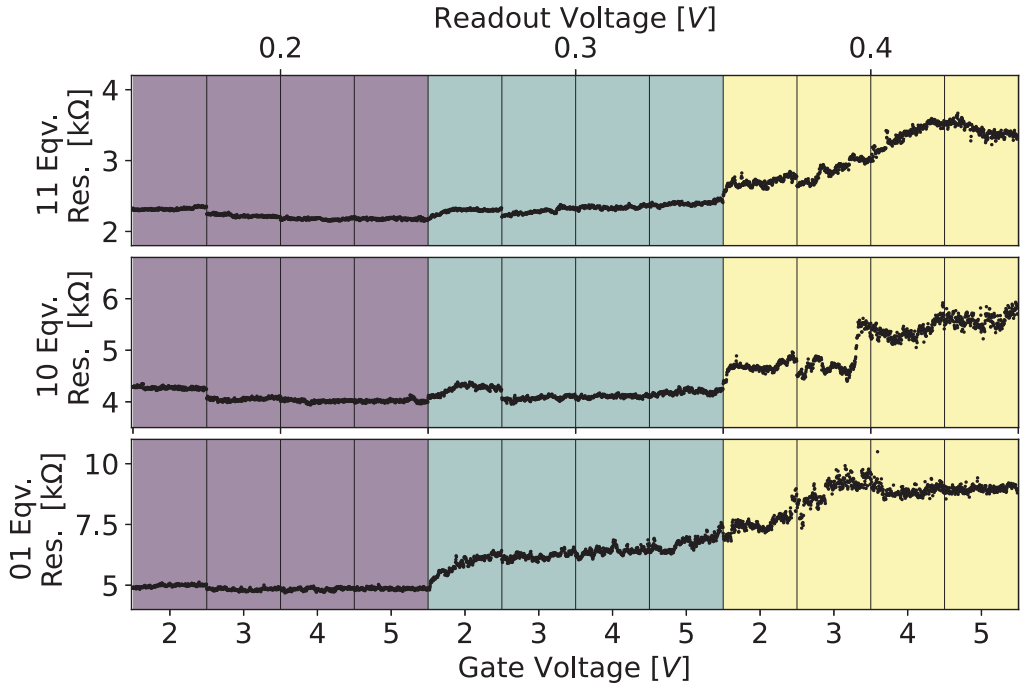


Figure 3.27: Experimental temporal evolution of the equivalent resistance, calculated as $V_{\text{read}}/I_{\text{read}}$, during the voltage dependent two-device measurement as depicted in Fig. 3.25. The labeled pulse combinations (V_{read},V_{G}) are applied sequentially from left to right, each tested 250 times. Please note that for the sake of a better visualization, the pulse combination for $V_{\text{G}} = 1\,\text{V}$ are not depicted. They were performed at the positions seen in Fig. 3.26. Furthermore as the measured HRS-HRS (00) data showed no disturbance in Fig. 3.25, it is also left out here for better visualization.

comparison to the case of $V_{\text{read}} = 0.2\,\text{V}$. This occurs for a relatively low quantity of $5\,000$ readout operations. In order to guarantee a state-conservation and reliable non-stateful logic operations, V_{read} should be limited to $\leq 0.3\,\text{V}$ or even $\leq 0.25\,\text{V}$, applied in a RESET polarity. On the other hand, it has to be ensured that the gate voltage, V_{G}, is set high enough to reach the transistor saturation.

In Fig. 3.28 an attempt was made to perform the distribution fitting and probability calculation as previously done for the single device data. However due to the limited sample size, the fitting resulted in very high residuals and therefore the meaningful calculation of error probabilities was not possible. The visible shift in the distribution means for higher readout voltages can only support the critical voltage-dependency in a qualitative manner. Instead the voltage-dependent two-device readout operation was simulated in the variability-aware framework. For each readout voltage $V_{\text{read}} \in [0.2, 0.3, 0.4]\,\text{V}$, $1\,000$ operations are simulated for both voltages polarities, SET and RESET direction.

The simulation results are depicted in Fig. 3.29. The currents in the RESET direction are consistently higher than those in the SET direction at the same voltages which is consistent with the observations from single-device measurements in Fig. 3.18. This behavior is caused by the inherent asymmetry of the VCM devices. Interestingly in the SET direction, the currents increase progressively for the input combinations which include a LRS device, 01/10, and 11, but not for 00. Therefore it cannot be explained by a full unintended input switching. Instead, the LRS devices in these combinations are initially in a comparably high low-resistance state (soft LRS). During the consecutive readouts at elevated voltages, they undergo an additional SET process which shifts them into the maximum

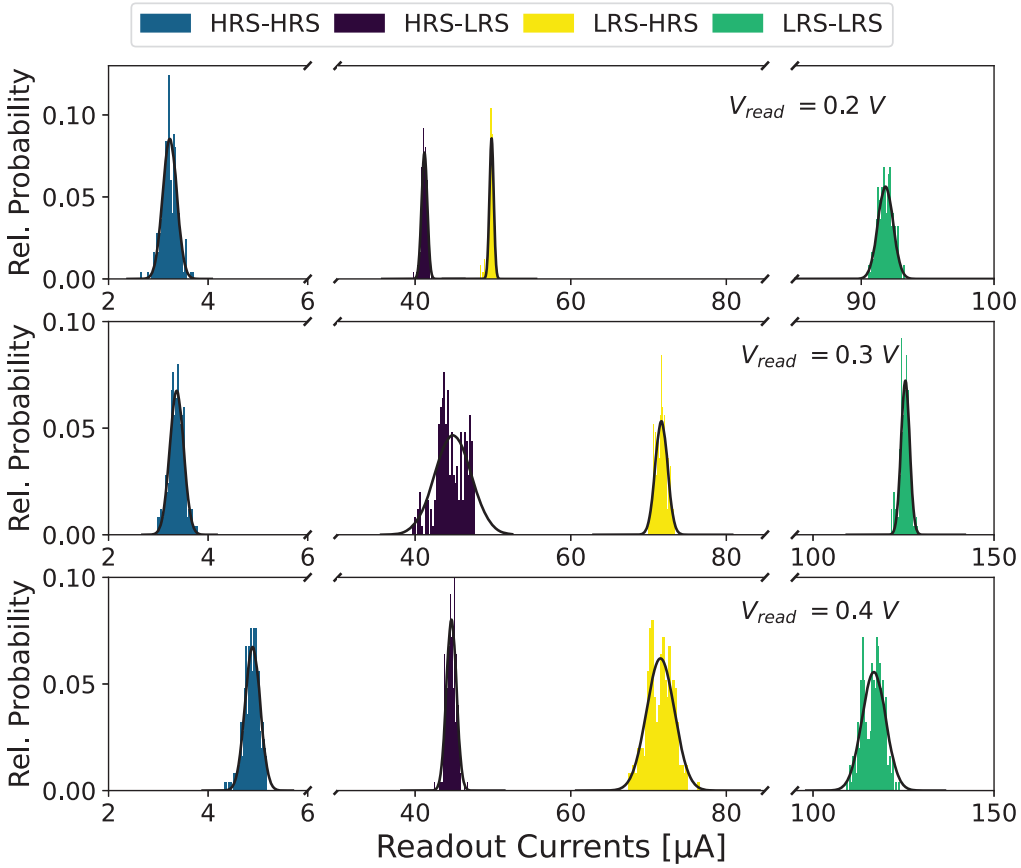


Figure 3.28: Experimental two-device readout current distributions and fitted distributions for increasing V_{read} values at a constant $V_{\text{G}} = 5\,\text{V}$. Each distribution contains 250 measured data points from 1 device.

achievable low-resistance state (hard LRS).

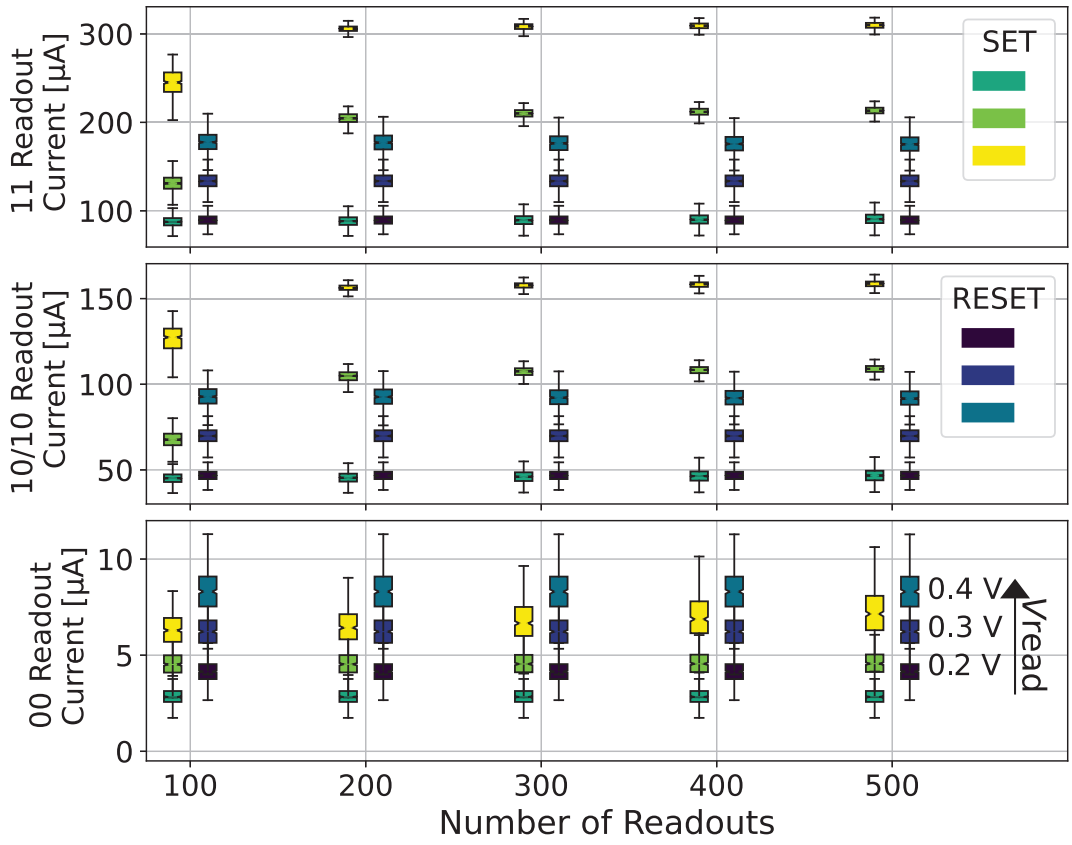


Figure 3.29: Boxplots of the simulated two-device readout current during 500 readout operations for increasing V_{read} values at a constant $V_{\text{G}} = 5\,\text{V}$. The readout is simulated 500 times for both polarities, in SET direction (left boxes) and RESET directions (right boxes). For better visualization, only every 100-th current is depicted. Simulated are the input combinations HRS-HRS (00) (bottom), HRS-LRS (01)/LRS-HRS (10) (middle) and LRS-LRS (11) (top). As the input combinations HRS-LRS (01) and LRS-HRS (10) are identical in the simulation, only one of them, HRS-LRS (01), was simulated. The simulations are conducted with the variability parameters as described in Table 2.1 and Table 2.2 and for all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed.

3.2.3.3 Influence of the Operation Position

For the two-device readout, two types of positional influences are possible: the *absolute position* of adjacent devices in the crossbar, for example the pairs (1,2) compared to (511,512), and their *relative position*, for example (1,2) compared to (1,512). Therefore, two different data sets are utilized.

The first one consists of 256 adjacent device pairs, i.e., (1,2), . . . ,(511,512). Each pair is programmed into all four combinations and read out 100 times to account for any read instability. The results are illustrated in Fig. 3.30. For a better visualization, the device pairs are grouped into 8 packs of 32 device pairs. From this data no clear position dependence is evident even for 512 Signal-Lines/Word-Lines. This matches to the single-device behavior in Fig. 3.14 where constant device currents were observed along 256 lines.

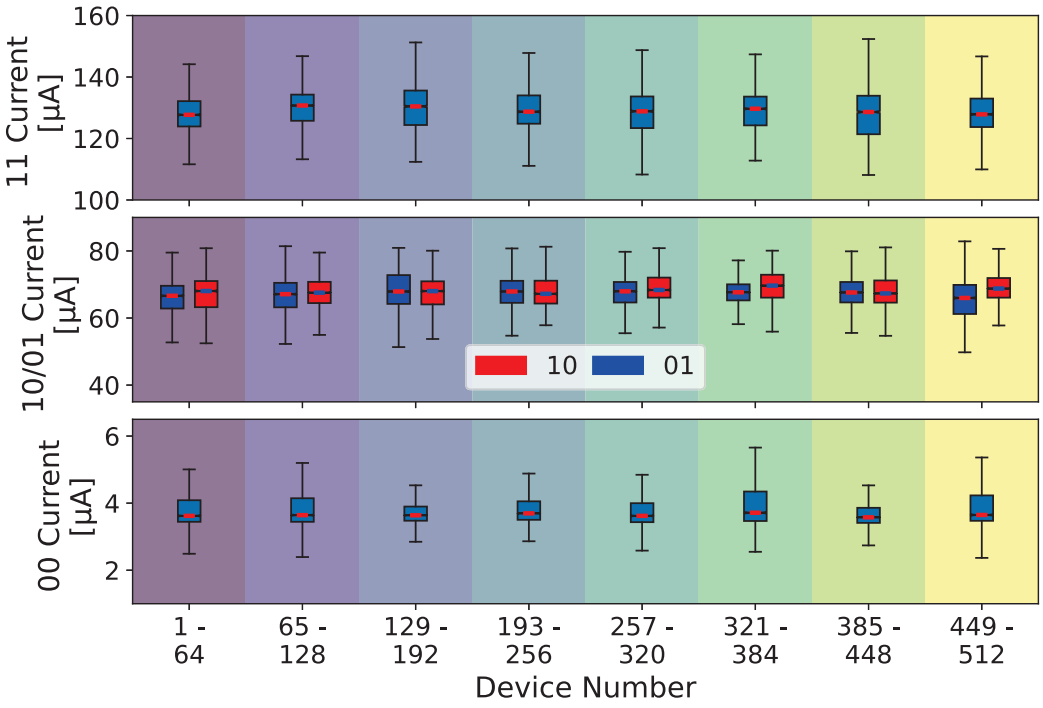


Figure 3.30: Experimental readout currents for two-device operations in a 1T-1R crossbar for SL/WL 1 to 512 at Bit-Line 16. Measured are the input combinations HRS-HRS (00) (bottom), HRS-LRS (01)/LRS-HRS (10) (middle) and LRS-LRS (11) (top). Device pairs are formed from adjacent devices in the SL/WL, e.g., SL/WL (1,2), SL/WL (3,4) and so on. The readout current is measured on the far left position (SL/WL 1). Each box contains 3 200 data points of 32 device pairs which are cycled once and readout 100 times.

In the second data set, 7 distinct device pairs are selected with varying distance across the whole 512 Signal-Lines/Word-Lines. These pairs were programmed 16 times into each input combination and then again read out 100 times. The results are illustrated in Fig. 3.31. Overall, the data shows no clear dependency on the relative device positions. However, the readout currents in the pair (127,128) are exceptionally low for the combinations 00, 01 and 11. A similar current is observed for the input combination 00 in the pair (383,384). Yet here the other combinations behave as expected. These exceptions could indicate an insufficient device programming. However, the data includes 16 individual switching steps. If only one programming is corrupted, it would results in a slight shift of the mean together with a larger whisker range. Yet for the samples above, the mean currents are shifted and the

whiskers as quartile intervals are not extended.

The behavior may be attributed to malfunctioning devices. Such a device damage can occur during fabrication or operation. The pair (127,128) shows for the 11 (LRS-LRS) combination approximately half of the expected current, around 50 µA-70 µA instead of 120 µA-150 µA. Additionally, the 01 (HRS-LRS) currents of this pair show more similarity to the 00 (HRS-HRS) currents of the other device pairs, ≤ 5 µA, as to the other 01 combinations. However, the 10 (LRS-HRS) combination of pair (127,128) matches the other pairs. This leads to the conclusion that device (128) might be *stuck-at-HRS* while device (127) operates normal. Therefore, the 11 combination behaves similar as a 10. As the 00 (HRS-HRS) current in this pair is also reduced, it indicates that the stuck-at-HRS resistance of device (128) is even lower than a regular HRS and it is likely that device (128) is fully broken. For the deviating behavior of device pair (383,384) in combination 00, there is no similar clear explanation. As the other combinations are unobtrusive, the deviations cannot be caused by a broken device or a stuck-at-HRS of both cells.

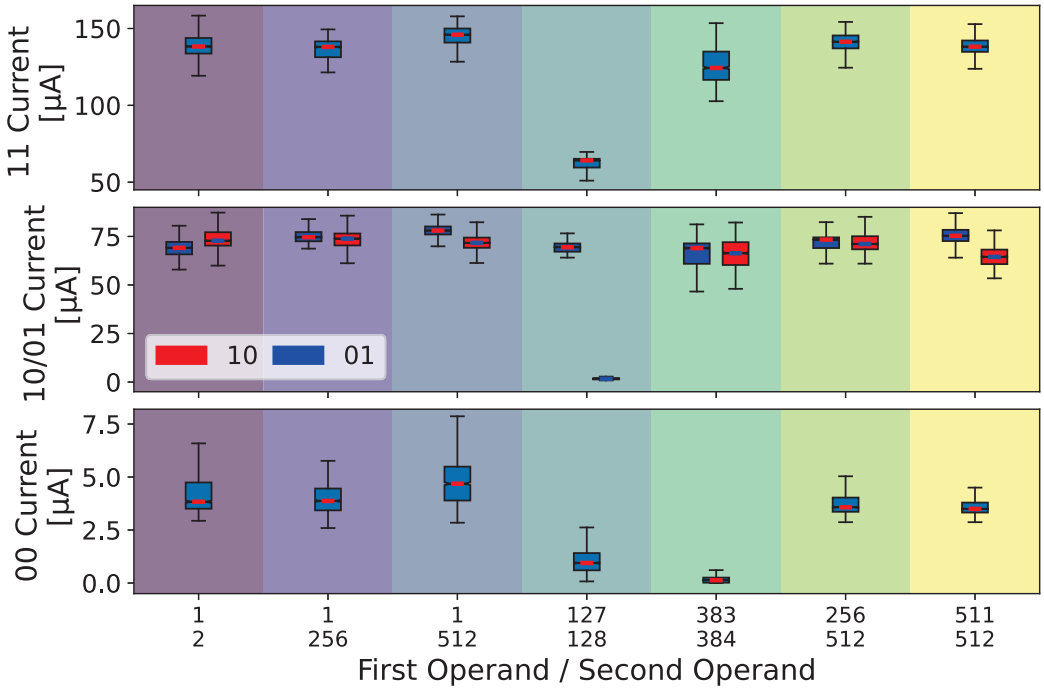


Figure 3.31: Experimental readout currents for two-device operations in a 1T-1R crossbar for selected device pairs at the labeled SL/WL positions. Measured are the input combinations HRS-HRS (00) (bottom), HRS-LRS (01)/LRS-HRS (10) (middle) and LRS-LRS (11) (top). The readout current is measured on the far left (SL/WL 1). Each box contains 1 600 data points as each input combination was programmed 16 in each device pair and readout 100 times.

The investigation confirms that the two-device logic operations are possible across a full 512 x 32 crossbar without any significant position dependency. However, the last analysis highlights the importance of uniformity in the 1T-1R cells. Non-functioning devices due to fabrication defects, aging or false operation have to be avoided. If such malfunctions occur, it is necessary to identify the corresponding devices and individually exclude them from the operation if possible. Also error correction codes (ECC) can be exploited which compensate logic errors later on.

3.2.4 Multi-bit Evaluation

Unlike the Two-Input boolean operations, a *Majority* (MAJ) operation typically requires an odd number of inputs. The MAJ gate outputs a logic 1 if and only if more than half of its inputs are 1; otherwise it outputs 0. This function can be inverted to a *Minority* (MIN) gate. While a Majority gate can operate on an arbitrary odd number of inputs, the discussion will focus on the simple scenario of a Three-Input Majority gate for demonstrative purposes.

In this case 8 input combinations are possible, 000, 001/010/100, 011/101/110 and 111, which are listed as logic truth table in Table 3.10. The output current distributions for the three permutations 001/010/100 and respectively 011/101/110 are expected to be identical. This leads to four distinct current levels which is similar to the behavior of 01/10 for the logic AND/OR. In contrast to the two-device AND or OR, the Majority gate is not widely recognized in literature and only investigated as simulation so far [160, 161]. The data and results in this chapter were presented in [204] and some of the text is taken from there.

Table 3.10: Logic truth table for the 3-Input Majority function.

A	**B**	**C**	**ReRAM representation**	MAJ $\langle A, B, C \rangle$	MIN $\langle A, B, C \rangle$
0	0	0	HRS-HRS-HRS	0	1
0	0	1	HRS-HRS-LRS	0	1
0	1	0	HRS-LRS-HRS	0	1
1	0	0	LRS-HRS-HRS	0	1
0	1	1	HRS-LRS-LRS	1	0
1	0	1	LRS-HRS-LRS	1	0
1	1	0	LRS-LRS-HRS	1	0
1	1	1	LRS-LRS-LRS	1	0

All possible input combinations are programmed 10 times into a device triplet in the MAD-200 setup and the resulting mean currents and fitted histograms are depicted in Fig. 3.32. The current distributions for the permuted combinations of 001/010/100 and respectively 011/101/110 are nearly identical. This supports the previous assumption of four distinct current levels. Although the readout currents are generally stable, some individual read failures are observed as current peaks towards 0. The recorded data for these events contains the value of exactly 0.0 A. If the peak would be related to any instability at least some noise recording is expected. However, the exact value of 0.0 suggests that its origin is within the measurement setup as the sampling step has failed.

During the Majority operation, the 8 possible combinations are assigned to the two logic levels of 0 and 1. In this classification, the threshold has to be selected in a way to reliably distinct the combinations 000, 001, 010 and 100 (logic 0) from 011, 101, 110 and 111 (logic 1). Based on the fitted current distributions in Fig. 3.32, the intersection values and error probabilities are calculated in a similar manner as before. The calculation results in Table 3.11 present several interesting observations. Firstly, the overlap probability is highest between the inner distributions, 001/010/100 and 011/101/110. This is expected as these distributions are closest to each other. Opposed to this, it is less likely for a 000 state to overlap with any of the logic 1 combinations. Similarly, a 111 state is unlikely to overlap with any of the logic 0 configurations.

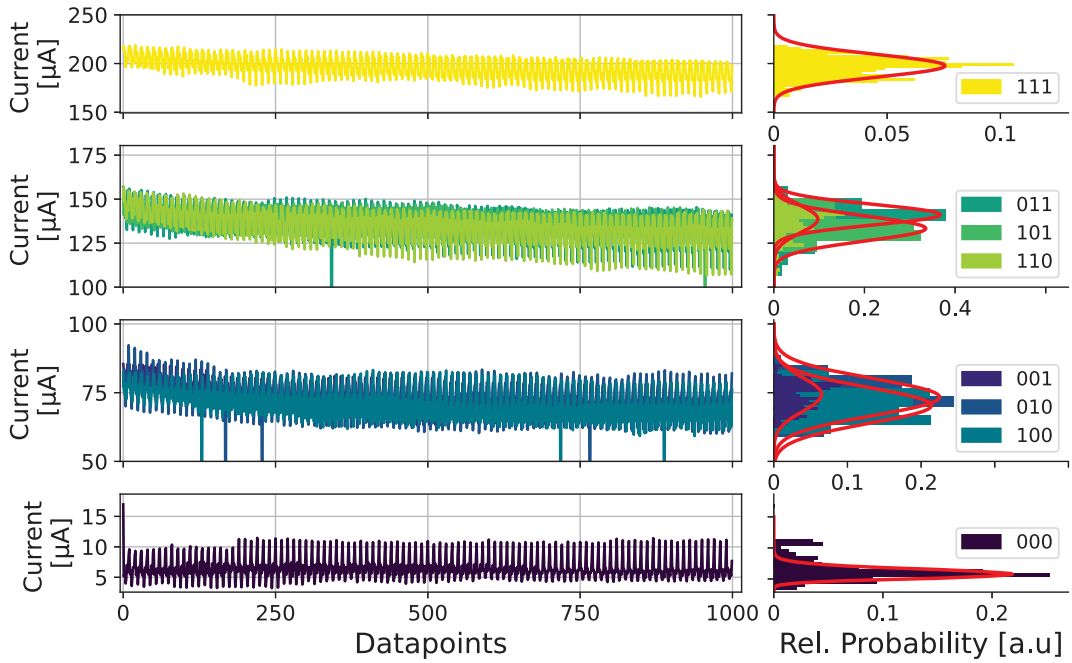


Figure 3.32: Experimental current means of a parallel 3 1T-1R device readout with corresponding distribution histograms for the input combinations HRS-HRS-HRS (000) (bottom), HRS-HRS-LRS (001) and its permutations (second from bottom), HRS-LRS-LRS (011) and its permutations (second from top), and LRS-LRS-LRS (111) (top). Each input combination was programmed 10 times into one device triplet. Following each programming, the readout is performed 100 times with a single read of 200 ms at $V_{\text{read}} = 0.25\,\text{V}$ and $V_{\text{G}} = 5\,\text{V}$.

The outermost distributions, 111 and 000, display another noteworthy asymmetry. As mentioned, for both of them it is unlikely to overlap. However, it appears that the combinations 001/010/100 are closer to 111 than 000 is to 011/101/110. This observation aligns with the previous observation from the two-device readout in Table 3.6. Also there, the inner distributions 01 and 10 demonstrated more overlap with 11 than 00. As seen in Fig. 3.24 this asymmetry was related to the absolute variation in the low resistive state and the distribution width.

Based on the values a threshold of $105\,\mu\text{A}$ is selected for the Majority gate. For this value the maximum error probability is around 1×10^{-5}. This exceeds the error probability of the measured two-device logic AND and OR by almost 3 orders of magnitude compared to Table 3.8. Yet, it is possible to implement a three-device Majority gate in resistive 1T-1R structures. The same circuitry can be also used to realize a 2-bit readout. Here, four instead of two logic states are distinguished: 000 (logic 0), 001/010/100 (logic 1), 011/101/110 (logic 2) and 111 (logic 3). This distinction requires at least a 2-Bit ADC as sensing circuitry.

The maximum error probabilities for this 2-bit classification are listed in Table 3.12. While it is possible to reliably separate the 000 combination, all other state demonstrate an increased error probability up to the order of 1×10^{-3}, i.e., 1 of 1 000 operations are corrupted. The data highlights the critical impact of the LRS variability on the operation reliability. Without any measures to control and limit the LRS variability, such a 2-Bit readout operation is not feasible. This issue will become worse with an increasing number of parallel devices, for instance in a binary vector-matrix-multiplication, where more than 4 levels have to be distinguished.

Table 3.11: Calculated error probabilities for the measured three-device non-stateful Majority (MAJ) function. The calculation is performed based on the data displayed in Fig. 3.32.

Distributions	Intersection [μA]	Probability 'False 1'	Probability 'False 0'
000 - 011	48.9	1.5×10^{-56}	1.2×10^{-57}
000 - 101	36.4	2.2×10^{-42}	2.6×10^{-43}
000 - 110	35.0	1.1×10^{-40}	4.6×10^{-41}
000 - 111	37.4	1.2×10^{-43}	6.1×10^{-44}
001 - 011	107.8	2.1×10^{-8}	3.7×10^{-9}
001 - 101	100.5	8.9×10^{-5}	1.7×10^{-6}
001 - 110	102.4	2.1×10^{-6}	1.5×10^{-6}
001 - 111	116.9	1.6×10^{-12}	1.4×10^{-12}
010 - 011	109.4	3.0×10^{-8}	1.8×10^{-8}
010 - 101	102.2	7.2×10^{-6}	4.9×10^{-6}
010 - 110	104.2	1.7×10^{-6}	4.4×10^{-6}
010 - 111	119.4	2.1×10^{-12}	6.6×10^{-12}
100 - 011	107.5	4.7×10^{-9}	2.8×10^{-9}
100 - 101	100.2	2.0×10^{-6}	1.3×10^{-6}
100 - 110	102.0	4.9×10^{-7}	1.2×10^{-6}
100 - 111	116.7	4.2×10^{-13}	1.3×10^{-12}

Table 3.12: Calculated maximum error probabilities for the experimental 2-bit readout with 4 logic states. The calculation is performed based on the data displayed in Fig. 3.32.

Logic State	Distributions	Thresholds [μA]	Error Probability
0	000	≤ 18	3.4×10^{-17}
1	001 010 100	$[18, 105]$	2.5×10^{-4}
2	011 101 110	$[105, 165]$	3.6×10^{-3}
3	111	≥ 165	5.6×10^{-4}

As further extension, the 5-Input Majority gate is demonstrated. In this arrangement 32 total input state combinations are possible which split into 6 distinct current levels [219]. Due to a limited measurement time, the sample size is reduced and each of the identical permutations was only programmed once. For instance, the combination 00001 was measured while the combinations 00010/00100/01000/10000 were left out as they are assumed to be identical. The remaining 6 combinations were programmed 16 times into one quintet of devices and read out 20 times. The resulting current means and histograms are depicted in Fig. 3.33. The limited number of data points results in narrowed distributions. Still, the feasibility of a 5-device readout is demonstrated as all 6 expected current levels are distinguishable. This can be exploited in a 5-Input Majority gate with logic outputs 0 for the combinations 00000, 00001 and 00011, and logic 1 for 00111, 01111 and 11111. Furthermore, a 2 1/2-bit or 6-state readout is possible by assigning each current level a unique logic state from 0 (00000) to 5 (11111).

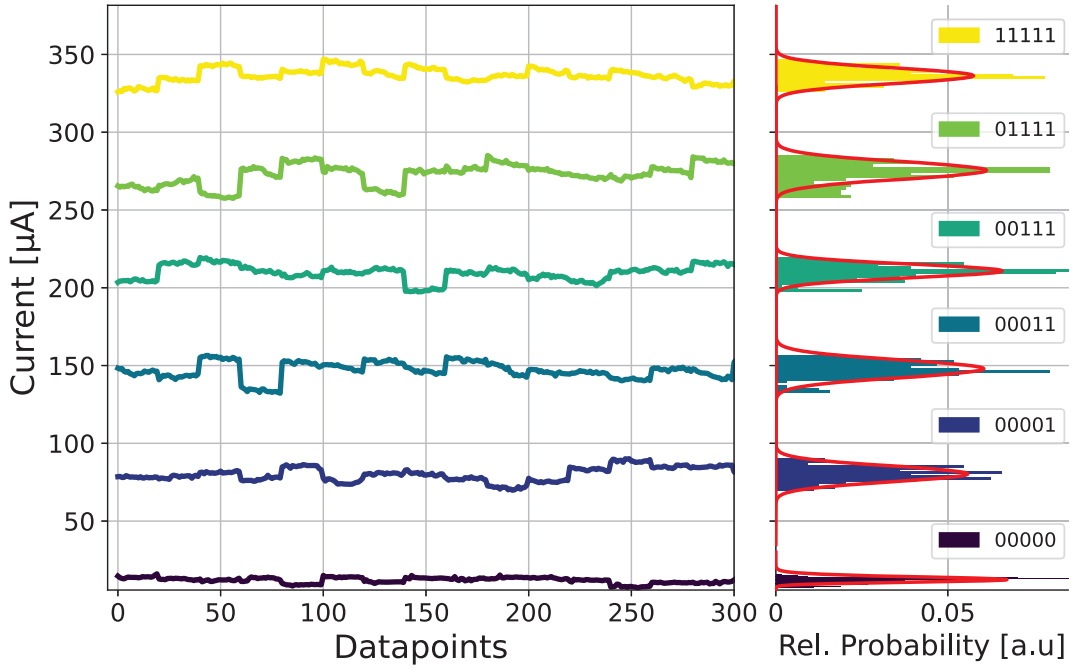


Figure 3.33: Experimental current means for a 5 1T-1R device parallel readout with corresponding distribution histograms. Measured and displayed are only the input combinations 00000, 00001, 00011, 00111, 01111 and 11111 without their respective permutations. Each combination is programmed 16 times into a device quintuple. Following each programming, the readout is performed 20 times to account the read instability with a single read of 200 ms at $V_{\mathrm{read}} = 0.25\,\mathrm{V}$ and $V_{\mathrm{G}} = 5\,\mathrm{V}$.

Attemps were made to fit the corresponding distributions and to calculate the error probabilities as before for both, the 5-Input Majority as well as for the 6-state readout. The calculation results are listed in Table 3.13 and Table 3.14. It has to be noted that as the data set is reduced and does not cover any device-to-device variability, the fitting resulted in very high residuals and while the calculation is mathematically possible, the physical meaning of the resulting error probabilities is limited.

For the 5-Input Majority, the calculated error probabilities are several orders of magnitude lower as for the 3-Input gate. This contradicts the expectation that each input causes additional variability. Due to the small sample size, the distribution width is not resolved appropriately. As the exploited error estimation method relies on the fitting of statistical distributions, it depends on the quality of the

underlying data sets to cover variability features. Despite the actual values being questionable low, their overall trend however matches to the observations of the 3-Input Majority in Table 3.11. The probability of a false classification is highest for the adjacent distributions 00011 and 00111, and lowest for the outermost distributions, 00000/00001 and 01111/11111 respectively. The situation worsens for the 6-state distinction in Table 3.14. While the 00000 state can be classified reliably, all other combinations show significant error probabilities in the order of 1×10^{-3}. While the actual values might vary due to the reduced fitting accuracy, the behavior is similar as for the 2-bit readout in Table 3.12.

Table 3.13: Calculated error probabilities for the measured 5 Input non-stateful Majority (MAJ). The calculation is performed based on the data displayed in Fig. 3.33.

Distributions	Intersection [µA]	Probability 'False 1'	Probability 'False 0'
00000 - 00111	136	7.8×10^{-66}	7.5×10^{-67}
00000 - 01111	170	1.9×10^{-78}	1.6×10^{-79}
00000 - 11111	232	1.2×10^{-97}	8.0×10^{-99}
00001 - 00111	150	3.8×10^{-44}	3.6×10^{-44}
00001 - 01111	173	1.5×10^{-75}	1.6×10^{-75}
00001 - 11111	210	6.2×10^{-145}	7.3×10^{-145}
00011 - 00111	183	1.3×10^{-10}	1.3×10^{-10}
00011 - 01111	212	2.1×10^{-30}	2.2×10^{-30}
00011 - 11111	248	1.4×10^{-71}	1.7×10^{-71}

Table 3.14: Calculated maximum error probabilities for the experimental readout with 2 1/2-bits/6 logic states. The calculation is performed based on the data displayed in Fig. 3.33.

Logic State	Distributions	Thresholds [µA]	Error Probability
0	00000	≤ 39	2.2×10^{-16}
1	00001	$[39, 112]$	4.5×10^{-4}
2	00011	$[112, 183]$	4.6×10^{-4}
3	00111	$[183, 239]$	6.1×10^{-4}
4	01111	$[239, 308]$	1.1×10^{-3}
5	11111	≥ 308	3.3×10^{-4}

In order to further evaluate the multi-input operations, both the 3-Input and 5-Input readout are investigated in simulations. The resulting histograms and distributions are depicted in Fig. 3.34 and Fig. 3.35. Additionally, the maximum error probabilities for the 2-bit and 6-state readout are calculated based on the simulated distributions.

The 3-Input 2-bit readout currents in Table 3.15 match to the measured data in Table 3.12. In both sets, 000 can be reliably distinguished from all other combination. However the distinction of those combinations demonstrates a higher error probability. A similar trend is observed also for the simulated

6-state simulations in Table 3.16 compared to the reduced measurement data in Table 3.14. In this case the distributions of 00011, 01111 and 11111 already visually overlap. Furthermore, the error probability for these distributions increased to approximately 2 %. At this point, an accurate state classification is not possible anymore.

The simulation results support the conclusions from measurement data, confirming that the multi-input readouts can indeed be executed as non-stateful logic operation. However, the variability observed in the low resistive state is crucial for an accurate state classification. The calculated error probabilities indicate that this becomes increasingly important for a growing number of inputs. Therefore, a precise control over the LRS variability is necessary to ensure a reliable performance.

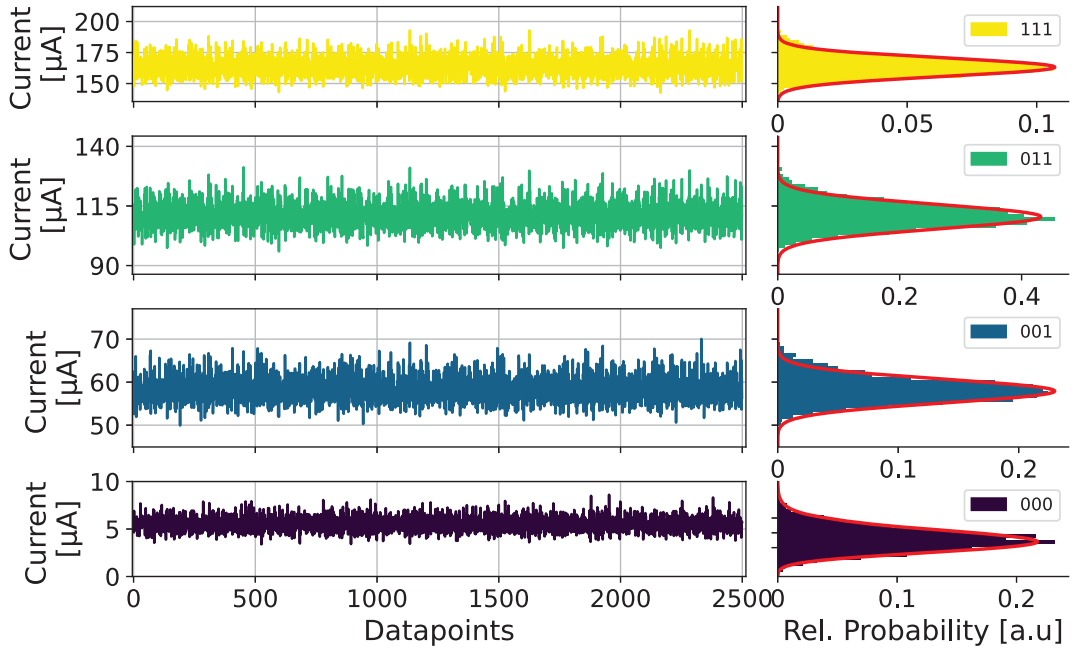


Figure 3.34: Simulated current means for a parallel 3 1T-1R device readout with corresponding distribution histograms for the input combinations HRS-HRS-HRS (000) (bottom), HRS-HRS-LRS (001) (second from bottom), HRS-LRS-LRS (011) (second from top), and LRS-LRS-LRS (111) (top). As there is no position dependence included in the simulation, only one permutation of (001) and respectively (011) was simulated. Displayed are 2 500 data points with the variability parameters as described in Table 2.1 and Table 2.2, and a readout voltage of $V_{\text{read}} = 0.25\,\text{V}$. For all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed.

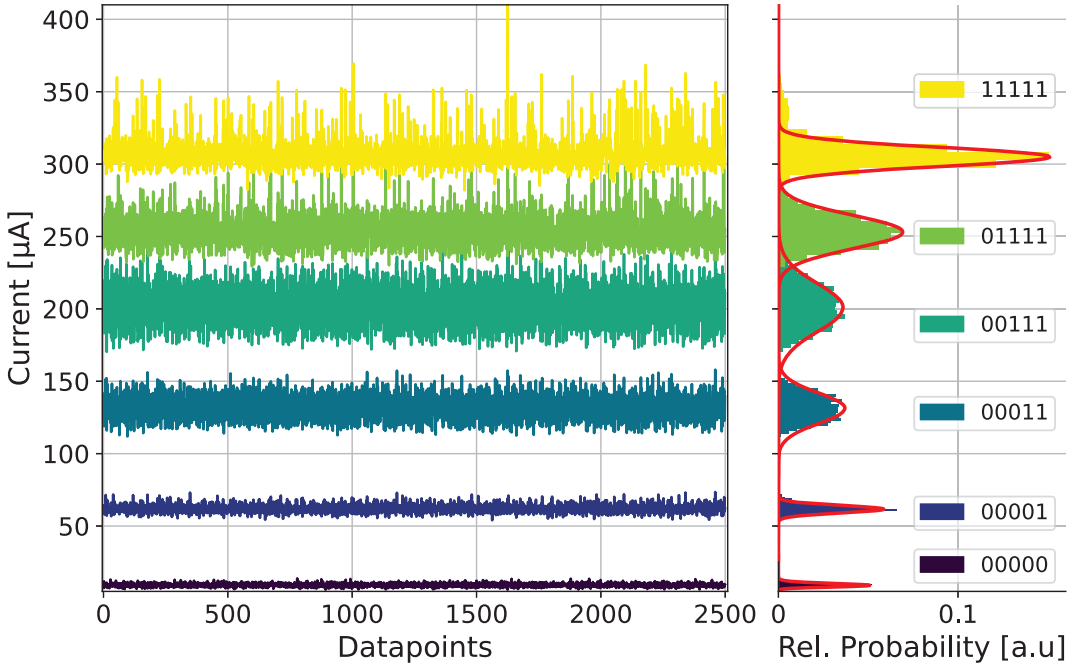


Figure 3.35: Simulated current means for a parallel 5 1T-1R device readout with corresponding distribution histograms for the input combinations 00000, 00001, 00011, 00111, 01111 and 11111. As there is no position dependence included in the simulation, only one permutation was simulated. Displayed are 2 500 data points with the variability parameters as described in Table 2.1 and Table 2.2, and a readout voltage of $V_{\text{read}} = 0.25\,\text{V}$. For all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed.

Table 3.15: Calculated maximum error probabilities for the simulated 2-bit readout with 4 logic states. The calculation is performed based on the data displayed in Fig. 3.32.

Logic State	Distributions	Thresholds [µA]	Error Probability
0	000	≤ 27	5.8×10^{-30}
1	001	$[27, 76]$	7.2×10^{-4}
2	011	$[76, 132]$	8.1×10^{-4}
3	111	≥ 132	6.2×10^{-5}

Table 3.16: Calculated maximum error probabilities for the simulated readout with 2 1/2-bits/6 logic states. The calculation is performed based on the data displayed in Fig. 3.35.

Logic State	Distributions	Thresholds [μA]	Error Probability
0	00000	≤ 35	1.0×10^{-23}
1	00001	$[35, 75]$	9.0×10^{-4}
2	00011	$[75, 158]$	6.0×10^{-3}
3	00111	$[158, 229]$	5.0×10^{-2}
4	01111	$[229, 285]$	2.1×10^{-2}
5	11111	≥ 285	1.4×10^{-3}

3.2.5 Transistor-specific Logic Operations

As seen in Section 3.2.2.2 and Section 3.2.3.2, the transistor gate voltage has a major impact on the operation results of readout-based logic. The gate voltage determines the transistor channel resistance and by this the total current flowing through a 1T-1R bitcell. Yet, so far the binary information was only stored within the switchable ReRAM state as LRS or HRS and the transistor served solely as selector device. In contrast to that, the concept of transistor-specific logic exploits the ability to adjust the total 1T-1R path resistance not only by switching the ReRAM state but also by tuning the transistor channel resistance via the applied gate voltage. This corresponds to a 2-Input logic AND function between the transistor gate voltage and the ReRAM resistance as logic inputs. The concept of exploiting the transistor gate as additional logic input was proposed by Wang et al. in 2017 [225]. One possible application for this additional input is the binary vector-matrix multiplication or hamming weight calculation. For the inputs 0/1, the logic AND resembles a binary multiplication. This application was already shown in a passive resistive crossbar based on complementary resistive switches [226].

The content of this section was already published in [206] and some texts are taken from there.

3.2.5.1 Single Device

Fig. 3.36 presents the circuit arrangement for a single-device transistor-specific logic operation. With A being the ReRAM state (HRS/0 or LRS/1) and X the applied gate voltage at the bitcell transistor (GND/0 or V_{DD}/1), four different input combinations are possible. The truth table for this operation is depicted in Table 3.17. The gate voltage for a logic 1, V_{DD}, has to be selected sufficiently high to ensure a minimal drain-source channel resistance. On the other hand if the gate voltage is significantly below its threshold, for instance at ground potential GND, only a sub-threshold leakage contributes to the source-drain current. This leakage ranges in the order of nA or less [227, 228] which should be substantially lower than the HRS currents in VCM-ReRAM. In the following, the combinations are labeled in the form of *X-A*.

Table 3.17: Logic truth table for the single device transistor-specific 1T-1R AND operation. The binary logic inputs are the transistor gate voltage X (GND/0 or VDD/1) and the non-volatile ReRAM state A (HRS/0 or LRS/1).

ReRAM State	Gate Voltage V_G	Readout Current	Logic State
HRS	0	$I_{Leakage}$	0
LRS	0	$I_{Leakage}$	0
HRS	V_{DD}	I_{HRS}	0
LRS	V_{DD}	I_{LRS}	1

Fig. 3.37 depicts the current traces for a single device transistor-specific logic operation. The data was measured with the Elmos setup for TaO_x-based VCM ReRAMs. Due to the low sample size of 10 traces per combination, a histogram fitting was not possible. However, it can be seen that the 0-HRS signals demonstrate currents around $0.5\,\mu\mathrm{A}$ which is higher than the expected transistor leakage. The spacing between the 0-HRS and 1-HRS combinations is only $\approx 1\,\mu\mathrm{A}$. Due to this narrow margin it is difficult to separate the three expected current levels which were derived in Table 3.17. Instead, 0-HRS and 1-HRS can be assumed as one shared combination.

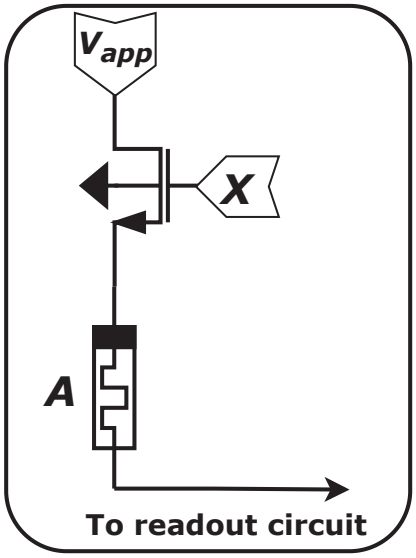


Figure 3.36: Circuit arrangement for Transistor-specific logic operations with of a single 1T-1R device. The ReRAM state is denoted as $A \in \{\text{HRS}, \text{LRS}\}$ and the transistor gate voltage as $X \in \{0, V_{\text{DD}}\}$ with V_{DD} being larger than the transistor threshold voltage.

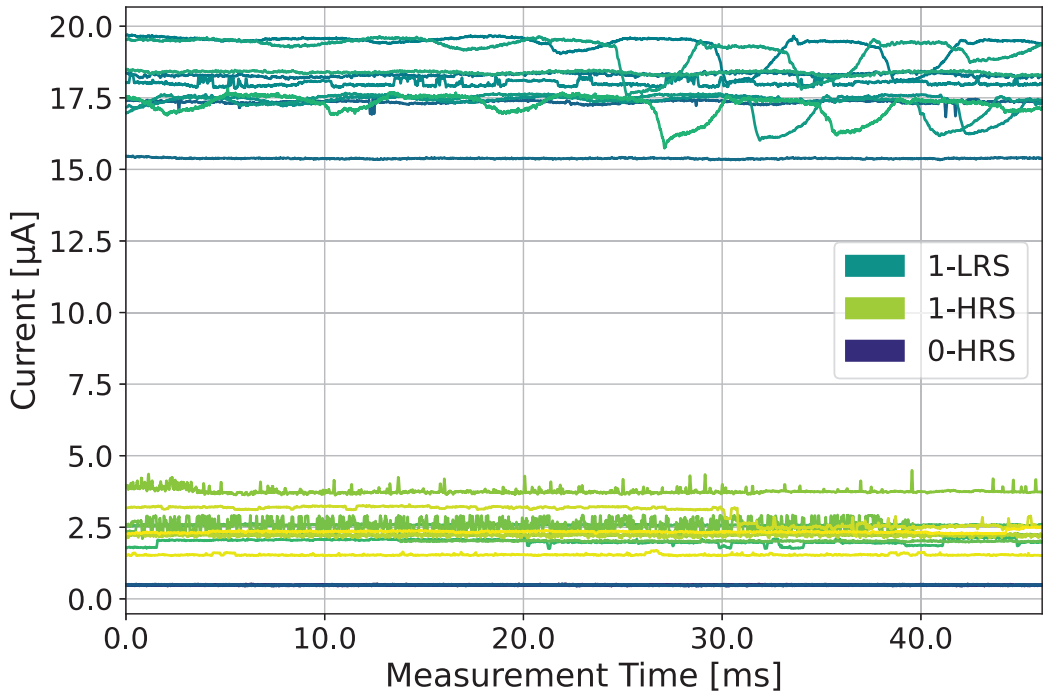


Figure 3.37: Experimental time-current data for a single transistor-specific logic operation the Elmos setup. Each state was cycled 10 times and readout with a pulse duration 500 x 100 µs, $V_{\text{G}} \in [0\,\text{V}, 6\,\text{V}]$ and $V_{\text{read}} = 0.2\,\text{V}$.

The same operation was measured using the MAD-200 setup. The resulting current histograms are depicted in Fig. 3.38. For this measurement, each input combination, 0-HRS, 0-LRS, 1-HRS, and 1-LRS, was programmed 12 times across 12 devices, with each configuration read out 10 times. The combinations 0-HRS and 0-LRS exhibit current values around $0.5\,\mu\mathrm{A}$ which is consistent to the Elmos data above. These values are again higher than the expected transistor leakage. However, the distributions appear to be symmetric around a mean of $0\,\mu\mathrm{A}$. This indicates that these current values are likely the result of noise introduced by the measurement setup rather than transistor leakage. The significant overlap between the current distributions of the leakage combinations, 0-HRS and 0-LRS, and 1-HRS is already visually evident, making their separation unreliable. This is further supported by the calculated error probabilities as listed in Table 3.18. While 1-HRS and 1-LRS are reliably distinguishable, the likelihood of incorrect classification between 0-HRS or 0-LRS and 1-HRS is larger than $1\,\%$. This means that 1 in 100 readouts is expected to be false.

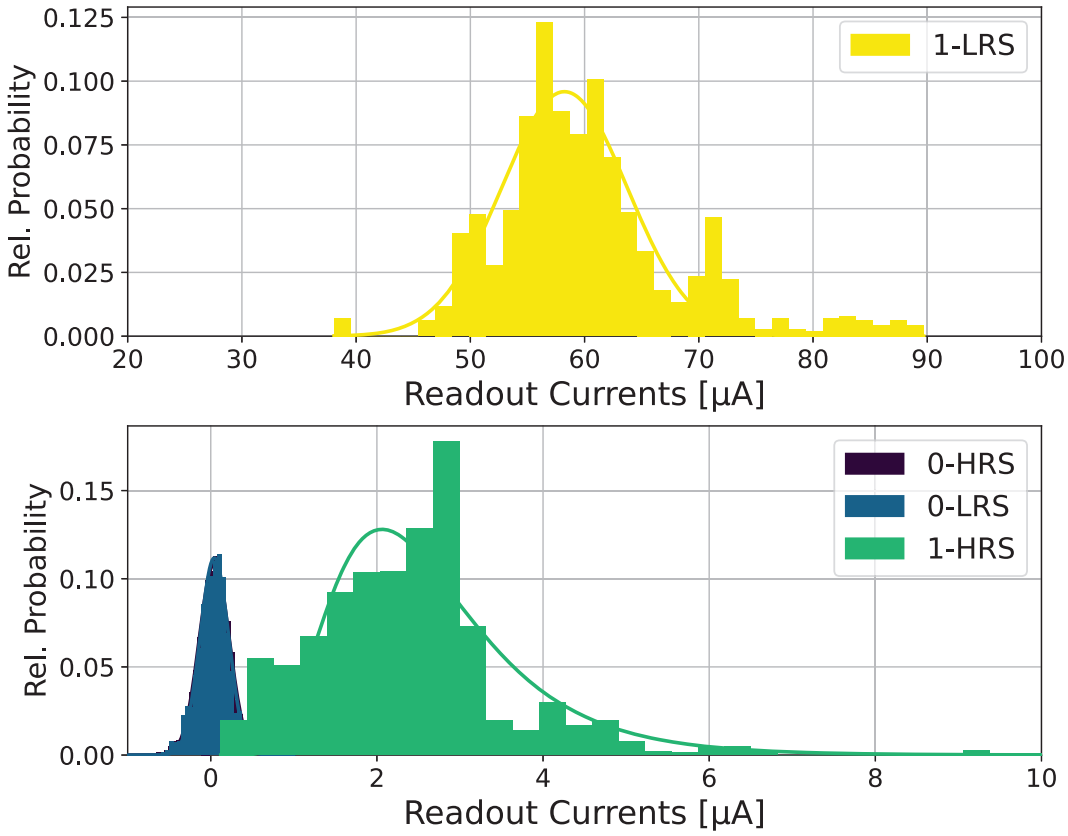


Figure 3.38: Histograms and fitted distributions for the experimental currents in a transistor-specific single device readout. Each of the four displayed combination was programmed 12 times into 12 different devices. Following each programming, the readout is performed 10 times to account the read instability with a single read of 200 ms at $V_{\mathrm{read}} = 0.25\,\mathrm{V}$ and $V_{\mathrm{G}} = 5\,\mathrm{V}$.

The operation was also simulated based on the setup described in Table 2.1 and Table 2.2. The resulting histograms are depicted in Fig. 3.39. In contrast to the measurements above, three different current levels can be distinguished in the simulations. In this case, the leakage combinations, 0-HRS and 0-LRS, are clearly distinguishable as their currents are far below the 1-HRS ones. Due to the very low absolute current values of the leakage combinations, 0-HRS and 0-LRS, it was not possible

Table 3.18: Calculated error probabilities for the measured single device transistor-specific readout. The calculation is performed based on the data displayed in Fig. 3.38.

Distributions	Intersection [μA]	Probability 'False 1'	Probability 'False 0'
0-HRS - 1-HRS	0.6	4×10^{-4}	0.022
0-LRS - 1-HRS	0.6	3×10^{-4}	0.08
1-HRS - 1-LRS	25	9.8×10^{-9}	1.1×10^{-9}

to calculate a reasonable intersection point between these combinations and 1-HRS or 1-LRS as the numerical equation solver reached its convergence limit. Therefore, the simulation results can only serve for a qualitative investigation. Yet, the simulations indicate that the concatenation of volatile gate voltage input and non-volatile ReRAM state behaves according to the expected logic AND, which contradicts the experimentally observed leakage currents in Fig. 3.37 and Fig. 3.38. An accurate current detection, even within the sub-threshold regime, is therefore essential for effectively separating more than 2 logic states. Without such precise measurement capabilities, only the distinction between the low current combinations, 0-HRS, 0-LRS, and 1-HRS, and the high current one, 1-LRS, can be implemented reliably which correlates to a logic AND.

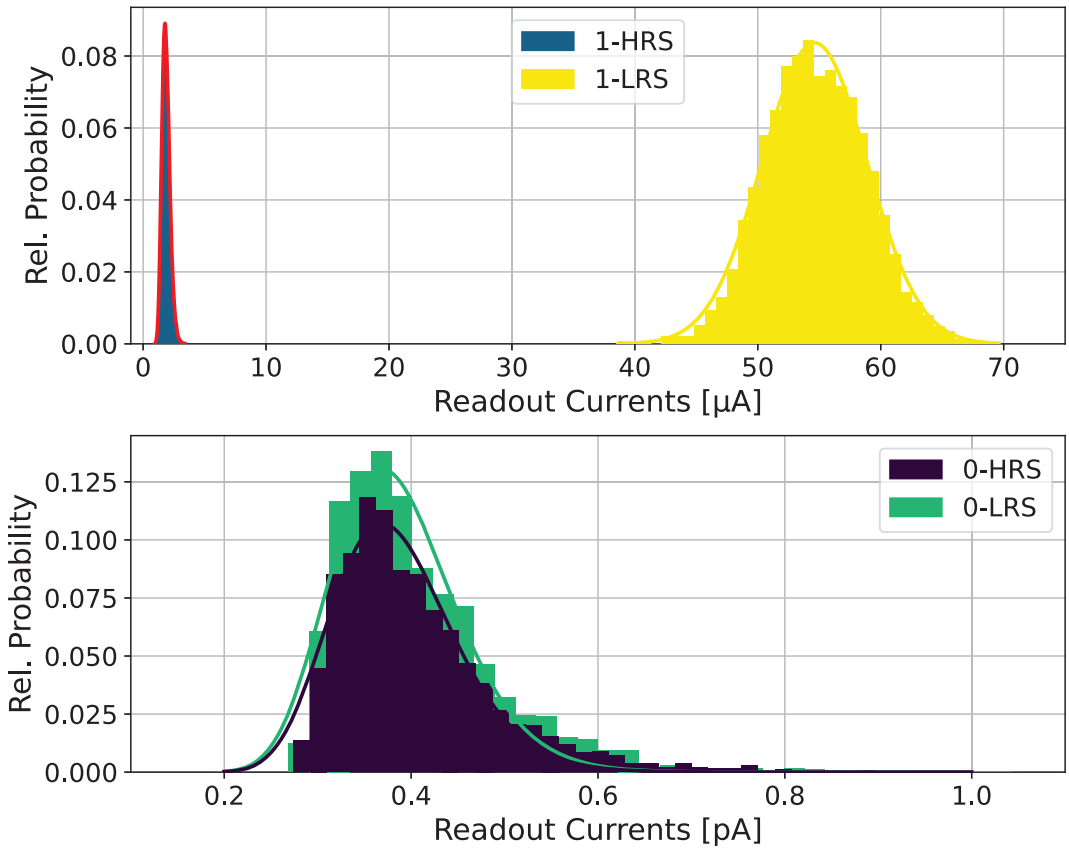


Figure 3.39: Simulated current histograms and fitted distributions for the transistor-specific single device readout. Displayed are 2 500 simulated data points for each combination with the variability parameters as described in Table 2.1 and Table 2.2, and a readout voltage of $V_{\text{read}} = 0.25\,\text{V}$. For all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed.

3.2.5.2 Two Device Operations

The previously demonstrated 2-Input operations AND ($*$), OR ($+$) and XOR ($\oplus$), can be concatenated with this 1T-1R specific AND, which results in the functions $((\mathrm{X} * \mathrm{A}) +, *, \oplus (\mathrm{Y} * \mathrm{B})]$ with $A, B \in \{\mathrm{HRS}, \mathrm{LRS}\}$ as ReRAM states and $X, Y \in \{0, V_{\mathrm{DD}}\}$ as transistor gate voltages. If both transistors are set to 1, it would correspond to the investigated standard gates in Section 3.2.3 again.

Table 3.19: Logic truth table for the two-device transistor-specific 1T-1R operations

No.	Gate Voltage $V_{\mathrm{G},1}$	ReRAM 1	Gate Voltage $V_{\mathrm{G},2}$	ReRAM 2	Readout Current
1	0	HRS	0	HRS	I_{Leakage}
2	0	LRS	0	HRS	I_{Leakage}
3	0	HRS	0	LRS	I_{Leakage}
4	0	LRS	0	LRS	I_{Leakage}
5	V_{DD}	HRS	0	HRS	I_{HRS}
6	V_{DD}	HRS	0	LRS	I_{HRS}
7	V_{DD}	LRS	0	HRS	I_{LRS}
8	V_{DD}	LRS	0	LRS	I_{LRS}
9	0	HRS	V_{DD}	HRS	I_{HRS}
10	0	HRS	V_{DD}	LRS	I_{LRS}
11	0	LRS	V_{DD}	HRS	I_{HRS}
12	0	LRS	V_{DD}	LRS	I_{LRS}
13	V_{DD}	HRS	V_{DD}	HRS	$2 \times I_{\mathrm{HRS}}$
14	V_{DD}	HRS	V_{DD}	LRS	$\approx I_{\mathrm{LRS}}$
15	V_{DD}	LRS	V_{DD}	HRS	$\approx I_{\mathrm{LRS}}$
16	V_{DD}	LRS	V_{DD}	LRS	$2 \times I_{\mathrm{LRS}}$

Table 3.19 displays all 16 possible input combinations of gate voltages $V_{\mathrm{G},1/2}$ and ReRAM states HRS/LRS. As before for the standard non-stateful gates, several of these combinations are permutations of the same distinct state. This is the case for the combinations 2 to 3, 5-8 to 9-12, and 14 to 15. Yet in theory at least four different currents levels should be distinguishable in this operation. Based on a single ADC readout, either a logic AND or OR can be computed depending on the threshold selection. If a second sense amplifier is used, a XOR is possible or with a 2-bit ADC even a 2-bit multi-level readout could be enabled.

As the combinations 13-16 are equivalent to the standard non-stateful readout in Section 3.2.3, their data is taken from the previously presented discussion in Fig. 3.20. All other unique combinations were measured, namely 1, 2 and 4-8. Fig. 3.40 depicts the mean currents and histograms for the transistor-specific operations. Each combination was programmed 256 times into a pair of two devices and readout 10 times. It is seen that even for gate voltages of $V_{\mathrm{G},1/2} = 0\,\mathrm{V}$, currents in the range of $0.5\,\mu\mathrm{A}$ are measured. These values match the single device sub-threshold leakage currents in Fig. 3.37 and Fig. 3.38 which are attributed to measurement noise.

In Fig. 3.41 the corresponding current histograms are presented including the combinations 1-HRS 1-HRS, 1-HRS 1-LRS, 1-LRS 1-HRS and 1-LRS 1-LRS from Fig. 3.20. As for the previous investigated

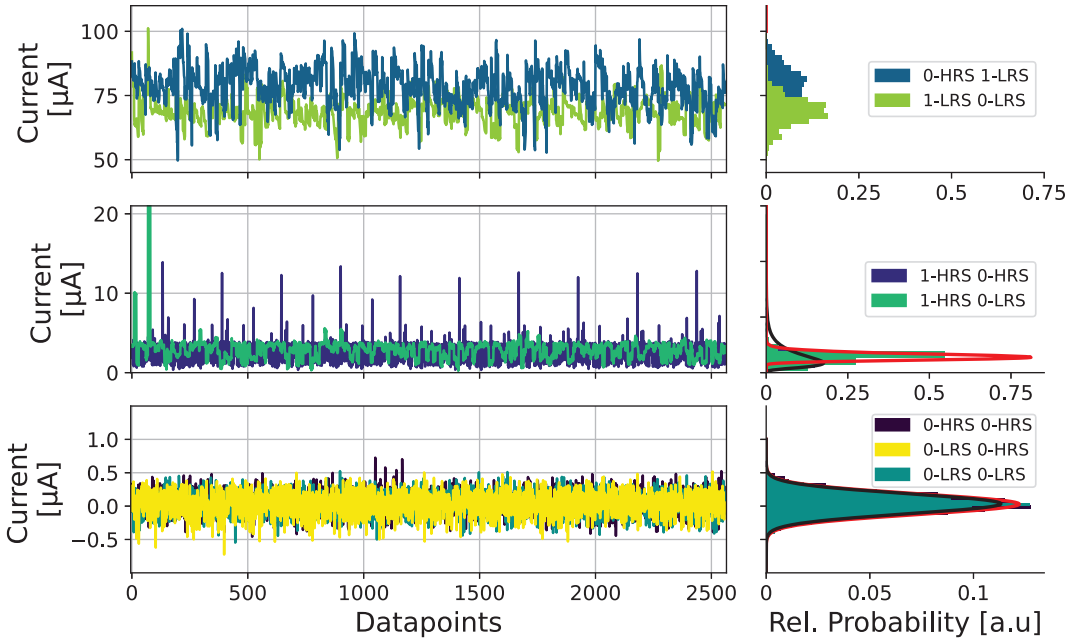


Figure 3.40: Experimental current means of a 2 1T-1R device transistor-specific parallel readout with corresponding distribution histograms for the input combinations 0-LRS 0-LRS, 0-HRS 0-HRS, 0-LRS 0-HRS, 1-HRS 0-HRS, 1-HRS 0-LRS, 1-LRS 0-LRS and 0-HRS 1-LRS. Each input combination was programmed 256 times into one device pair and readout 10 times.

standard non-stateful 2-Input gates, the currents split in 3 ranges: a low-current state (1-6,9,11,13), a broad mid-current state (7,8,10,12,14,15) and one high-current state (16). In particular, the combinations 1-HRS 0-HRS and 1-HRS 0-LRS demonstrate a high overlap with 1-HRS 1-HRS which is similar as for 0-HRS 1-LRS and 1-LRS 0-LRS to 1-HRS 1-LRS and 1-LRS 1-HRS. Therefore, the transistor-specific logic does not lead to the expected additionally current levels.

In addition, all 10 unique current combinations were simulated in the variability-aware setup as described in Table 2.1 and Table 2.2. The resulting histograms are depicted in Fig. 3.42. In general, the simulated currents are similar to the measured distributions in Fig. 3.41. However, the simulated transistor leakage in the combinations 0-HRS 0-HRS, 0-HRS 0-LRS and 0-LRS 0-LRS is significantly lower. This supports the assumption that these high current are related to measurement noise. Furthermore, while the combinations 0-HRS 1-HRS and 0-LRS 1-HRS show less overlap with 1-HRS 1-HRS compared to the measurements, the corresponding error probability between these combinations is still around 8×10^{-3}. Therefore, also in the simulations no additional current level besides the transistor leakage can be identified.

For the transistor leakage combinations 0-HRS 0-HRS, 0-HRS 0-LRS and 0-LRS 0-LRS, the error probability calculation failed again as the numerical intersection solver reached its convergence limit. Therefore the error probabilities are calculated only between the three distinguishable cases: low current (1-HRS 0-HRS, 1-HRS 0-LRS and 1-HRS 1-HRS), medium current (0-HRS 1-LRS, 1-LRS 0-LRS, 1-HRS 1-LRS and 1-LRS 1-HRS) and high current (1-LRS 1-LRS). For both logic gates, AND and OR, the simulated and measured maximum error probabilities are listed in Table 3.20. The calculated values are very similar to the standard non-stateful 2-Input gates in Table 3.8.

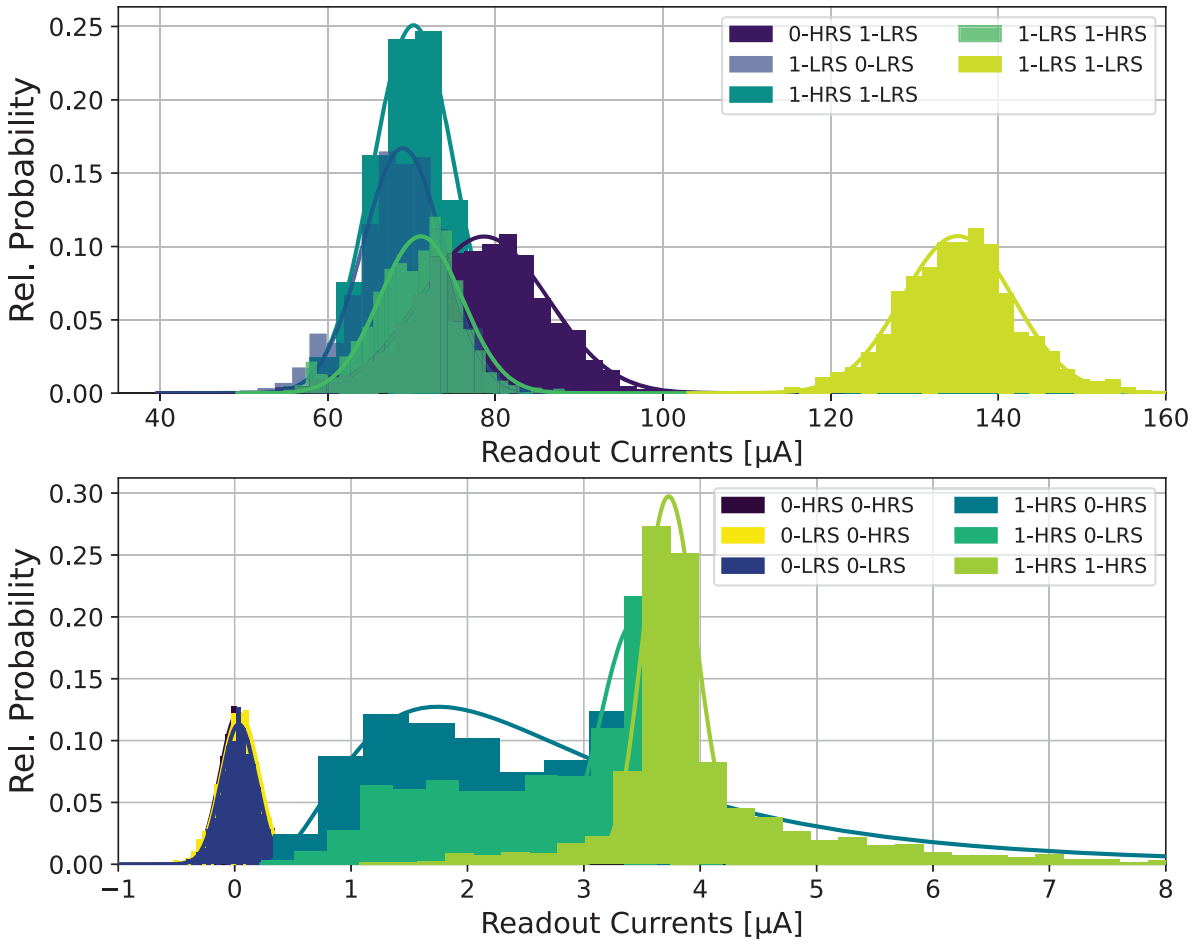


Figure 3.41: Current histograms and fitted distributions for the transistor-specific two-device operations from Fig. 3.40. Additionally, the currents from the standard two-device readout in from Fig. 3.20 are included which resemble the transistor-specific gates 1-HRS 1-HRS, 1-HRS 1-LRS, 1-LRS 1-HRS and 1-LRS 1-LRS.

Table 3.20: Calculated maximum error probabilities for the measured and simulated non-stateful transistor-specific operations OR and AND based on the distributions in Fig. 3.41 and Fig. 3.42.

	Measured			Simulated		
Logic Function	Threshold [µA]	Probability 'False 1'	Probability 'False 0'	Threshold [µA]	Probability 'False 1'	Probability 'False 0'
OR	17	9.1×10^{-24}	1.3×10^{-25}	15	1.3×10^{-20}	8.6×10^{-21}
AND	98	2.3×10^{-8}	1.0×10^{-8}	75	3.9×10^{-6}	5.7×10^{-6}

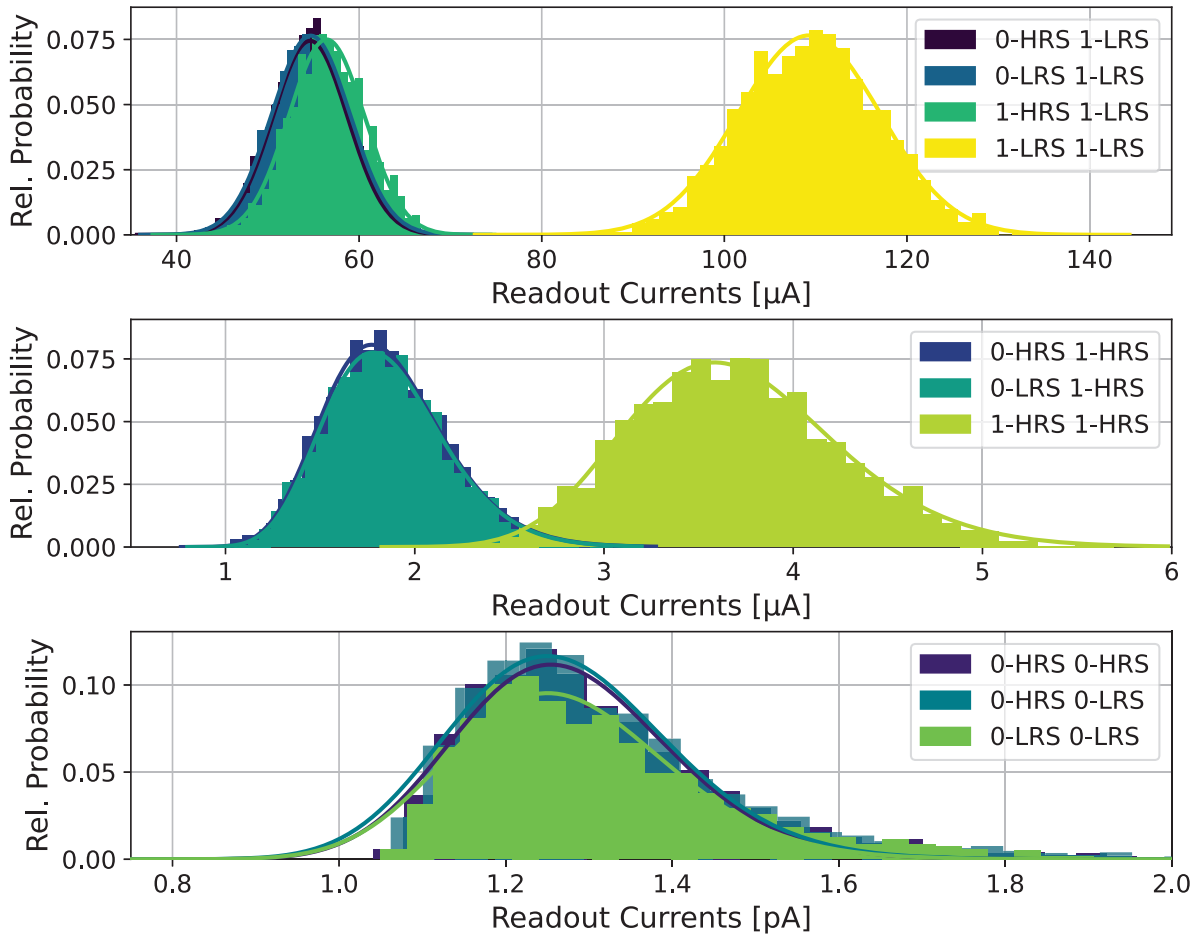


Figure 3.42: Simulated current histograms and fitted distributions for the transistor-specific two-device readout. Displayed are 2 500 simulated datapoints for each combination with the variability parameters as described in Table 2.1 and Table 2.2, and a readout voltage of $V_{\text{read}} = 0.25\,\text{V}$. For all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed.

3.3 Summary of Non-Stateful Readout Logic

This chapter dealt with the investigation of non-stateful readout operations which have been studied in simulation and measurements. These operations are based on the current-dependent state classification of one or several resistive devices. One important factor in the correct classification is understanding its threshold dependence. As example, higher threshold values will ensure that no HRS device is classified as LRS. However, devices which are in a higher LRS might be wrongly sensed as HRS. Vice versa, for lower thresholds a low HRS can be falsely sensed as LRS. Here it is important to characterize the state distributions and to calculate the optimal intersection which minimizes the distribution overlap. One possibility to improve the overlap in the single device readout is a complementary or differential sensing scheme. Here, two complementary devices get paired, either as HRS-LRS or as LRS-HRS [174]. Similar approach as complementary resistive switches [159]. One binary information in two devices improves the reliability by adding redundancy but doubles the required memory area.

A second and arguably the most crucial element for a successful operation are the applied voltages, V_{read} and V_{G}. A high gate voltage is necessary in order to reduce the transistor series resistance and to operate it in saturation. This will increase the readout current margin, especially for the low resistive state. Both, the LRS and HRS currents, will scale with increasing voltages as they show an approximately ohmic behavior. Yet, at least in one measurement setup, a current asymmetry is observed as the currents depend on the direction in which the readout is performed, either in SET or RESET. The second measurement setup however did not reproduce this effect. In any case for the selected configuration of active electrode connected to the transistor drain, the RESET readout is the better choice.

It has further been shown that an increase in readout currents will improve the separation between distributions of individual logic states. The clearer those states are separated, the more reliably can they be distinguished by a sensing circuitry. However, the effect of voltage scaling on these distributions is limited. A too high readout voltage V_{read} might lead to a read disturb. The performed measurements lead to the selection of $V_{\mathrm{read}} \leq 0.3\,\mathrm{V}$ as a stable readout margin. In contrast to the read instabilities, which were visible as randomly occurring current jumps, a read disturb is a permanent shift of the resistive state. The important goal is to find the voltage which maximizes the readout current margin but still does not influence the resistive states, at least for the expected number of consecutive readouts between two programming steps.

Related to this voltage dependence is the influence of the physical operation position within the crossbar structure. The parasitic circuit elements, e.g., line resistances, cause voltage drops. As result the actual operation voltage will differ throughout the crossbar. Depending on the ratio between the resistive state LRS and HRS and the parasitic components, the crossbar dimensions for a reliable performance are limited. In the measured example, the dimension of 512 1T-1R bit cells demonstrated no critical impact, while the simulations indicated that for dimensions above $1\,000$ the parasitic voltage drop cannot be neglected anymore.

The finding for single devices were also relevant in two-input logic functions. In the voltage range of $0.25\,\mathrm{V}$ - $0.3\,\mathrm{V}$, the two-input readout demonstrated the best separable current margins, while voltages above can already downgrade the reliability. The actual voltage range will depend on many factors such as the selected transistor technology or the ReRAM material stack. In particular, the non-linear *I-V*-characteristic of the oxide material will determine how likely a permanent read disturb is at a specific voltage. However, the underlying trade-off between increasing the current resolution by higher readout voltage and ensuring a non-destructive readout will occur for all bipolar switching resistive technologies.

For two inputs, as well as for the 3- and 5-input functions, the accurate state classification of all states containing at least one LRS device becomes more complex. The increased overlap between those distributions is caused by the higher absolute current variability of the low resistive state and its larger impact on the readout current, which results in broader distributions. Especially for a multi-state classification, neighboring distributions at higher currents are moving closer together. The separation of these states can be improved by narrowing the individually programmed LRS distributions.

The transistor-specific logic was introduced as a possible extension of the non-stateful readout operations. In this investigation, no benefit in number of separable current levels was observed. While the measured transistor-specific logic does not hold any advantage regarding the operation reliability, it also shows no critical negative effects. If the sensing circuitry is able to reliably resolve the transistor leakage current, the transistor-specific logic might be beneficial as another current level is accessible. Otherwise, adding an additional logic input in form of the transistor gate voltage might also be useful for combining operations in specific application with only a small degrade in reliability compared to the standard non-stateful readout. Yet, applying the gate voltages as logic inputs requires additional hardware. Therefore the applications of the transistor-specific logic have to be carefully selected.

CHAPTER 4

Switching-based Logic Operations

4.1 Fundamentals of Switching-based Logic Operation

Switching-based logic operations, as the term suggests, execute a logic function by conditionally switching one or more resistive output devices based on the provided inputs. Similar as for the readout-based logic, the analog resistances ranges are mapped to discrete logic values in order to realize digital logic functions. For instance a boolean switching-based logic operates with the binary states '0' and '1' which are represented by the high-resistive and low-resistive state, HRS and LRS.

The conditional switching is triggered by applying a voltage V_{cond} to the output devices. However, this voltage is not directly applied at the outputs but it depends on the N logic inputs, $V_{\text{cond}} = V_{\text{cond}}(\text{In}_1, \ldots \text{In}_n)$. The simplest case for such a conditional input is the programming step of a single 1T-1R bitcell. Independent of the applied drain-source voltage, V_{DS}, the voltage drop across the ReRAM device depends also on the gate-source voltage, V_{GS}. In this case the gate voltage is the only logic input. If it is high enough, the voltage drop across the ReRAM is sufficient for switching, otherwise the ReRAM will remains its state.

In order to implement a stateful and input conserving switching-based logic function, at least one input should be represented as the non-volatile resistive state of one or several ReRAM devices. This will create a voltage-divider arrangement in which the voltage drops across the individual memristive devices are determined by their resistive state in relation to the total resistance of all devices in the circuit. Given that the number of input and output states is finite, there is only a limited number of possible logic combination. As example, in a two-input one-output binary operation, $2^3 = 8$ unique operations are possible. For each combination, the specific voltage drop configuration can be calculated using Ohm's law and Kirchhoff's current laws. This voltage drop, V_{cond}, depends on the input voltage, the resistive states of the input and output devices, and the ratio between high-resistive and low-resistive state. The magnitude of this voltage drop is critical as it determines whether an individual device will switch its state or remain unchanged. If it exceeds the threshold required for a ReRAM switching, the resistive state will change. Conversely if the voltage drop is too low, it will not induce a state change and the output device will retain its current state. Depending on the specific combinations under which the output device switches its state, the various logic combinations can be systematically organized into a logic truth table. This table essentially outlines the voltage drop for all possible input combinations and indicates whether a combination results in a output switching or maintaining state. The two extreme cases are *contradiction* and *tautology*. In the case of a contradiction or *zero-function*,

the output device does not change its state for any input combination. Opposed to this, in the case of a tautology or *one-function*, the output device switches its state for any set of inputs. Between these two extremes lies a range of feasible logic functions that can be realized by appropriately configuring the input conditions. The example of a simple switching-based gate with one logic input, R_1, and one logic output, R_2, in a serial connection is provided in Fig. 4.1. If the input voltage, V_1, is applied at R_1 and ground potential at R_2, the conditional voltage across R_2 is calculated according to a serial voltage divider, $V_2 = V_1 \times \frac{R_2}{R_1 + R_2}$. The executable logic function then depends on the relation between V_2 and the threshold voltage which is required for switching R_2, $V_{\text{threshold}}$. The logic truth table for this calculation is depicted in Table 4.1.

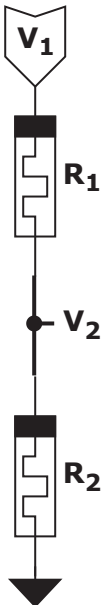


Figure 4.1: Schematic representation of a switching-based logic with one input, R_1, and one output, R_2, in serial connection. The voltage drop across the output, V_2 is set by the voltage divider arrangement in dependence of R_1, R_2 and the input voltage, V_1.

R_1	$V_2(V_1)$
$10\,\Omega$	$\frac{10}{11} \times V_1$
$100\,\Omega$	$\frac{1}{2} \times V_1$
Function	**Voltage condition**
contradiction	$\frac{1}{2} \times V_1 < \frac{10}{11} \times V_1 < V_{\text{threshold}}$
identity	$\frac{1}{2} \times V_1 < V_{\text{threshold}} < \frac{10}{11} \times V_1$
tautology	$V_{\text{threshold}} < \frac{1}{2} \times V_1 < \frac{10}{11} \times V_1$

Table 4.1: Logic truth table for the switching-based logic in Fig. 4.1. The conditional output voltage, V_2 is calculated from a serial voltage divider based on the input voltage, V_1, the resistive input $R_1 \in \{10\,\Omega, 100\,\Omega\}$, and the resistive output, $R_2 = 100\,\Omega$. Based on the relation between V_2 and the switching threshold voltage of R_2, $V_{\text{threshold}}$, the 3 stated logic functions can be implemented.

Similar as for the readout-based logic operations, the inherent variability in ReRAM devices introduces significant challenges to the predictability and reliability of the switching-based logic. Due to variability in the resistive states, the voltage drop across the output device will fluctuate. Therefore the output voltage is part of a distribution rather than being a fixed value. Consequently, the same logic operation performed on two separate circuits might result in different output states. As explained above, this effect can be mitigated by using write-verify based device programming.

Although these techniques can reduce the variation in the resistive states, the inherent stochastic nature of the VCM switching will lead to a variability in the switching threshold voltage, $V_{\text{threshold}}$, the effective voltage at which the output device will transition its state. This issue poses significant challenges for achieving a consistent logic performance. Due to this stochastic nature, the operation shifts from a deterministic to a probabilistic event. Instead of fixed truth tables, switching probabilities have to be introduced as a success measure of how likely it is for a specific input combination to induce

a correct output switching or prevent an incorrect switching. In this case, a *switching error* is defined as a discrepancy between the expected ReRAM behavior and the actual observed event. A possible way to reduce the likelihood of switching errors is to incorporate a variability-aware robust margin between the conditions which define distinct logic states. This will compensate some of the switching threshold variability. Otherwise if the voltage conditions are too close to each other, a false switching might be induced or a desired switching might not occur. Therefore it is favorable to take variability into account when assigning the output voltages to logic states.

A crucial part of the switching probability improvement is the careful adjustment of the input voltages. The inputs will determine the absolute voltage drop across each cell, independent of the voltage-divider portions. As the occurrence of a switching event is closely linked to the applied voltage, this will significantly determine the switching probability. For instance, all input voltages below the theoretical minimum threshold required for switching, $\approx 0.4\,\mathrm{V}$ for VCM devices, will never cause a switching regardless of the state arrangement. The goal for switching-based operations is to enable the desired switching while decreasing the possibility of a false switching at the same time. However this manifest the most severe issue in switching-based logic operations, the voltage-optimization dilemma. Both processes, increasing the switching rate while reducing unintended switching, contradict each other. Any change in the parameters that will lead to an increase in the switching probability for all desired combinations will also increase the switching probability for the undesired cases. Vice versa anything that inhibits the false switching will also worsen the desired switching operations.

4.2 Primitive Gates for Switching-based Logic

Over the years, several resistive-switching logic gates have been proposed. Among the first examples are the *stateful implication* (IMPLY) from Lehtonen et al. in 2009 [6] or logic based on *complementary resistive switches* (CRS-logic) by Linn et al. in 2010 [159]. In 2010, Borghetti et al. experimentally demonstrated the execution of the logic NOT and material implication (IMP) with two resistive switching devices [229]. In combination, NOT and IMP are logically complete, therefore all logic operation can be executed within these concepts. However in both gates at least one of the logic inputs is overwritten during the operation and therefore they do not fulfill the state conservation requirement. In 2014, the prominent *Memristor-Aided-Logic* (MAGIC) by Kvatinsky et al. was published which proposed a logic NOR gate based on two or more inputs and one output device [8]. This concept relies on a conditional RESET switching and is state conserving. Later on, the concept got extended for instance with the *Fast and Energy-Efficient logic* (FELIX) [230] or *Three Memristor Stateful logic* (TMSL) [231]. It is debatable whether these are stand-alone concepts or only improvements of the MAGIC concept from 2014. To not favor any group, all concepts in the following are referred to as *process*-based switching logic where *process* denotes either SET or RESET, depending on the VCM switching direction. The schematics of several exemplary switching-based logic gates are depicted in Fig. 4.2.

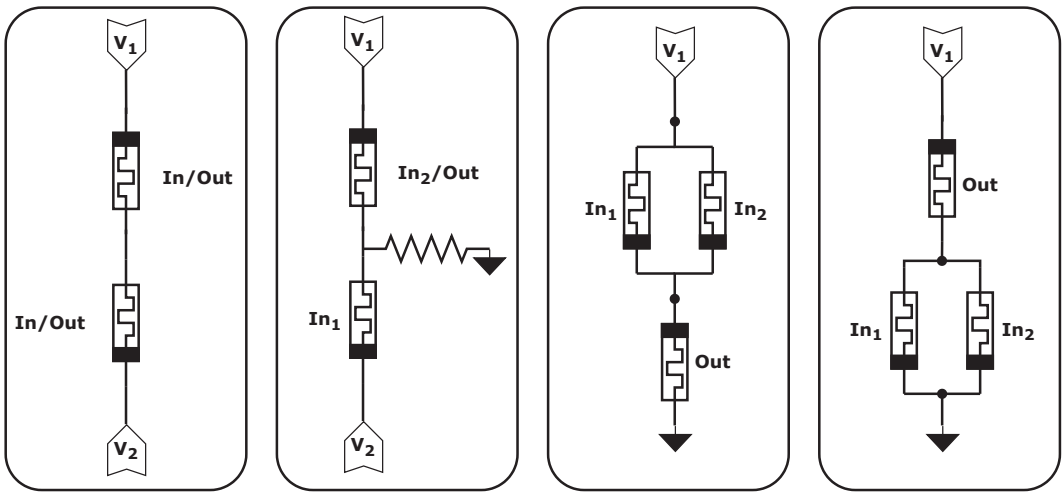


Figure 4.2: Exemplary primitive gates for switching-based logic operations which were proposed in literature, (a) Complementary resistive switches logic from Linn et al. [159], (b) Material implication from Lehtonen et al. [6], (c) Memristor-Aided Logic from Kvatinsky et al. [8], (d) Fast and Energy-Efficient Logic by Gupta et al. [230]. All circuit diagrams are self-drawn according to the original publications.

4.2.1 Error Estimation Framework for Switching-Based Logic

The overall goal of an error-estimation framework for switching-based logic is to approximate how likely it is for a specific output device A to switch if the operation X is performed. Similar as for the readout-based logic, several approaches are available for this task. For instance, it is possible to perform Monte-Carlo simulations or measure the individual gate and determine the error rate [199, 206,

232]. While this method is straight-forward, it requires a high number of simulations/measurements or dedicated high-sigma approaches in order to resolve low error rates [233]. The proposed framework extends the Monte-Carlo approach by introducing a probabilistic voltage-dependent switching. This framework was published in [190, 234] and parts of the text is taken from there.

As the resistive switching of VCM devices is an inherent stochastic mechanism, it is not possible to state fixed switching threshold voltages. Instead, depending on the applied pulse conditions, e.g., pulse width and amplitude, the probability for a state transition can be given [235–237]. Fig. 4.3 shows these probabilities for a simulated 1T-1R SET and RESET process with fitted Weibull-functions. The voltage- and time-dependent single-device switching probabilities state how likely it is for a device A to switch under the applied voltage V. This probability marks the first part of the framework.

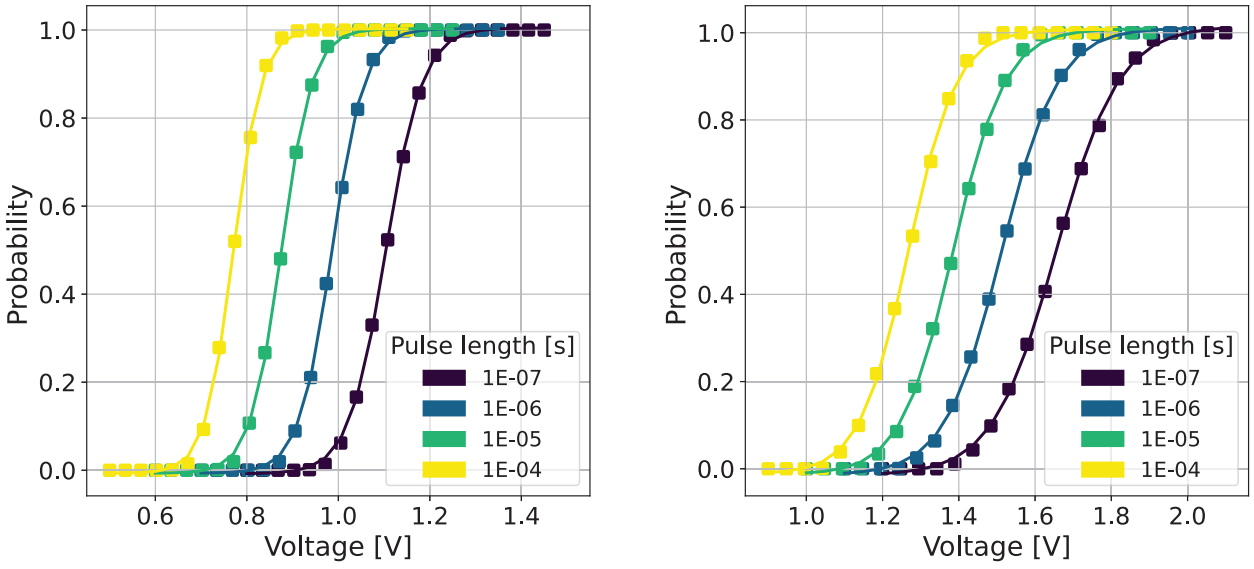


Figure 4.3: Voltage-dependent switching probability for the simulated (a) SET process and (b) RESET process of a single JART-VCM ReRAM device according to the model in Section 2.4.1.1, for varying pulse lengths. A successful switching is defined if the device shows a readout current of $I_{\text{post-switching}} > 26\,\mu\text{A}$ in SET (a), $I_{\text{post-switching}} < 26\,\mu\text{A}$ in RESET (b) for a readout voltage of $V_{\text{read}} = 0.25\,\text{V}$. This threshold is based on the results in Table 3.4. Each pulse is simulated 10 000 times. The simulations are performed based on the variability parameters in Table 2.1 and Table 2.2.

As second step, the actual switching-based logic operation is either simulated or measured. In contrast to standard Monte-Carlo simulations, the focus here not lies on evaluating the actual switching events. Instead, the voltage distribution across the output device is estimated as this quantity will determine a successful switching [197]. This framework step estimates, how likely it is that the output device in operation X undergoes a voltage drop of V. As example, Fig. 4.4 shows the simulated output voltage distributions for the switching-based logic gate in Fig. 4.1.

If both distributions are convoluted, it will result in the desired probability that operations X will lead to a switching event. This is mathematically expressed as

$$P_{\text{switching,i}} = \int_0^{V_{\text{max}}} P_{\text{i}}(V) \times S(V)\ \mathrm{d}V \tag{4.1}$$

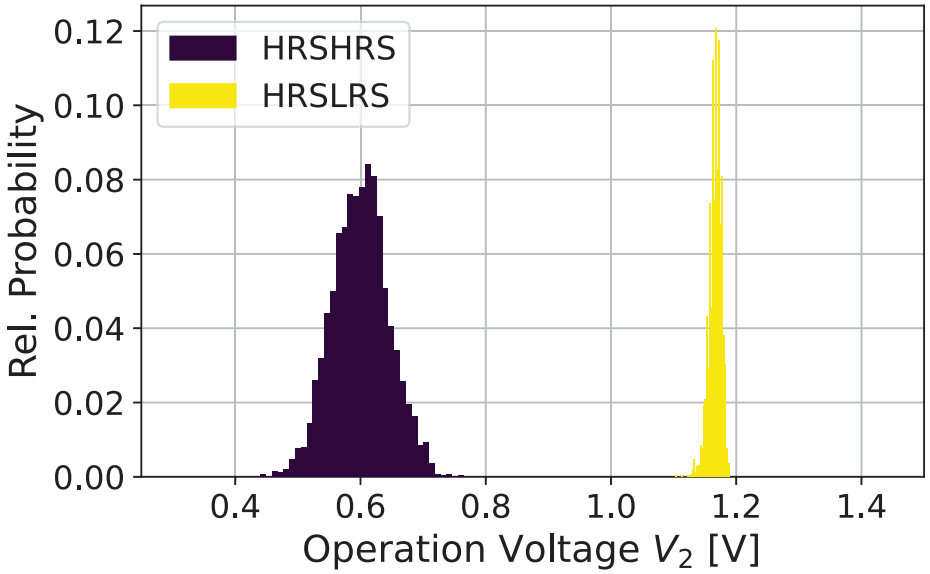


Figure 4.4: Simulated output voltages V_2 for a voltage-divider arrangement with a fixed R_1 in HRS state and R_2 either in HRS or LRS state according to the JART-VCM variability simulation parameters in Table 2.1 and the schematic as presented in Fig. 4.1. The input voltage, V_1, is set as $1.4\,\mathrm{V}$. Each combination is simulated $10\,000$ times.

with $P_i(V)$ being the output voltage distribution for input combination i, $S(V)$ the single device switching probability, and V_{max} the maximum applied voltage during the operation. The framework therefore combines aspects of a probabilistic switching [55, 65, 235, 236] with input-specific voltage distributions [197, 201].

4.2.2 Operative Conditions

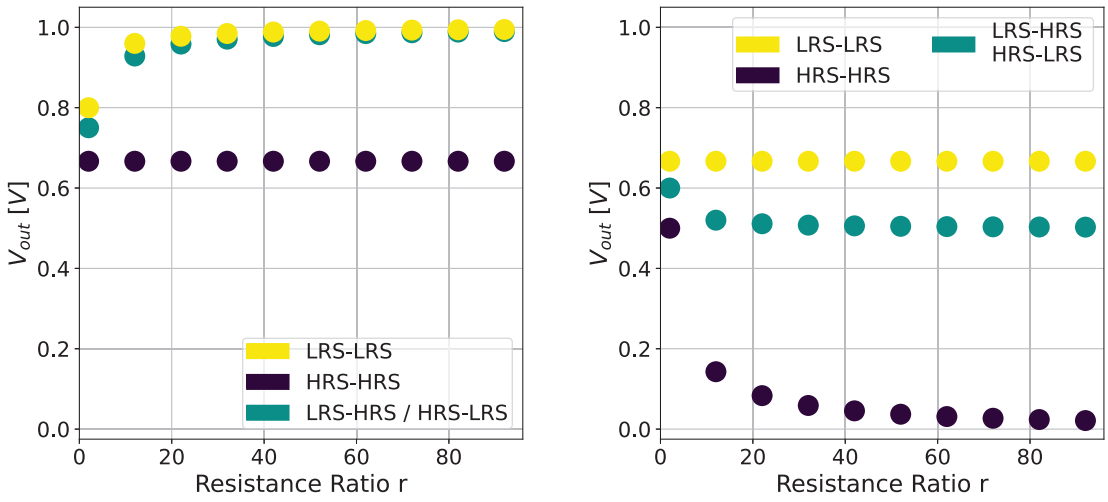
In the following, the focus is laid on switching logic gates with two or more logic inputs. This is due to the fact that one-input one-output operations such as IMPLY or CRS are not state conservative [232, 238]. For the simple case of two inputs, In_1 and In_2, and one output, Out, as in Fig. 4.2(c), the conditional switching voltage, V_{out}, is set as

$$V_{\text{out}} = V_{\text{in}} \times \frac{Out}{Out + \frac{In_1 \times In_2}{In_1 + In_2}} \tag{4.2}$$

If Out is initially in the high-resistive state, the operation is a SET-based switching logic as the output device switches conditionally into the low-resistive state. Respectively if Out is initially in the low-resistive state, the operation is a RESET-based switching. For both processes, four input combinations are possible, HRS-HRS, LRS-HRS, HRS-LRS and LRS-HRS. In case that both inputs received the same input voltage, the combinations LRS-HRS and HRS-LRS are symmetric. A successful switching-based logic function requires a certain amount of voltage separability. Only if the output voltages ,V_{out}, of different input combinations are distinguishable, they can be exploited for a conditional switching. To identify possible operations, Fig. 4.5 shows the calculated output voltages for a binary two-input one-output voltage divider. In general, a higher resistance ratio $r = HRS/LRS$ is favorable in order to separate different logic combinations. For the SET-based switching in Fig. 4.5(a), the cases of HRS-LRS/LRS-HRS and LRS-LRS behave nearly identical. Contrary, the RESET-based output voltage in Fig. 4.5(b) separates into three distinct voltages levels. This indicates that a SET-switching operation can only execute a logic OR while in the RESET-based switching both operation, NOR and NAND, are possible. However due to the proximity of HRS-LRS/LRS-HRS to LRS-LRS in this case, the logic NOR is more favorable. The logic truth table for these switching operations is listed in Table 4.2. The identification as OR and NOR in this case is based on the state assignment HRS = '0' and LRS = '1'. If this is inverted, the SET-based logic OR turns into a AND while the RESET-based NOR changes to logic NAND.

In case that the voltage V_{in} is applied at the two inputs, these devices will experience a voltage of $(V_{\text{in}} - V_{\text{out}})$ while the output device sees a voltage drop of $(V_{\text{out}} - 0) = V_{\text{out}}$ for an ideal ground. At least for the above identified cases, this V_{out} should be high enough to induce a switching. In order to maintain a state conservation, $(V_{\text{in}} - V_{\text{out}})$ needs to be sufficiently low to avoid an input drift. However, V_{out} is not a static voltage but adapts dynamically in case that the output device changes its state which will influence the total voltage divider arrangement. Fig. 4.6 depicts the simulated transient voltages, V_{out} and $(V_{\text{in}} - V_{\text{out}})$ for the four possible input combinations as SET- and RESET-based switching.

The SET in Fig. 4.6(a) shows an abrupt switching for the combinations LRS-HRS/HRS-LRS and LRS-LRS. This leads to a reduction in V_{out} from $V_{\text{out}} \approx V_{\text{in}}$ to $V_{\text{out}} \approx \frac{1}{3}V_{\text{in}}$. Vice versa, $(V_{\text{in}} - V_{\text{out}})$ increases to $\approx \frac{2}{3}V_{\text{in}}$ as their sum is constant. The RESET in Fig. 4.6(b) demonstrate a similar abrupt switching for the input combination LRS-LRS. However the combination LRS-HRS/HRS-LRS transitions its state only gradually. As the RESET is induced by local joule heating, a longer transition time is required for the lower V_{out} of this combination. Yet after the operation, the voltages arrange as $V_{\text{out}} \approx \frac{9}{10}V_{\text{in}}$ and $(V_{\text{in}} - V_{\text{out}}) \approx \frac{1}{10}V_{\text{in}}$. In the RESET-based switching, both nodes, $(V_{\text{in}} - V_{\text{out}})$ and V_{out}, operate on the total available voltage. This is favorable as the switching margin is extended. However, in case of the non-switching combination HRS-HRS, the input devices experience a very high applied voltage of $\approx \frac{9}{10}V_{\text{in}}$. In contrast to this, the SET-based margin is reduced to approximately $\frac{1}{3}$ of the total voltage.

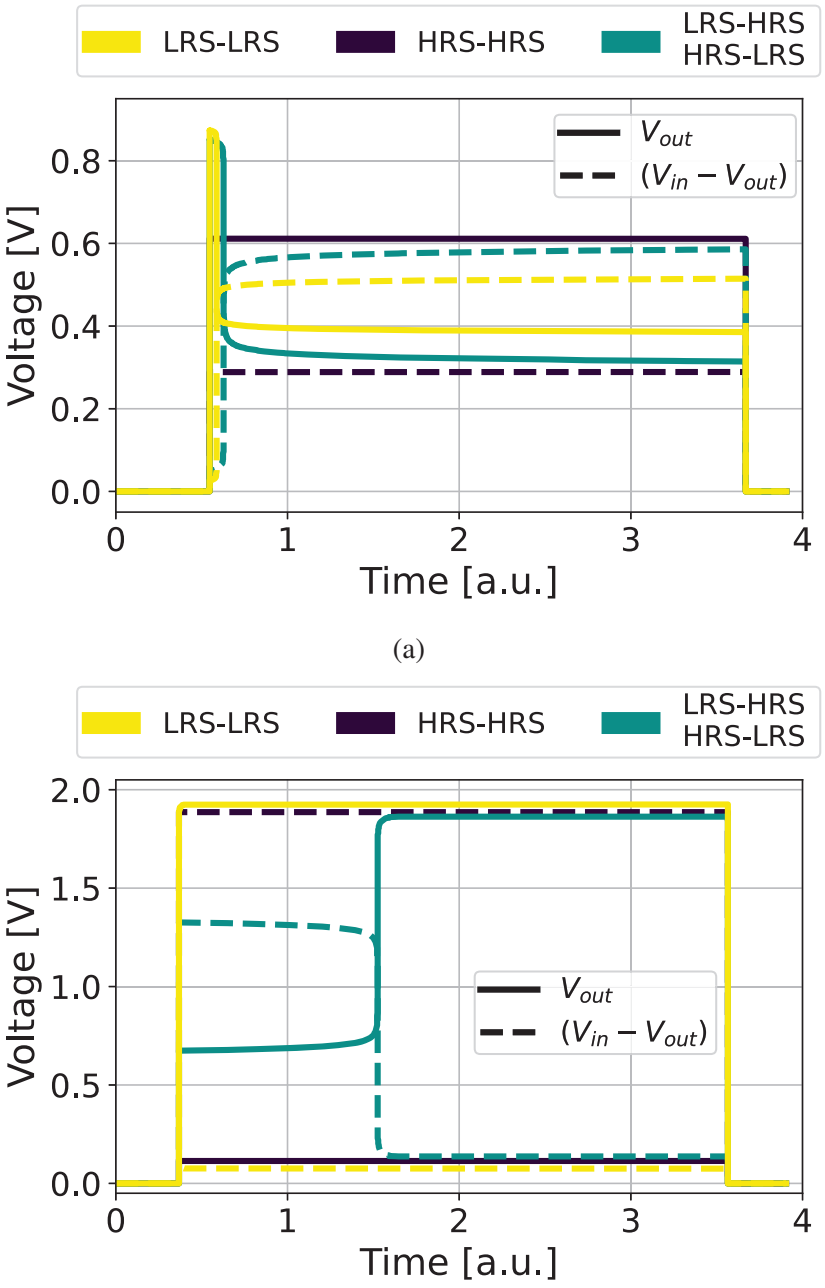


(a) (b)

Figure 4.5: Calculated conditional voltage divider output, V_{out} in dependence of the resistance ratio between HRS and LRS for (a) SET process and (b) RESET process. The calculation is performed based on Eq. (4.2) with $In_{1,2} \in \{LRS, HRS\}$, $LRS = 5\,\text{k}\Omega$, $HRS = r \times LRS$ and $V_{\text{in}} = 1\,\text{V}$. The initial output resistance is (a) $Out = HRS$ and (b) $Out = LRS$.

Table 4.2: Logic truth table for SET- and RESET-based two-input one-output switching operations. In the used assignment of HRS = '0' and LRS = '1', the SET-based operation resembles a logic OR, the RESET-based operation a logic NOR. It has to be noted that if the assignment is inverted, HRS = '1' and LRS = '0', the operations turn into a logic AND (SET-based) and logic NAND (RESET-based) respectively.

SET-based			RESET-based		
In_1	In_2	Out	In_1	In_2	Out
HRS 0	HRS 0	HRS 0	HRS 0	HRS 0	LRS 1
HRS 0	LRS 1	HRS → LRS 0 → 1	HRS 0	LRS 1	LRS → HRS 1 → 0
LRS 1	HRS 0	HRS → LRS 0 → 1	LRS 1	HRS 0	LRS → HRS 1 → 0
LRS 1	LRS 1	HRS → LRS 0 → 1	LRS 1	LRS 1	LRS → HRS 1 → 0



(a)

(b)

Figure 4.6: Simulated transient behavior of the conditional voltage divider for (a) SET-based and (b) RESET-based two-input one-output binary switching operation. The simulations are based on the parameters in Table 2.1 with (a) $V_{\text{in}} = 0.9\,\text{V}$ and (b) $V_{\text{in}} = 2.0\,\text{V}$. In order to isolate the ideal output behavior, the input devices are simulated as static i.e., not switchable VCM devices. For all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed.

This is problematic as even in the non-switching case, $\approx \frac{2}{3}V_{\text{in}}$ is applied to the output device which might cause a false switching. Another important aspect for the switching-based operation is the device polarity. In a crossbar, all devices should be able to serve as inputs and outputs which requires a identical device orientation. This means that for a RESET-based process the active ReRAM electrodes have to be commonly connected and the ohmic electrodes for the SET-based process. Furthermore, a potential input corruption also depends on the device orientation as the VCM switching is polarity-depending. Under this restriction, an increase voltage drop across the HRS-HRS inputs in a RESET-based switching might cause an issue. As V_{in} is applied at the ohmic electrode it might induce a SET process in the input devices. Both processes, SET-based and RESET-based, are investigated in detail in the following.

4.2.3 RESET-based Switching Logic

As the name states, a RESET-based logic exploits the resistive RESET process by conditionally switching an output device from the low-resistive into a high-resistive state. This was for example exploited in the *MAGIC* concept from Kvatinsky et al. who proposed logic NOR, NAND, OR, AND and NOT gates based on a RESET switching [8].Of these gates only NOR and NOT were crossbar compatible in terms of the above mentioned symmetric device arrangement.

Firstly, the successful conditional switching and the input conservation are investigated with variability-aware simulations of a two-input one-output RESET-based switching gate. For this purpose, the logic operation is simulated $1\,000$ times for different voltages and the three unique input combinations HRS-HRS, LRS-HRS/HRS-LRS and LRS-LRS. Fig. 4.7 presents the corresponding readout currents for the output device and the input devices after the operation. At a voltage of $0.6\,\text{V}$ no combinations demonstrate an output switching. Only for voltages $\geq 1.2\,\text{V}$, the outputs of HRS-LRS / LRS-HRS and LRS-LRS perform a stable conditional switching towards the high-resistive state. Furthermore in these cases, both LRS and HRS inputs retain their initial state. However the output device of HRS-HRS also shows a significant current decrease above $1.2\,\text{V}$ although the calculated voltage drop approaches 0 as seen in Fig. 4.5(b). This incorrect behavior is related to an unintended input switching in these cases. Already at voltages $\geq 1.0\,\text{V}$ the resistances of both HRS inputs decreases. Not in all cases both devices perform a full SET operation as indicated by the large error bars. Instead it is more likely that as one input begins to decrease its resistance, the total voltage divider turns into an effective HRS-LRS / LRS-HRS arrangement. This leads to a steady increase of the voltage drop across the output device until it switches at last. This in turn will reduce the voltage drop across the inputs and inhibits a further input switching.

A similar observation of an undesired input switching and following output switching was found by Hoffer et al. In their experiments based on TaO_x-VCM devices they demonstrated the severe input stability for the RESET-based conditional switching with two inputs [239]. Based on the voltage divider they stated the switching threshold restriction $|V_{\text{SET}}| \geq |V_{\text{RESET}}|$. Only if this requirement is met, a successful RESET-based NOR operation can be executed. However in filamentary VCM-ReRAM the SET voltage amplitude tends to be lower as the one for a RESET, $|V_{\text{SET}}| \leq |V_{\text{RESET}}|$ [52]. This is also indicated by the 1T-1R measurements on the MAD-200 design seen in Fig. 3.3. Therefore a RESET-based switching logic cannot be reliable implemented with VCM-ReRAM devices.

For other bipolar memories which demonstrate different switching conditions the RESET-based logic operations might be successful. This was demonstrated for instance in phase-change-memory [240] or magnetic spin-torque-transfer memory [241].

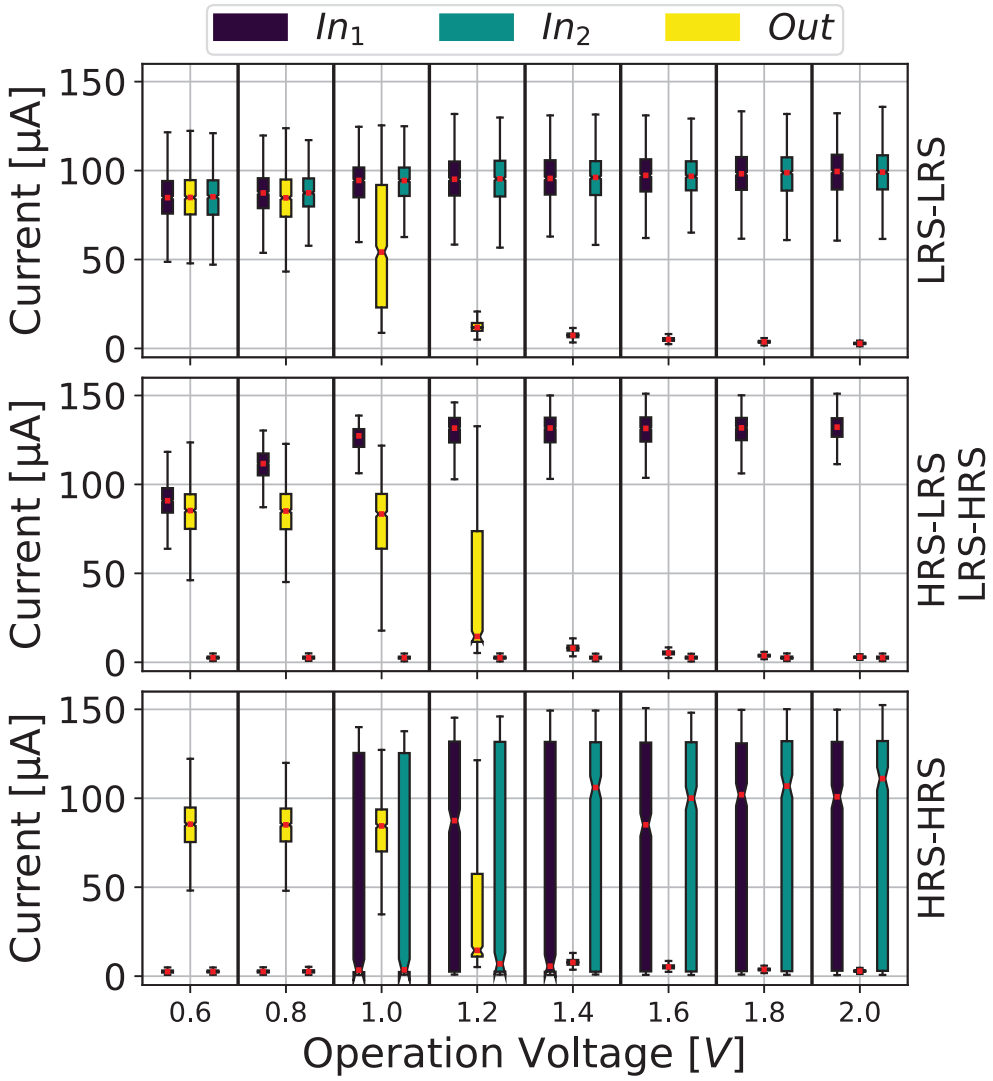


Figure 4.7: Simulated voltage-dependent readout currents after a two-input one-output RESET-based logic operation as depicted in Fig. 4.2(c). The currents are acquired by applying a readout voltage of $0.25\,\mathrm{V}$. For each input combination and operation voltage, $1\,000$ simulations are performed based on the variability parameters in Table 2.1. For all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed.

4.2.4 SET-based Stateful Switching Logic

For a SET-based two-input one-output switching, a voltage drop from the OE to AE across the output device is required. Therefore compared to the circuit displayed in Fig. 4.2(c) all device orientations have to be inverted. The modified circuitry as passive 1R and active 1T-1R structure is depicted in Fig. 4.8 with commonly connected ohmic electrodes. Parts of the measurement results in this chapter have already been published in [206].

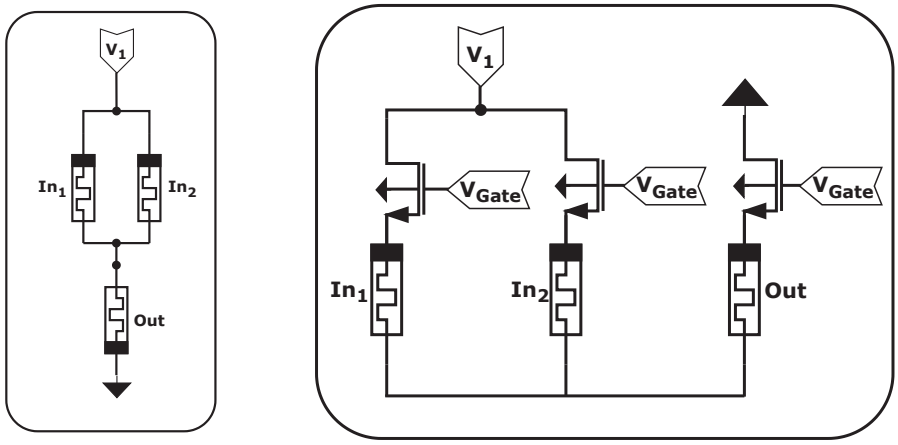


Figure 4.8: Primitive two-input one-output gate for a SET-based switching operations as (a) passive 1R and (b) active 1T-1R configuration. For the SET-based switching, the ohmic electrodes of the input and output devices are connected in common.

The general feasibility of a SET-based logic switching is investigated with similar variability-aware simulations as performed above. For the 3 unique input combinations, Fig. 4.9 depicts the corresponding readout currents for the output device and the input devices after the operation. The input combinations LRS-LRS and LRS-HRS / HRS-LRS demonstrate a significant current increase and therefore a resistance decrease for voltages of $\approx 1.0\,\mathrm{V} - 1.2\,\mathrm{V}$. This correlates to the desired logic OR. Furthermore, the HRS-HRS only causes a switching at voltages $\geq 1.4\,\mathrm{V}$. In contrast to the RESET-based operation above, the SET-based switching at least offers a stable voltage window for a correct operation output. However, the undesired input switching states remains a critical issue. At $1.2\,\mathrm{V}$, the low-resistive device in the HRS-LRS / LRS-HRS combinations undergoes at least a partial RESET. This worsens for higher voltages as the device performs a full RESET which results in an effective HRS-HRS input combination. The transient simulation results in Fig. 4.6(a) show that after the output switching roughly $\frac{2}{3}$ of the input voltage drops across the inputs in RESET polarity. This voltage induces the RESET or at least a resistance increase in the LRS device. Therefore while the SET-based switching operation is in general possible in VCM devices, the operation voltage conditions are a critical factor and have to be adjusted carefully. Otherwise, the logic operation might be performed as expected but it will violate the input state conservation. This would require additional copy steps in order to re-use logic inputs which significantly increases the operation costs in terms of latency and energy.

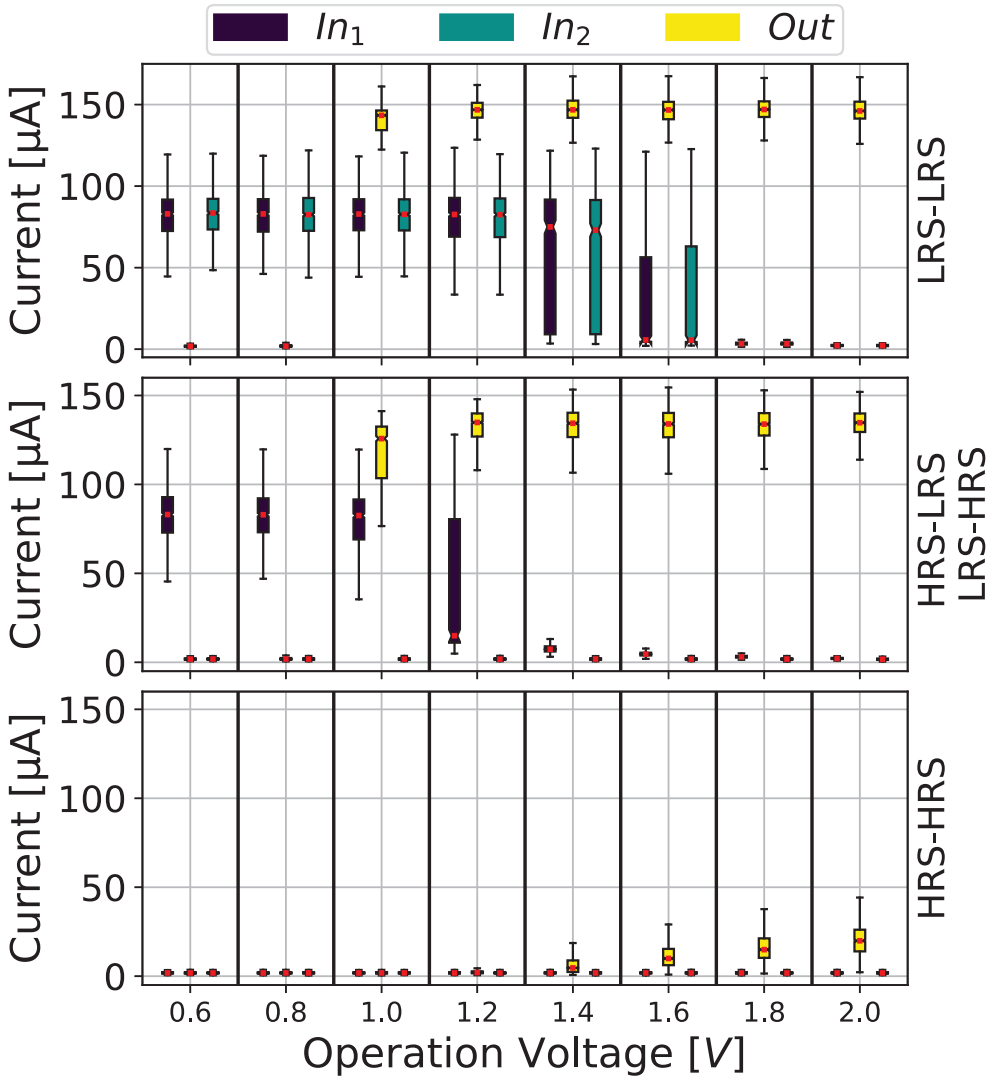


Figure 4.9: Simulated voltage-dependent readout currents after a two-input one-output SET-based logic operation as depicted in Fig. 4.8(a). The currents are acquired by applying a readout voltage of $0.25\,\mathrm{V}$. For each input combination and operation voltage, $1\,000$ simulations are performed based on the variability parameters in Table 2.1. For all varied parameters a normal distribution with $\sigma = 10\,\%$ and a truncation of 3σ is assumed.

4.2.4.1 Characterization of Operation Reliability

As a first step, the influence of device-to-device variability on the operation is investigated. For this investigation, the two-input one-output SET-based logic operation is performed in a pseudo-crossbar array in the MAD-200 test setup. Each input combination, HRS-HRS, LRS-HRS, HRS-LRS and LRS-LRS, is measured 100 times in four different device triplets. The operation voltage, V_{op}, is applied at the Signal-Lines (SL) of the input 1T-1R bitcells which connect to the transistor source and the SL of the output device is connected to ground. The Bit-Line which is shared by the ohmic electrode of all three devices in one triplet is left on floating potential to enable the voltage divider arrangement. The 1T-1R bitcells are selected via their Word-Lines where the gate voltage, V_{gate}, is applied. Before and after this operation the resistive state of each device is measured individually. The operation is performed with the voltage set (V_{op} = 1.0 V / V_{gate} = 1.5 V).

The results are depicted in Fig. 4.10. It is seen that all triplets behave very similar in the 00 input combination and none demonstrates a false output switching. However for the expected switching in 11, the triplets 1 and 2 show a high output current variability. The high-resistive outliers suggest that some outputs do not change their resistance at all. In contrast to that, triplets 3 and 4 demonstrate significantly higher post-operations currents. The non-sufficient switching issue is prominent also for the input combinations 01 / 10. Especially in Triplet 2 the outputs barely switch as indicated by the low mean values. Yet all triplets demonstrate switching issues in these two input combinations. This is confirmed by the quantitative switching rates in Table 4.3. Based on the previous readout investigation in Table 3.4 the value of 26 µA is used to separate HRS and LRS. Triplets 3 and 4 demonstrate the best output performance, however their success rate is still below 90 % even in the 11 operation. The resistive device-to-device variability influences the initial HRS variability as well as the switching threshold variability. If both are low, the switching-based logic operation is easier to reproduce.

Table 4.3: Experimental Switching Rates (SR) in [%] for a SET-based switching OR. The rates denote the number of post-operation output devices (out of 100) to be in LRS state. A LRS state is defined as a readout current value above 26 µA based on the voltage-dependent readout investigation in Table 3.4. All data sets were collected with V_{op} = 1.0 V and V_{gate} = 1.5 V.

	Switching Rates [%]			
	Triplet 1	Triplet 2	Triplet 3	Triplet 4
00	0	0	0	0
01	1	4	17	7
10	2	0	29	15
11	39	43	89	85

Two components mainly determine the voltage configuration in a SET-based switching gate: the initial resistances and the applied voltages. However the resistance ratio in filamentary VCM devices is rather limited with a ratio of $HRS/LRS \leq 100$ [221, 222]. Furthermore the resistances are set by the material stack which cannot be changed after the fabrication. In contrast to that, the applied voltages offer a high potential for improving the operation reliability. The simulation results in Fig. 4.9 indicate that the operation voltage has to be balanced between a stable and correct output switching and preventing an input switching.

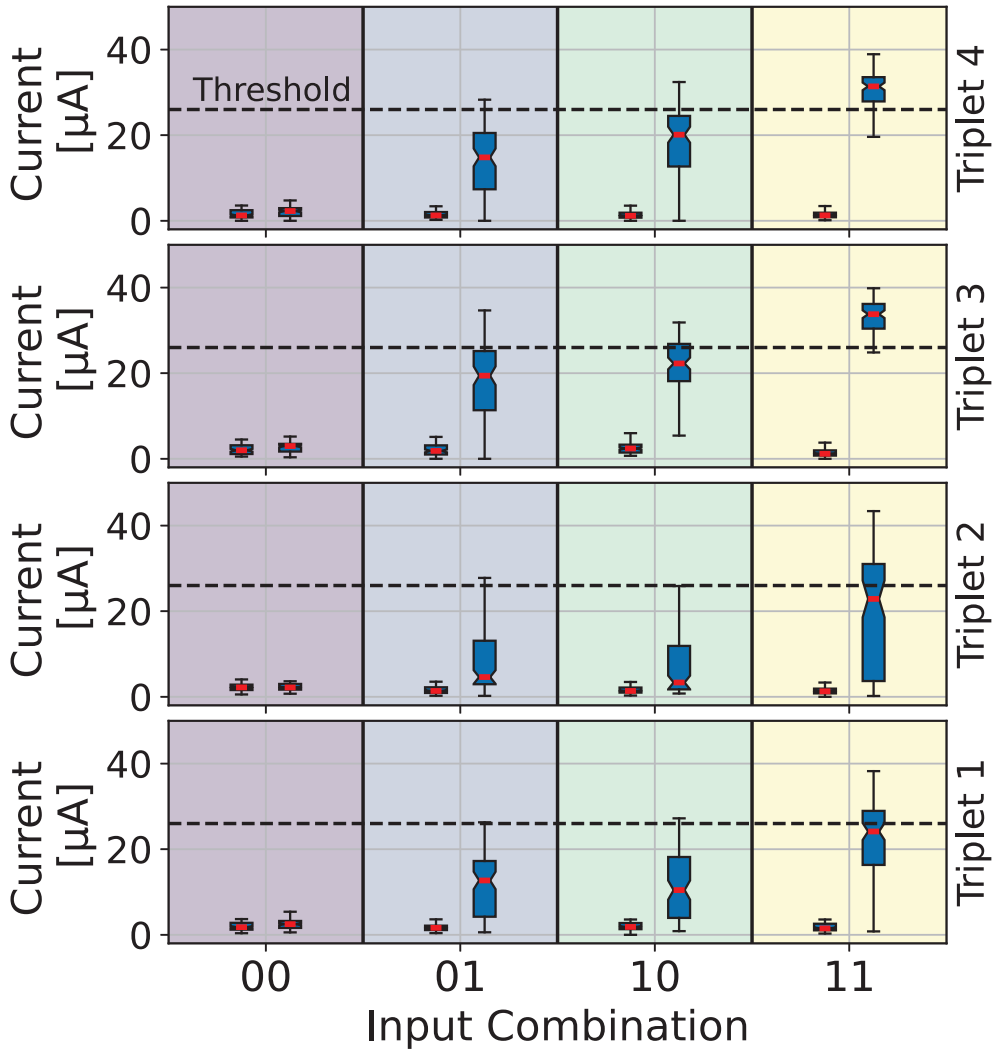


Figure 4.10: Boxplots of the experimental readout current of the output device before (left) and after (right) a two-input one-output SET-based logic circuit with 1T-1R bitcells as depicted in Fig. 4.8(b). The operations are performed in a pseudo-crossbar array on the MAD-200 setup. During the logic operation the voltage V_{op} is applied at the Signal-Lines of both inputs while the Signal-Line of the output device is connected to ground. At all three Word-Lines the voltage V_{gate} is applied. Each input combination is programmed and measured 100 times in four different device triplets for the voltage set $V_{op} = 1.0\,\mathrm{V}$ and $V_{gate} = 1.5\,\mathrm{V}$. The currents are acquired by applying a readout voltage of $0.25\,\mathrm{V}$ to the device before and after the logic operation. The threshold line marks the value of $26\,\mu\mathrm{A}$ which is based on the results in Table 3.4.

For a correct logic OR a comparably large voltage margin between $1\,\mathrm{V} - 1.4\,\mathrm{V}$ is available and only at very high voltages also the HRS-HRS input combination undergoes an output switching. However the input state conservation is a severe issue as the low resistive device in HRS-LRS / LRS-HRS might perform a RESET switching. While this second aspect is investigated later, this section focus on the reliable output switching. To investigate the voltage dependency, the previously explained measurement procedure is now performed for three different voltage sets (V_{op} / V_{gate}) in one device triplet. Here, the previous Triplet 3 from Fig. 4.10 is selected as it showed the best performance.

Fig. 4.11 depicts the corresponding readout currents as boxplots. In general, the input combination 11 (LRS-LRS) will lead to a lower LRS state after the operation compared to the inputs 01/10 (HRS-LRS / LRS-HRS). This matches with the theoretical analysis in Fig. 4.5(a) which revealed that although both cases behave very similar, the output voltage of LRS-LRS is always higher than for HRS-LRS / LRS-HRS. While the difference approaches 0 for increasing resistance ratios, the LRS-LRS input will perform a slightly better switching. Furthermore as for the device-dependent data set, no incorrect output switching was observed for the input 00 (HRS-HRS) in all voltage sets. However the expected output switching is highly voltage-dependent. At a gate voltage of $V_{\mathrm{gate}} = 1.5\,\mathrm{V}$, the outputs demonstrate overall lower current values and a large variability. It is seen that an increase of the operation voltage from $1.0\,\mathrm{V}$ to $1.2\,\mathrm{V}$ does not improve the output switching. However an elevated gate voltage of $1.7\,\mathrm{V}$ results in a significant increase of the output currents even for a low V_{op}. This qualitative result is confirmed by the successful switching rates which are displayed in Table 4.4 and are again based on the threshold value of $26\,\mu\mathrm{A}$ between LRS and HRS. All output currents above this threshold are counted as success. The switching rates indicate that the increased gate voltage, $V_{\mathrm{gate}} = 1.7\,\mathrm{V}$, improves the switching performance from below $10\,\%$ to above $90\,\%$.

Table 4.4: Experimental Switching Rates (SR) in [%] for a SET-based switching OR. The rates denote the number of post-operation output devices (out of 100) to be in LRS state. A LRS state is defined as a current value above $26\,\mu\mathrm{A}$ based on the voltage-dependent readout investigation in Table 3.4.

	Switching Rates [%]		
Inputs	$V_{\mathrm{op}} = 1.0\,\mathrm{V}$ $V_{\mathrm{gate}} = 1.5\,\mathrm{V}$	$V_{\mathrm{op}} = 1.2\,\mathrm{V}$ $V_{\mathrm{gate}} = 1.5\,\mathrm{V}$	$V_{\mathrm{op}} = 1.0\,\mathrm{V}$ $V_{\mathrm{gate}} = 1.7\,\mathrm{V}$
00	0	0	0
01	7	8	97
10	15	3	90
11	85	77	99

As the amount of performed measurements is limited, the error probability of a two-input 1T-1R SET-based switching is further investigated with variability-aware simulations. The simulations are conducted again according to the parameters in Table 2.1 and Table 2.2 for the input combinations HRS-HRS (00), LRS-HRS (10) / HRS-LRS (01), and LRS-LRS (11). In order to exploit the described framework in Section 4.2.1, the voltage drop across the output device is extracted from the simulation data. As this voltage drop is a transient quantity, i.e., it varies during the switching, only the maximum value is used for the calculation. This estimation holds based on the behavior seen in Fig. 4.6(a). If the output device does not switch its state, the maximum voltage will drop across it. However in case of a switching, the voltage drop will decrease. Therefore it can be estimated that the maximum voltage is applied for the total pulse duration for inducing a switching.

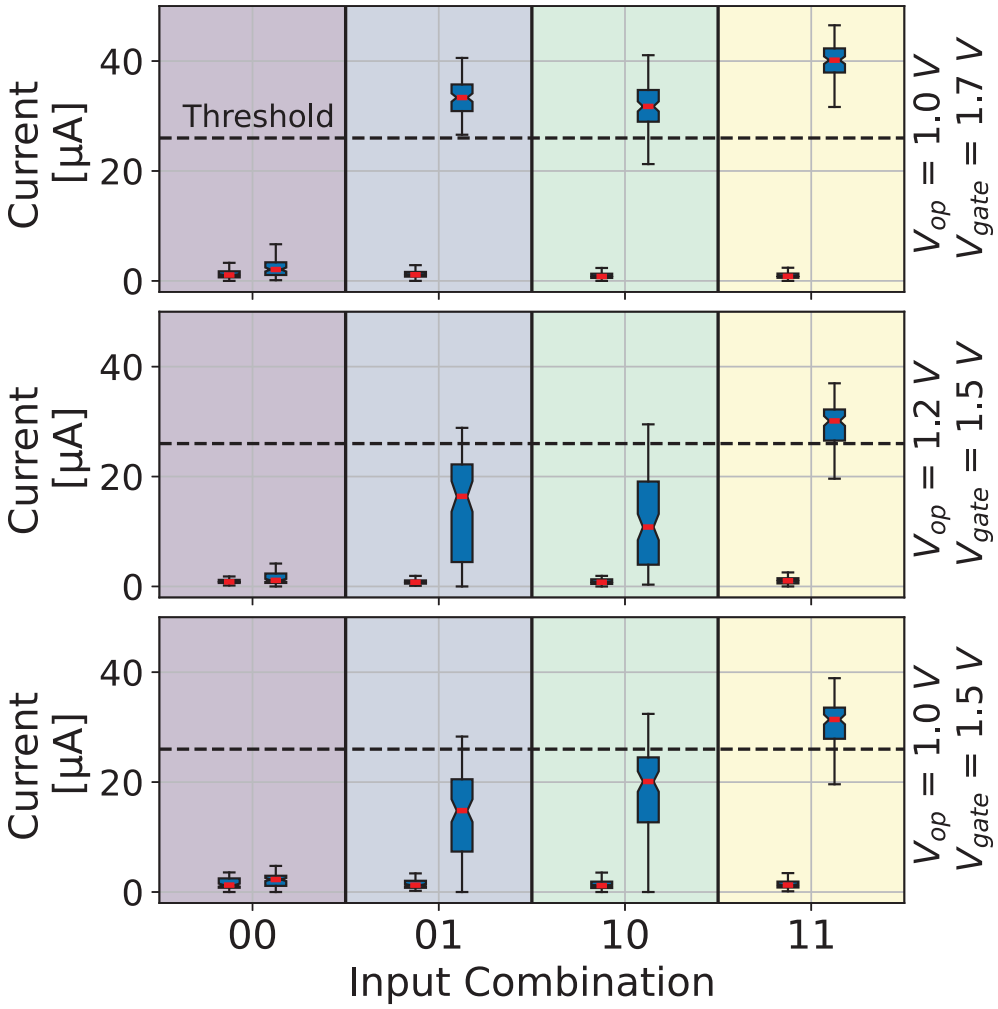


Figure 4.11: Boxplots of the experimental readout current of the output device before (left) and after (right) a two-input one-output SET-based logic circuit with 1T-1R bitcells as depicted in Fig. 4.8(b). The data is sampled similar as in Fig. 4.10. Now each input combination and voltage set is programmed and measured 100 times in the same device triplet. The threshold line marks the value of 26 µA which is based on the results in Table 3.4.

The calculated error probability in dependence of the applied operation voltage, V_{op}, are depicted in Fig. 4.12 (a). For $V_{op} < 0.8\,\mathrm{V}$, the operation voltage is not high enough to switch the output device in any case, therefore the combinations 01 / 10 and 11 yield an error probability of 1 while 00 has 0, respectively. Around $V_{op} = 1.0\,\mathrm{V}$, the output device starts to switch for the combinations 01 / 10 and 11 as expected. However the raised operation voltage comes in hand with an increased probability of an incorrect output switching in the 00 case. Yet even at high operation voltages, the input combinations 01 / 10 and 11 still demonstrate a high error probability and low reliability. While the high voltages make a switching more likely it remains a stochastic process and might not be successfully performed at this pulse configuration. The overall lowest error probability is achieved for an operation voltage $V_{op} \approx 1.1\,\mathrm{V}$ with a probability of 1×10^{-3}.

The previous simulation was conducted with a high initial ReRAM state variability by setting the distribution parameters to $\sigma = 10\,\%$ and a truncation of 3σ. In order to investigate the initial state impact, the simulations are repeated with a low variability for $\sigma = 5\,\%$ and a truncation of 2σ, respectively. The corresponding results in Fig. 4.12 (b) show a similar qualitative behavior as for the high variability but overall lower error probabilities. For narrowly confined initial states it becomes more likely that all outputs will switch uniformly. Here, a balanced error probability is achieved at a similar voltage, $V_{op} \approx 1.05\,\mathrm{V}$, however now with a probability of 1×10^{-5}. Therefore the initial device programming is crucial in order to maintain a reliable and uniform switching-based logic operation.

An additional measurement is performed for the input combination 10 (LRS-HRS) to further investigate the influence of the initial resistances states, for both inputs and the output. The data set consists of $1\,000$ operations in the voltage set ($V_{op} = 1.0\,\mathrm{V}$ / $V_{gate} = 1.5\,\mathrm{V}$). For this data set, an additional oscilloscope is used to measure the conditional voltage drop across the output device as simulated in Fig. 4.6(a). 10 exemplary transient voltage signals are displayed in Fig. 4.13. In case that the output resistance decreases due to a SET-based switching, the corresponding voltage will decrease too as the voltage divider arrangement constantly changes. Fig. 4.14 contains the output device currents after the performed operation in dependence of the input LRS state (Input 1), the input HRS state (Input 2) and the initial output HRS. The prominent observation is that only 63 of $1\,000$ operations achieved a successful switching with an output current above $26\,\mu\mathrm{A}$. This corresponds to an error rate of $94\,\%$. For the reduced number of 100 operations in Table 4.4 the same voltage set demonstrated an error rate of $93\,\%$ (01 input) and $85\,\%$ (10 input) respectively. These very high error rates strongly indicate that the lower operation voltage is significantly unreliable.

If the operation output current would depend on one of initial resistances states, the corresponding plot should have a curved shape instead of a horizontal line. However, the visual representation reveals only weak correlations. A lower LRS input and a higher HRS input seem to increase the operation output currents. Yet for the initial output HRS, no effect on the post-operation state is visible. These first observations are cross-check with the Pearson correlation coefficient. This measure ranges between -1 in case of a perfect negative linear correlation and respectively 1 for a perfect positive linear correlation. A coefficient of 0 corresponds to uncorrelated quantities. The calculated coefficients for the post-operation output current are -0.07 (input HRS) and 0.03 (initial output HRS), which indicates no correlation. The input LRS has a weak linear dependence with a coefficient of 0.28. The theoretical analysis already implied that a higher HRS/LRS ratio is beneficial for increasing the voltage drop across the output device. This in turns enhances the output switching. Improving the LRS input is one option to achieve such an increased ratio which results in the linear dependence. However the measurements do not confirm the observations based on the simulation results that a confined initial output state is beneficial for the operation reliability. These mismatching results might be caused by the differences between the simulated setup and the measurements.

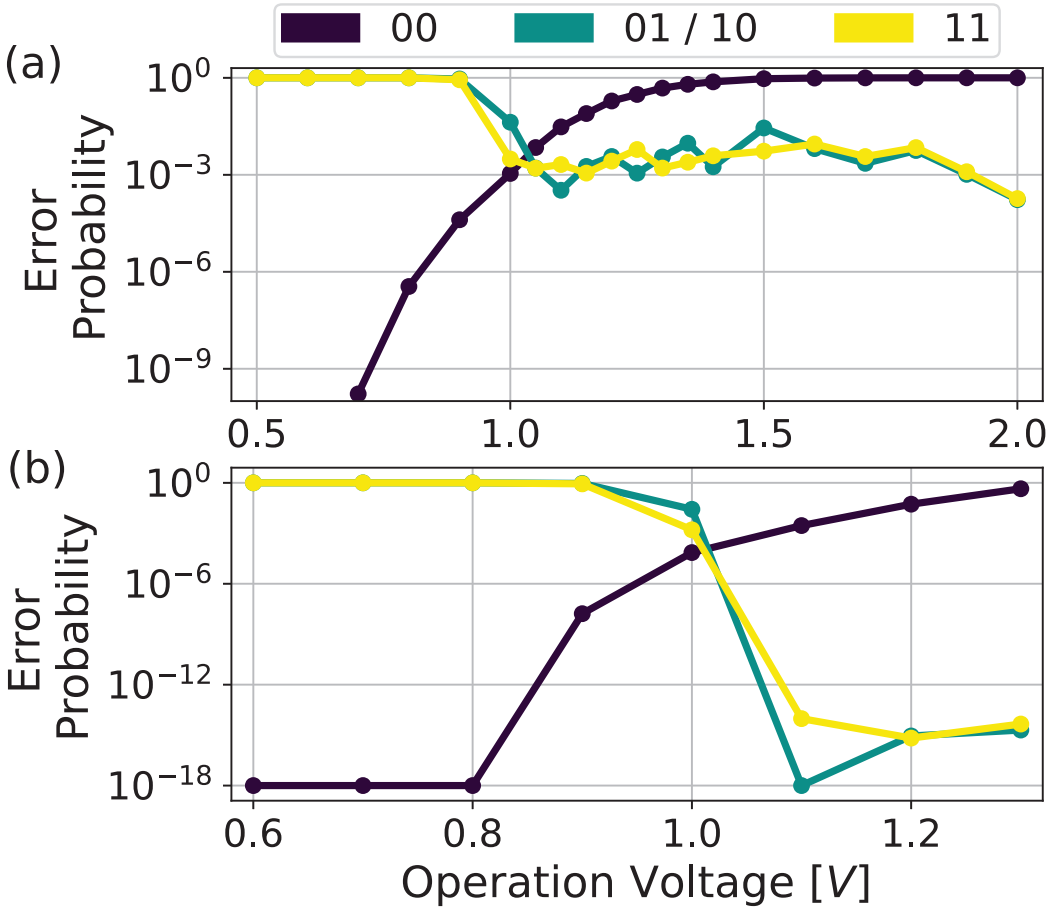


Figure 4.12: Calculated voltage-dependent error probabilities for a simulated two-device SET-based switching operation with (a) high and (b) low initial state variability. The calculation is based on the simulated output voltage distributions of 2 000 simulation runs and the framework explained in Section 4.2.1. As the input combinations HRS-LRS (01) and LRS-HRS (10) are identical in simulation, they are combined. The variability-aware simulations are conducted according to Table 2.1 and Table 2.2 with $V_{\text{gate}} = 5\,\text{V}$ For all varied parameters a normal distribution with (a) $\sigma = 10\,\%$ and a truncation of 3σ and (b) $\sigma = 5\,\%$ and a truncation of 2σ, are assumed.

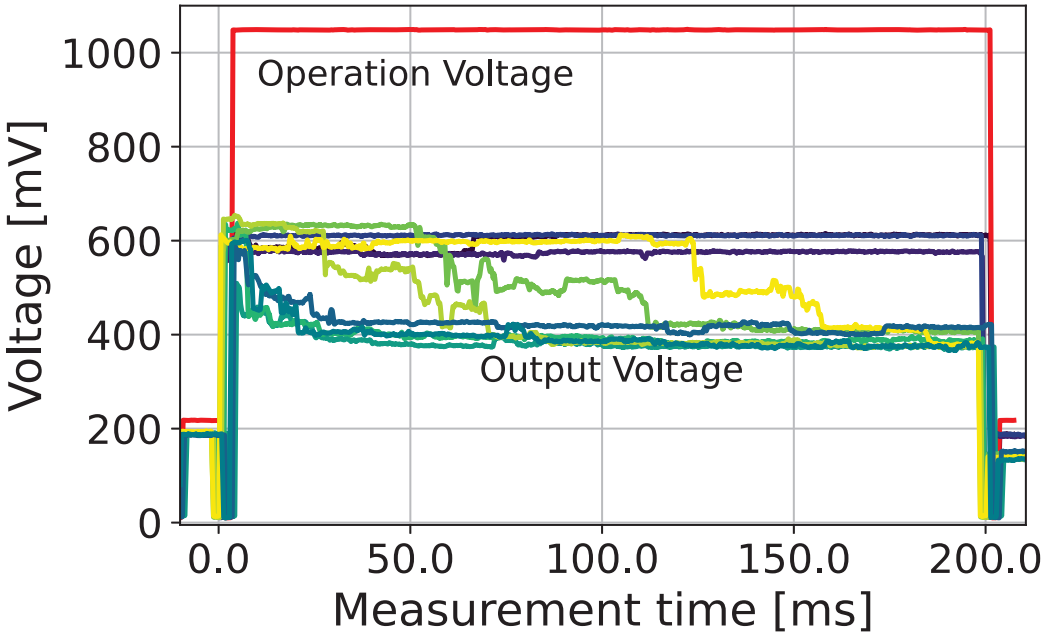


Figure 4.13: Exemplary measured transient time-voltage signals for a two-input one-output SET-based logic operation in the input combination LRS-HRS (10). The data is sampled similar as in Fig. 4.11 with V_{op} = 1.0 V and V_{gate} = 1.5 V. A decreasing output voltage corresponds to a reducing resistance and therefore SET transition occurring in the output device.

Overall, this analysis focused on achieving a reliable conditional SET-based switching in a two-device 1T-1R logic operation. The feasibility of this operation has already been demonstrated [206, 242]. However the conducted simulations and measurements have shown that although the two-device SET-based logic switching is possible, it faces severe reliability issues. Even for the best-case scenarios, the simulated error probability was $\geq 1 \times 10^{-5}$ while the best achievable experimental failure rate was $\leq 10\,\%$ for a two-input logic OR. This exceeds the corresponding error probabilities for the same logic gate implemented in a readout-based logic as depicted in Table 3.8.

One key result is that ensuring a successful switching is more critical than avoiding unintended switching events. Furthermore, the experimental measurements revealed that the gate voltage plays a more crucial role than the operating voltage. The VCM SET process is triggered by a thermal runaway which requires sufficiently high currents to flow through the device to achieve a switching. Therefore, the current limitation in a 1T-1R bitcell is a critical parameter to control when and how a switching occurs. Changing the gate voltage directly impacts the current limitation through the total 1T-1R configuration by modifying the transistor's effective series resistance. Already during the readout operation in Fig. 3.9(a) and Fig. 3.9(b), a higher gate voltage correlated to a lower series resistance and a higher current through the total 1T-1R structure. A similar effect is visible in the switching-based logic. On the other hand, reducing the gate voltage is an effective method to prevent switching by limiting the total device current.

The simulations suggest that a reduced variability in the initial ReRAM states is beneficial for a more reliable output switching. Furthermore, the experiments have shown that the device-to-device variability has a critical impact on the post-operation output states. Yet, only for the Input LRS also a quantitative correlation was found. The impact of resistive state programming is indeed obvious. A homogeneous

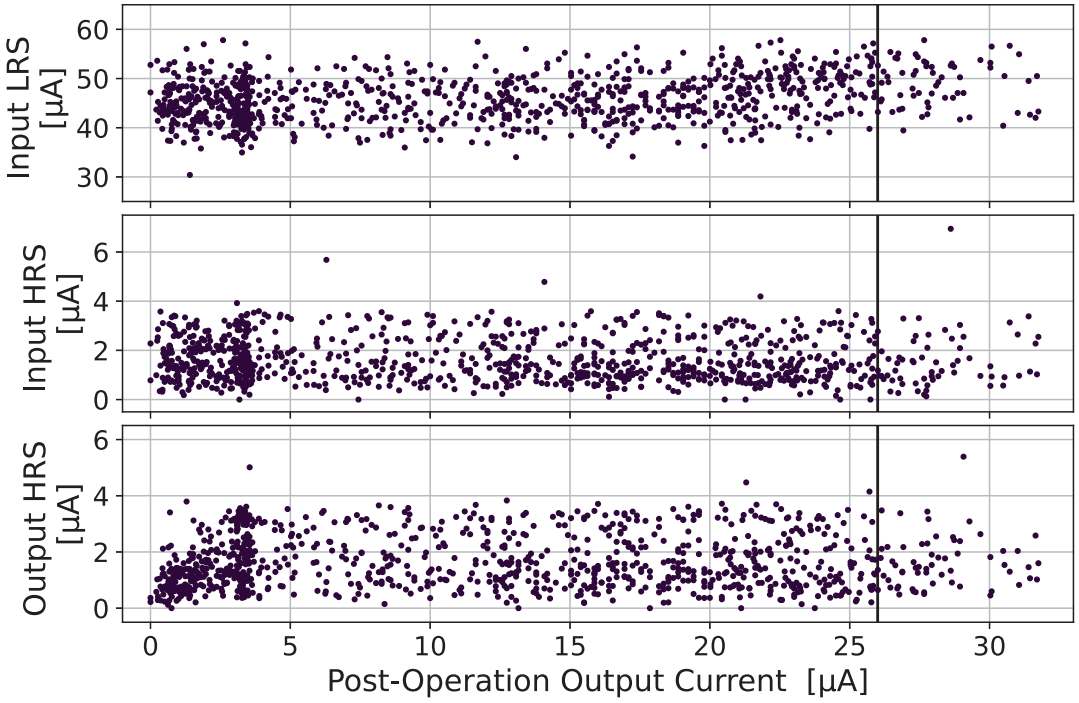


Figure 4.14: Post-operation readout current of a two-input one-output SET-based logic circuit with 1T-1R bitcells as depicted in Fig. 4.8(b) in dependence of the initial device states. The data set contains 1 000 operation for the input combination LRS-HRS (10). The sampling is conducted similar as in Fig. 4.11 with $V_{\mathrm{op}} = 1.0\,\mathrm{V}$ and $V_{\mathrm{gate}} = 1.5\,\mathrm{V}$. The currents are acquired by applying a readout voltage of $0.25\,\mathrm{V}$ individually to all three devices before and after the logic operation. The threshold line marks the value of $26\,\mu\mathrm{A}$ which is based on the results in Table 3.4.

distribution of the initial states of inputs and output will lead to a uniform conditional voltage divider configuration. As result, the voltage drop across the output device will show less variability which is in general advantageous for achieving a predictable switching behavior. However this improvement is limited. As previously discussed, the ReRAM switching is an inherent stochastic process which is challenging to predict. Therefore two devices undergo a different switching even though the identical voltage is applied to both.

As conclusion, the feasibility of a two-device SET-based logic switching operation is confirmed both in simulation and experiment. However, while the operation output is non-volatile compared to readout-based logic, achieving a correct operation output is challenging.

4.2.4.2 Input State Conservation

A correct result is one part of a successful stateful SET-based logic operation. The second important part for exploiting the advantages of this logic type is the input state conservation. Only if this feature is given, the logic states can be re-used in consecutive operations without additional copy steps. In case that the logic inputs are overwritten or corrupted during the operation, the result might still be correct, as demonstrated for instance in the IMPLY concept [6]. However if the states are destroyed, the stored information is lost has to be re-programmed after the operation which increases the overall performance cost [243].

The effects on the input devices is investigated with the same measured data set as before from Fig. 4.11. For each input combination and all three voltage sets (V_{op} / V_{gate}), 100 operations are performed. This time the readout currents of both inputs before and after the operation are depicted in Fig. 4.15. The boxplots indicate no significant input current change during the operation. However especially in the low-resistive states, the device-to-device variability between Input 1 and Input 2 is prominent. The LRS of Input 1 is in general lower than for Input 1 which is visible in the 11 (LRS-LRS) combination.

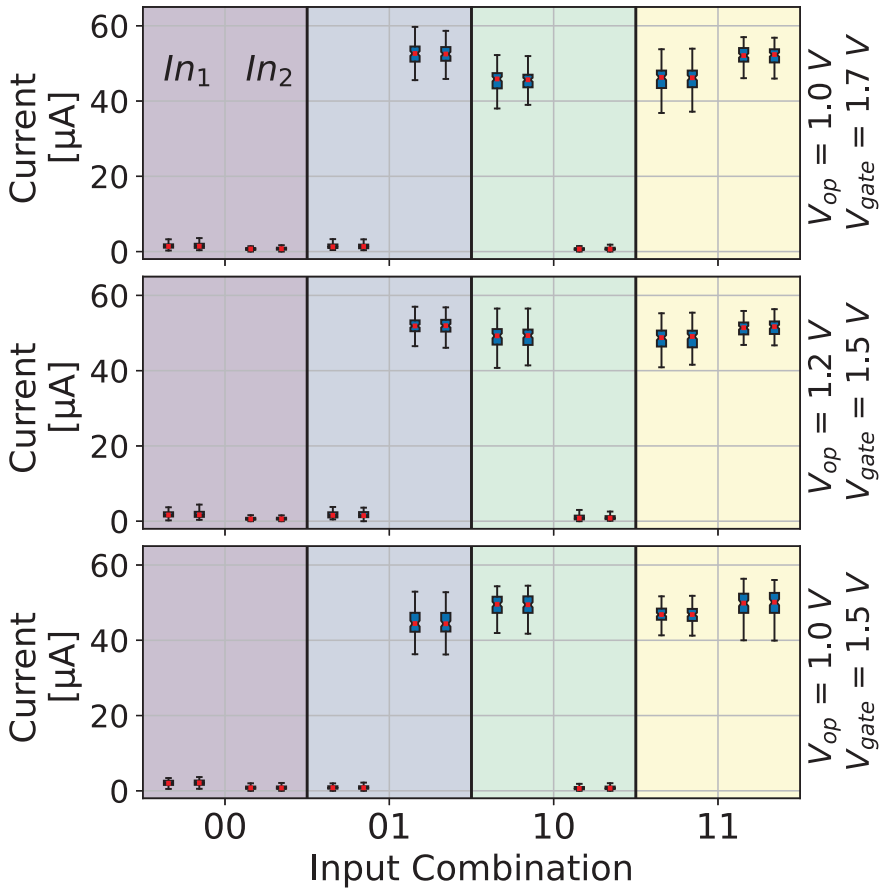


Figure 4.15: Boxplots of the experimental readout currents for both inputs before (left) and after (right) a two-input one-output SET-based logic circuit with 1T-1R bitcells as depicted in Fig. 4.8(b). The operations are performed in a pseudo-crossbar array on the MAD-200 setup. During the logic operation the voltage V_{op} is applied at the Signal-Lines of both inputs while the Signal-Line of the output device is connected to ground. At all three Word-Lines the voltage V_{gate} is applied. The currents are acquired by applying a readout voltage of $0.25\,\text{V}$ individually to both devices before and after the logic operation. Each input combination and voltage set is programmed measured 100 times in the same device triplet.

A similar observation is found for the device-dependent data set consisting of four different device triplets. The corresponding readout currents are displayed in Fig. 4.16. Especially triplet 1 demonstrates a large device-to-device variability in the low-resistive state. Furthermore, both LRS states in triplet 2 are significantly lower than $50\,\mu\text{A}$ while in the other triplets at least one LRS is higher. This variability negatively influences operation as it shifts the voltage divider. The effect is most severe for the input combinations 01 / 10 which become asymmetric for a large input device-to-device variability. The

optimization goal here is to increase the HRS / LRS ratio for the inputs by increasing the high-resistive state while lowering the LRS at the same time. In both measured data sets, no significant change in the input states for any device triplet or voltage pulse was observed. Therefore while the above investigated input correctness is critical, at least the input stability is sufficient in the MAD-200 setup.

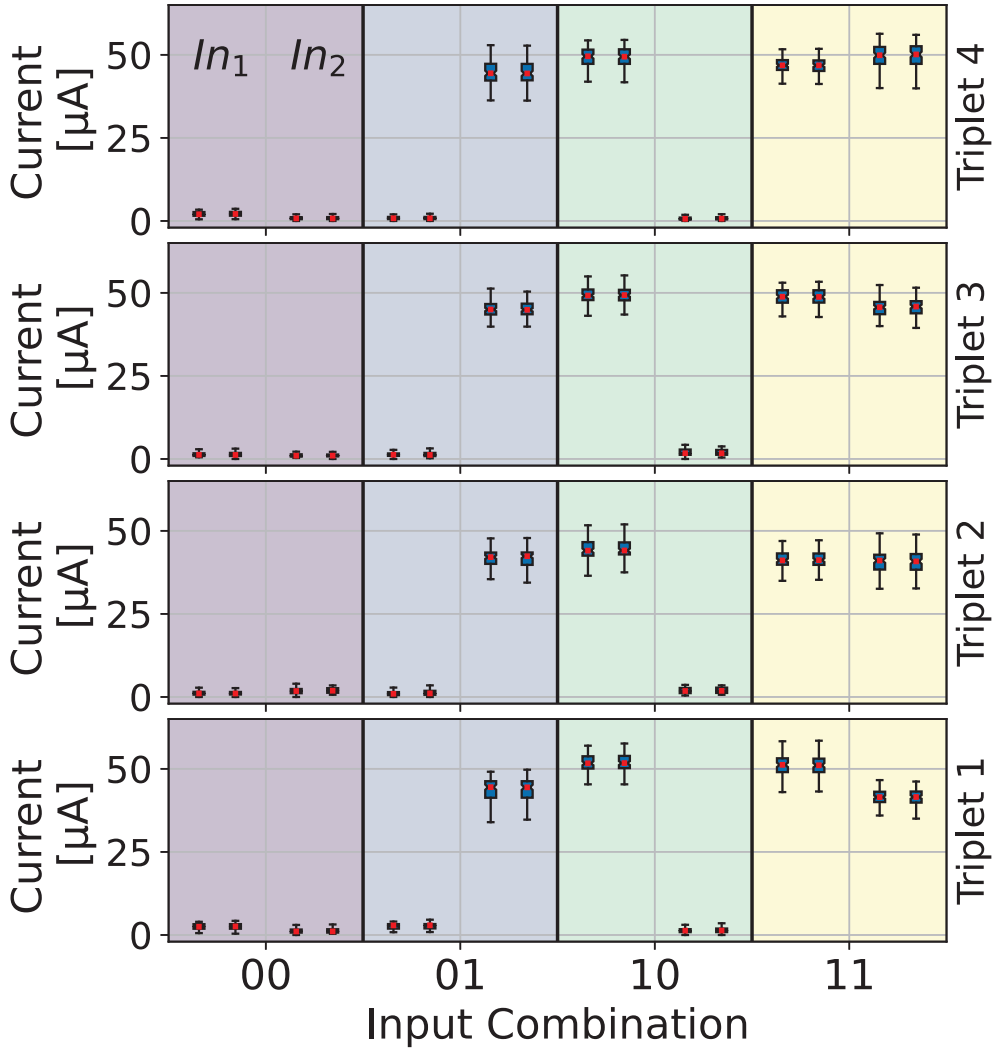


Figure 4.16: Boxplots of the experimental readout currents for both inputs before (left) and after (right) a two-input one-output SET-based logic circuit with 1T-1R bitcells as depicted in Fig. 4.8(b). The data is sampled similar as in Fig. 4.11. Each input combination is programmed and measured 100 times in four different device triplets for the voltage set V_{op} = 1.0 V and V_{gate} = 1.5 V. The currents are acquired by applying a readout voltage of 0.25 V individually to both devices before and after the logic operation.

The state conservation is also investigated with the error probability framework. For this calculation the voltage drop across the input devices is used. Although the applied input voltage, V_{op}, is constant during the pulse, the voltage divider output is a transient quantity as seen in Fig. 4.13. In case that the output device switches to the LRS, the voltage drop across it reduces which in turns increases the voltage drop across the inputs. Therefore, the voltage amplitude at the pulse end is exploited for the calculation as throughout the whole pulse duration the actual voltage is always equal or lower than the

end value. This worse-case analysis might lead to an overestimation of the input switching which is more favorable than an underestimation. Additionally, the voltage polarity now plays a role. To account for this, the single device RESET probability instead of the SET probability is used for this calculation. Due to the polarity, only the inputs which are in the low-resistive state are investigated at this point. The applied voltage might cause the HRS inputs to increase their resistance, yet this will not corrupt the stored binary information. The rest of the calculation is performed as before according to Section 4.2.1 and the results are depicted in Fig. 4.17 for both, a high and low initial state variability. The probability of a corrupted input RESET increases with higher voltages which is expected. Furthermore the LRS inputs in 10 / 01 and 11 are similar prone to a false RESET. Around the optimal value for the output correctness, $1.1\,\mathrm{V}$, the probability for a false input switching is $\approx 1 \times 10^{-5}$ for a high state variability and $\approx 1 \times 10^{-6}$ for a low variability, respectively. This is significantly lower than the output correctness, however it still implies that in 1 of $100\,000$ operations, the input is corrupted.

The investigation has shown that the input state conservation in a SET-based logic operation is not as unreliable as a correct output switching. This origins from the bipolar nature of VCM-ReRAM devices and their non-linearity, which make a false input RESET less likely then an output SET switching.

4.2.4.3 Framework Benchmarking

The issues of existing error investigation methods were discussed in the beginning of this chapter. Based on this, a novel error probability framework was proposed which has been exploited in the previous sections. An important part of proposing such a novel framework is to compare it to the state-of-the art methods, in this case is a Monte-Carlo benchmark simulation. For this, $2\,000$ samples are simulated for each voltage set as above. With this number of samples, the lowest resolvable error rate is $\frac{1}{2\,000} = 5 \times 10^{-4}$. A similar comparison was already conducted and published as part of [234]. However the simulation parameters used in that study deviate slightly from the ones in this investigation as for instance a different CMOS technology PDK was exploited. Therefore a new comparison is performed for this thesis.

The comparison of Monte-Carlo error rate and calculated error probabilities is depicted in Fig. 4.18. In general, the calculated results match in their qualitative behavior to the Monte-Carlo values as the framework is able to cover the error trends. However the calculated values deviate from the Monte-Carlo results in their specific voltage values. This is strongly visible for the input switching in Fig. 4.18(b). While the shapes of both curves match, the error probabilities are shifted approximately $0.15\,\mathrm{V}$ to higher voltage values. Therefore, the framework underestimates the occurrence of input corruptions. In contrast to that, the framework overestimates the error probability for the combination 00 output switching. These deviations can origin from multiple reasons. As the Monte-Carlo benchmark has only a limited number of samples, it might not be able to cover especially low error rates precisely. Furthermore the framework estimates the device switching by assuming a fixed voltage amplitude throughout the whole pulse duration. While calculating the convolution becomes easier, this assumption ignores transient processes within the voltage divider arrangement which are covered by the Monte-Carlo benchmarking. Both parts of the framework, the single device switching probability and the conditional voltage distributions, can be further improved for instance by a better fitting procedure. Additionally, the transient behavior could be included by splitting the voltage pulse into multiple time steps. Then the switching probability is calculated for each step individually and concatenated e.g., in a probability tree. Overall, the proposed framework is not able to replace a Monte-Carlo simulation or to predict the switching-based error probabilities perfectly. However it provides a robust and easy adaptable method to estimate logic switching errors.

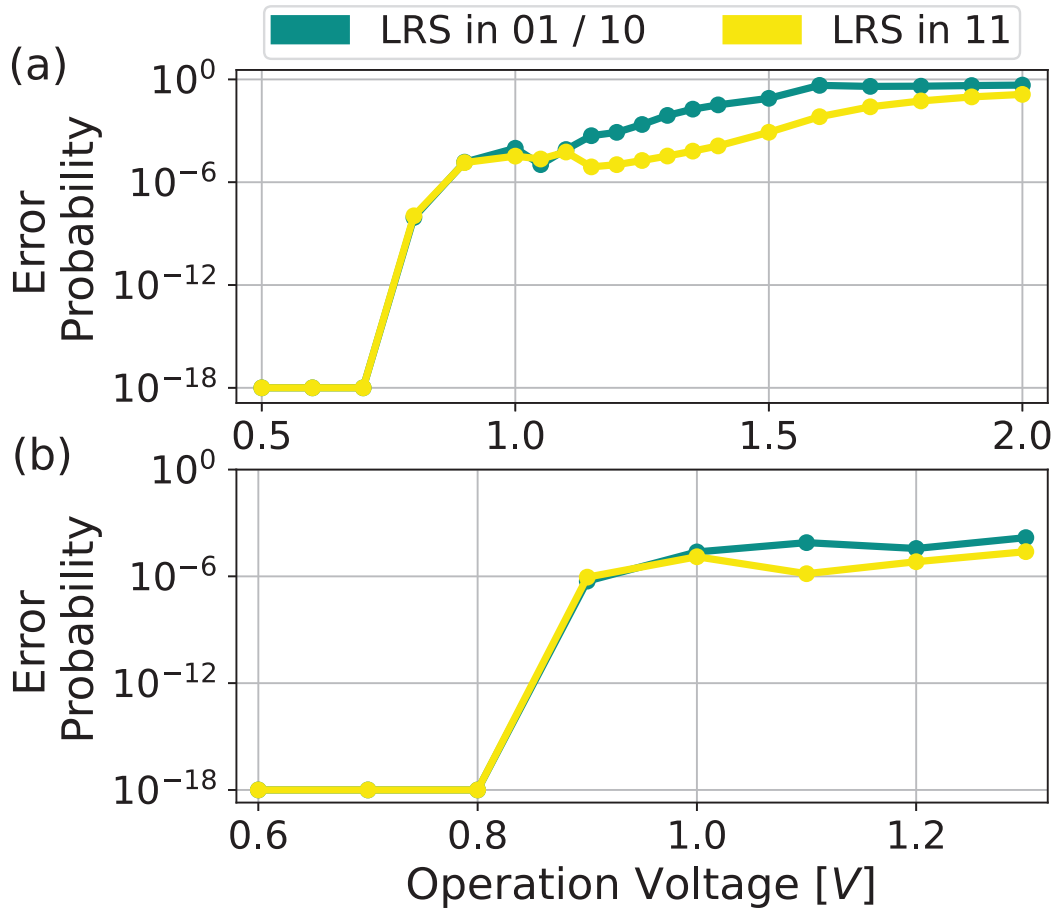
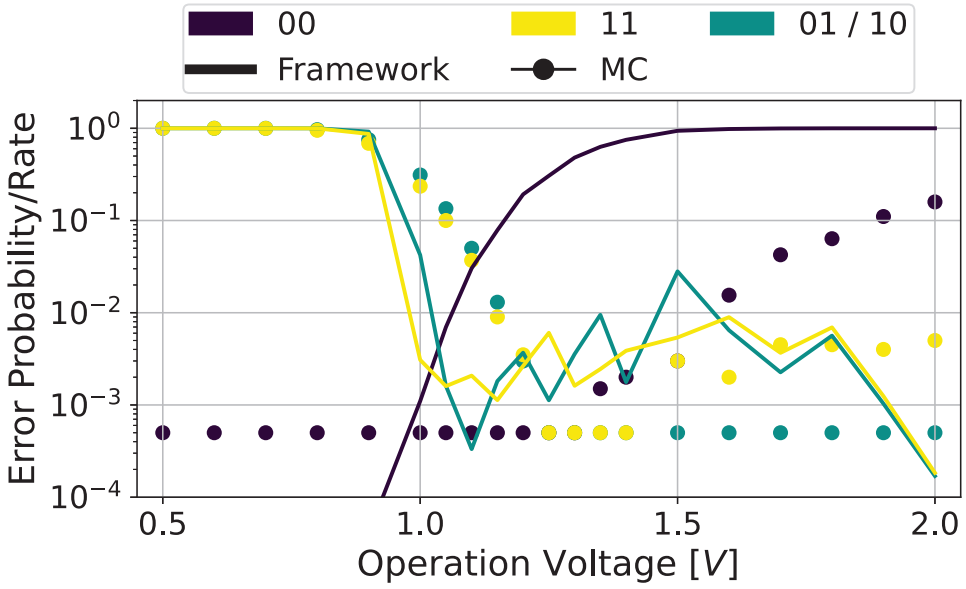
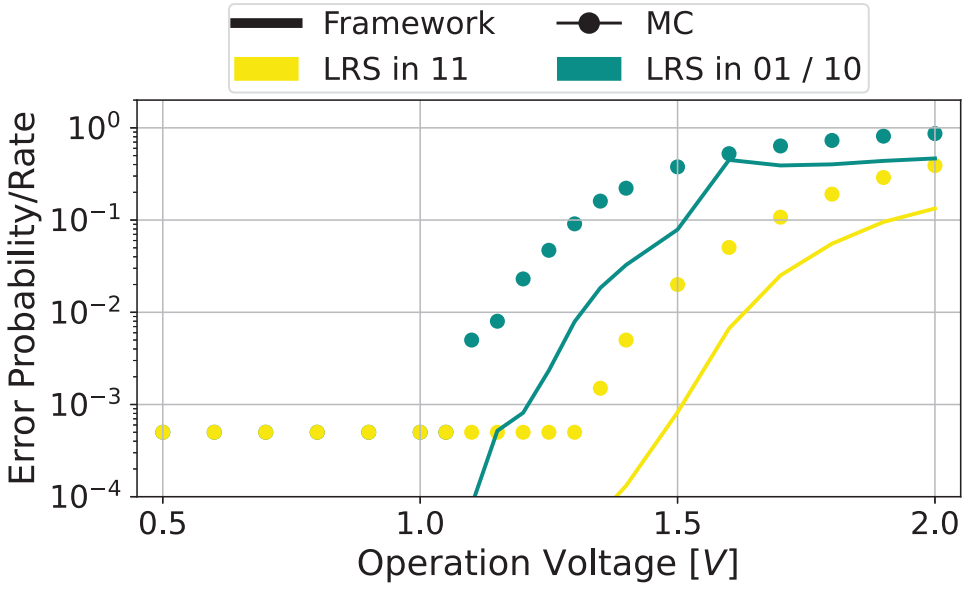


Figure 4.17: Calculated voltage-dependent input error probabilities for a simulated two-device SET-based switching operation with (a) high and (b) low initial state variability. The calculation is based on the simulated output voltage distributions of 2 000 simulation runs and the framework explained in Section 4.2.1. As the voltage drop occurs in the RESET polarity across the bipolar devices, only the low-resistive state inputs in the combinations HRS-LRS / LRS-HRS (01 / 10) and LRS-LRS (11) might perform a unintended switching. The variability-aware simulations are conducted according to Table 2.1 and Table 2.2 with $V_{\text{gate}} = 5\,\text{V}$ For all varied parameters a normal distribution with (a) $\sigma = 10\,\%$ and a truncation of 3σ and (b) $\sigma = 5\,\%$ and a truncation of 2σ, are assumed.



(a)



(b)

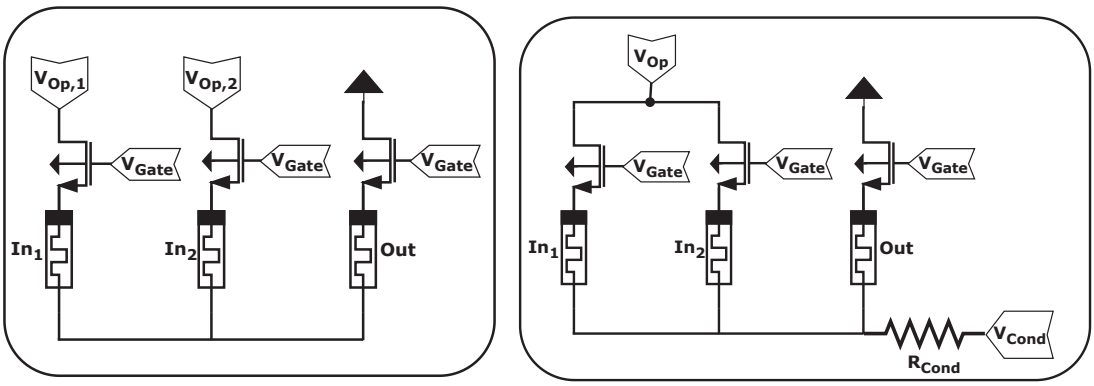
Figure 4.18: Benchmarking of the switching-based error probability framework. Displayed are calculated error probabilities for (a) the output device switching of Fig. 4.12, and (b) the input switching from Fig. 4.17, and the corresponding error rates of a Monte-Carlo simulation. The Monte-Carlo simulations are performed for 2 000 samples according to Table 2.1, Table 2.2 and normal distributed parameters with $\sigma = 10\,\%$ and a truncation of 3σ. An error in the Monte-Carlo simulation is defined as a device in expected HRS which shows a readout current above $26\,\mu\text{A}$ which is based on the results in Table 3.4. Vice versa an expected LRS device is counted as error if its readout current is below this threshold.

4.2.5 Extensions of the Two Input Switching-based Logic

The investigated two-input SET-based logic can be extended by increasing the number of voltage inputs. In the original concept for a logic NOR (RESET-based) and logic OR (SET-based), only three different potentials are required, the operation voltage, V_{op}, ground potential, *GND*, and the transistor gate voltage, V_{gate}, as both logic inputs receive the same voltage input. However if this is separated and $V_{\mathrm{op,1}}$ is applied to Input 1 while $V_{\mathrm{op,2}}$ is applied at Input 2, a logic NIMP operation can be implemented [239, 242]. The logic NIMP itself is of limited usage, but the concatenation of two NIMPs resembles a logic XOR. Furthermore, performing a logic NIMP with Input 1 in LRS is similar to a logic NOT of Input 2. Both truth tables are depicted in Table 4.5.

Table 4.5: Logic truth table for a switching-based two-input logic NIMP, NOT and XOR. The emphasized subsets for $A = 1$ resembles the logic inversion of input B, $1 \not\rightarrow B \equiv \overline{B}$.

A	**B**	NIMP ($A \not\rightarrow B$)	NIMP ($B \not\rightarrow A$)	XOR (($A \not\rightarrow B$) + ($B \not\rightarrow A$)) $\equiv A \bigoplus B$
0	0	0	0	0
0	1	0	1	1
1	**0**	**1**	0	1
1	**1**	**0**	0	0



(a) (b)

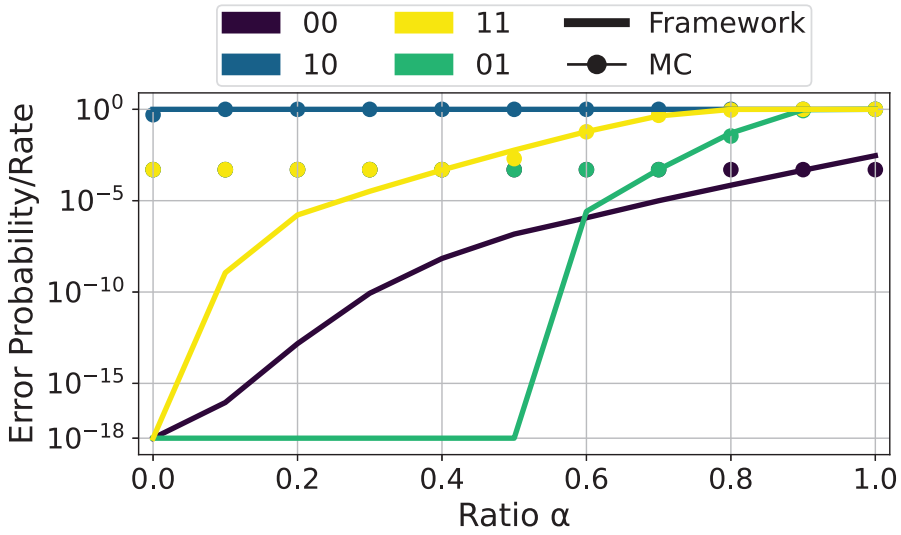
Figure 4.19: Primitive two-input one-output gates for the extended SET-based switching operations (a) NIMP logic according to Hoffer et al. [239] and (b) Three-Memristor-Stateful Logic (TMSL) according to Huang et al. [231].

The logic NIMP operation offers an additional degree of freedom compared to the standard two-device SET-based switching as two, instead of one, operation voltages are applied. For this, the parameter α is introduced which describes the ratio between $V_{\mathrm{op,1}}$ and $V_{\mathrm{op,2}}$, $\alpha = \dfrac{V_{\mathrm{op,2}}}{V_{\mathrm{op,1}}}$. The influence of this ratio is investigated again in variability-aware simulations with a similar setup as before. While $V_{\mathrm{op,1}}$ is set constant to $1.1\,\mathrm{V}$, the ratio α is varied between $[0, 1]$ which correlates to $V_{\mathrm{op,2}} \in [0, 1.1]\mathrm{V}$. Based on

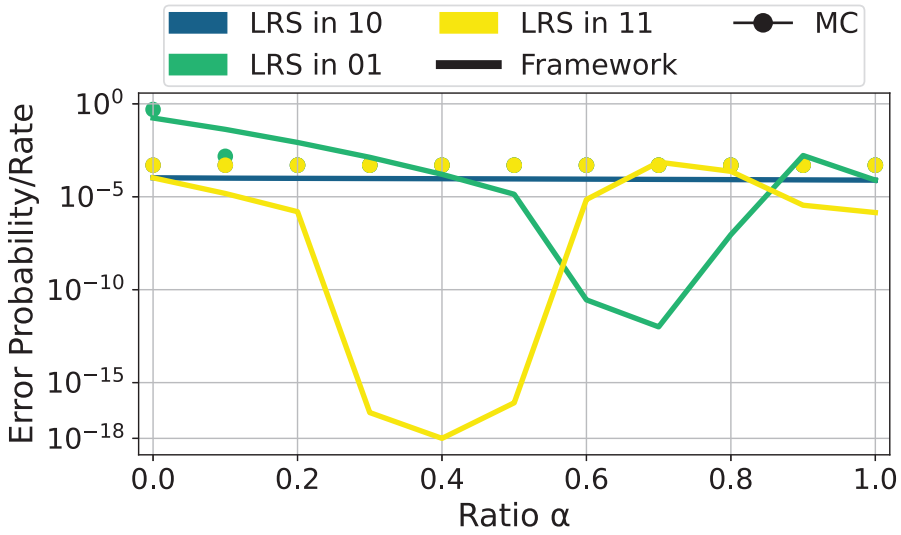
the simulation framework in Table 2.1 and Table 2.2, 2 000 operations are simulated each. Here, a low initial state variability is selected for all ReRAM devices. From the simulations, both the error probabilities according to the proposed framework as well as the pure Monte-Carlo error rates, are extracted.

The output switching probabilities are depicted in Fig. 4.20(a). In this case, the framework results agree well with the Monte-Carlo benchmarking. The combination 10, which is expected to switch, shows a stable output switching for all values of α. However as α increases, also the 01 and 11 outputs become more likely to switch. If α gets closer to 1, $V_{\mathrm{op},2}$ approaches $V_{\mathrm{op},1}$ and at $\alpha = 1$, both input voltages are equal which is identical to the logic OR. In this case, the output might switch incorrectly also in the 00 combination. This issue of a false output switching was already visible for the SET-based OR in Fig. 4.12. The results imply that α should be selected as 0 in order to optimize the output correctness. However in a stateful switching-based logic, also the input state conservation is important. The corresponding results are displayed in Fig. 4.20(b). It is seen that for α values close to 0, the voltage drop across the LRS input in the 01 combination is critically high which might induce a false input RESET. In contrast to that, all other low-resistive inputs are permanently stable. Therefore, α has to be selected low enough to benefit from an improved output correctness while being high enough in order to prevent the 01 input corruption. Hoffer et al. solves this trade-off by selecting $\alpha = \frac{1}{3}$ [239]. Based on the data in Fig. 4.20, $\alpha = 0.3$ is selected for all following NIMP operations.

A further extension of the two-input switching-based logic was proposed as *Fast and Energy-Efficient Logic in Memory* (FELIX) by Gupta et al. [230]. In this concept, both the RESET-based and SET-based switching are combined in order to implement a NOR, NOT, NAND and OR as single executable logic gates. However this RESET/SET combination suffers from the same voltage issues as the standalone RESET-based operation above. Furthermore even if the thresholds are adjusted to stabilize the RESET-based switching, the SET-based process will worsen on the other hand. Therefore this mixed-process implementation faces at least one reliability issue in all cases. In contrast to that, the concept of *Three-Memristor-Stateful Logic* (TMSL) by Huang et al. implemented two input resistive AND and NAND gates, both with a SET-based switching operation [231]. A primitive TMSL gate can be implemented either in passive 1R or active 1T-1R structures and consists of three memristive bitcells, two logic inputs and one logic output. Similar to the NIMP concept, four different voltage inputs are required, V_{op}, *GND*, V_{gate}, and a condition voltage, V_{cond}. Similar as for the standard two-input OR, both ReRAM receive same input voltage, V_{op}, while the second input potential, V_{cond}, is applied at an ohmic resistor, R_{s}, which is connected to the common ReRAM electrode of all devices as depicted in Fig. 4.19(b). By varying the pulse conditions, i.e., pulse width and amplitude of V_{cond} and V_{op}, either a logic AND or NAND is executed. It has to be noted that the original two-device one-resistor TMSL concept implemented both boolean function based on the assignment HRS = 1 and LRS = 0. For the selected assignment in this thesis, HRS = 0 and LRS = 1, the operations resemble an OR instead of AND and NOR instead of NAND. This nomenclature is also used in the following.



(a)



(b)

Figure 4.20: Calculated error probabilities and benchmarking for a simulated two-device SET-based NIMP according to the circuit in Fig. 4.19(a). The parameter α refers to the ratio between the applied operation voltages, $\alpha = \frac{V_{\mathrm{op},2}}{V_{\mathrm{op},1}}$. Displayed are calculated error probabilities for (a) the output, and (b) the input switching, as well as the corresponding error rates of a Monte-Carlo simulation. The simulations, both for the Monte-Carlo benchmarking as well as for the framework calculations, are performed with 2 000 samples according to Table 2.1, Table 2.2 and with normal distributed parameters with $\sigma = 5\,\%$ and a truncation of 2σ. An error in the Monte-Carlo simulation is defined as a device in expected HRS which shows a readout current above $26\,\mu\mathrm{A}$ at $V_{\mathrm{read}} = 0.2\,\mathrm{V}$ which is based on the results in Table 3.4 Table 3.4. Vice versa an expected LRS device is counted as error if its readout current is below this threshold.

The resistance R_s adds an additional parameter for optimizing the TMSL operation [232]. While the pulse conditions can be modified and optimized during the operation, the resistance value of R_s is set already in the design. Originally, a value of $R_{LRS} < R_s \approx R_{HRS}/2$ was proposed as optimal value [231]. For the HfO_2-based VCM 1T-1R devices in the MAD-200, this results in $R_s \approx 50\,k\Omega - 75\,k\Omega$ based on the measured equivalent resistances in Fig. 3.11(a). However such an assignment of ohmic resistance values is vague due to the non-linear VCM characteristics. Shen et al. experimentally demonstrated the logic NOR gate based on this circuit in 2019 with TaO_x/HfO_2-based 1T-1R devices [244]. In the study they selected $R_s = 1\,M\Omega$ based on the resistances $R_{LRS} \approx 100\,k\Omega$ and $R_{HRS} \approx 100\,M\Omega$. Zhang et al. published in 2021 1T-1R experimental results for the logic NOR based on a similar material stack, TaO_x/HfO_2 [245]. Their reported device resistances are $R_{LRS} \approx 100\,k\Omega$ and $R_{HRS} \approx 10\,G\Omega$ but unfortunately, they have not stated the value of R_s in their work.

Due to the complex circuit structure with an additional series resistance, it was not possible to measure the operation in the MAD-200 or Elmos setup. However the two-device one-resistor SET-based operation is investigated in simulation. At first, the simulations are performed without any variability based on Table 2.1 and Table 2.2 with V_{op} /$V_{cond} \in [0.25\,{-}2\,]$ V and $R_s \in [5\,,50\ 100\,]\,k\Omega$. The output current change during the operation, $\left|\frac{I_{after} - I_{before}}{I_{before}}\right|$ for each combination is visualized in Fig. 4.21.

Only for $R_s = 5\,k\Omega$ a significant current change is observed in the simulations. The combination of a high V_{op} and a low V_{cond} results in an output switching for the input combinations 01, 10 and 11, which corresponds to a logic OR. Vice versa, at low V_{op} and high V_{cond}, solely the output of 00 switches which implements a logic NOR. For both of these two voltage regions, variability-aware simulations are conducted to calculate their error probabilities which are displayed in Table 4.6. The results reveal a pronounced asymmetry between the gates along with a substantial asymmetric dependence on the inputs. As seen before, intended switching events become more reliable at higher voltages. Yet this comes in hand with an increased risk of unintended switching events. Conversely, lower voltages can minimize false outputs at the expense of potentially hindering desired switching events. The NOR operation obtains error probabilities similar as for the two-device SET-based OR in Fig. 4.12. In contrast to this, the two-device one-resistor SET-based OR operation is very unstable for the inputs 01 and 10. This introduces a reliability issue which was already found by Disfani et al. [232].

Table 4.6: Calculated voltage-dependent error probabilities for the simulated two-device one-resistor SET-based operation outputs. The lowest computable probability was 1×10^{-18}.

V_{cond} [V]	V_{op} [V]	Gate	Input combinations 00	01	10	11
0.25	1.4	OR	1×10^{-18}	9.67×10^{-1}	9.67×10^{-1}	1.91×10^{-1}
0.25	1.5	OR	1×10^{-18}	8.1×10^{-1}	8.1×10^{-1}	1.72×10^{-5}
1.4	0.25	NOR	5.6×10^{-5}	6.9×10^{-7}	6.88×10^{-7}	1.13×10^{-4}
1.5	0.25	NOR	2.89×10^{-7}	4.88×10^{-6}	4.88×10^{-6}	1.13×10^{-4}

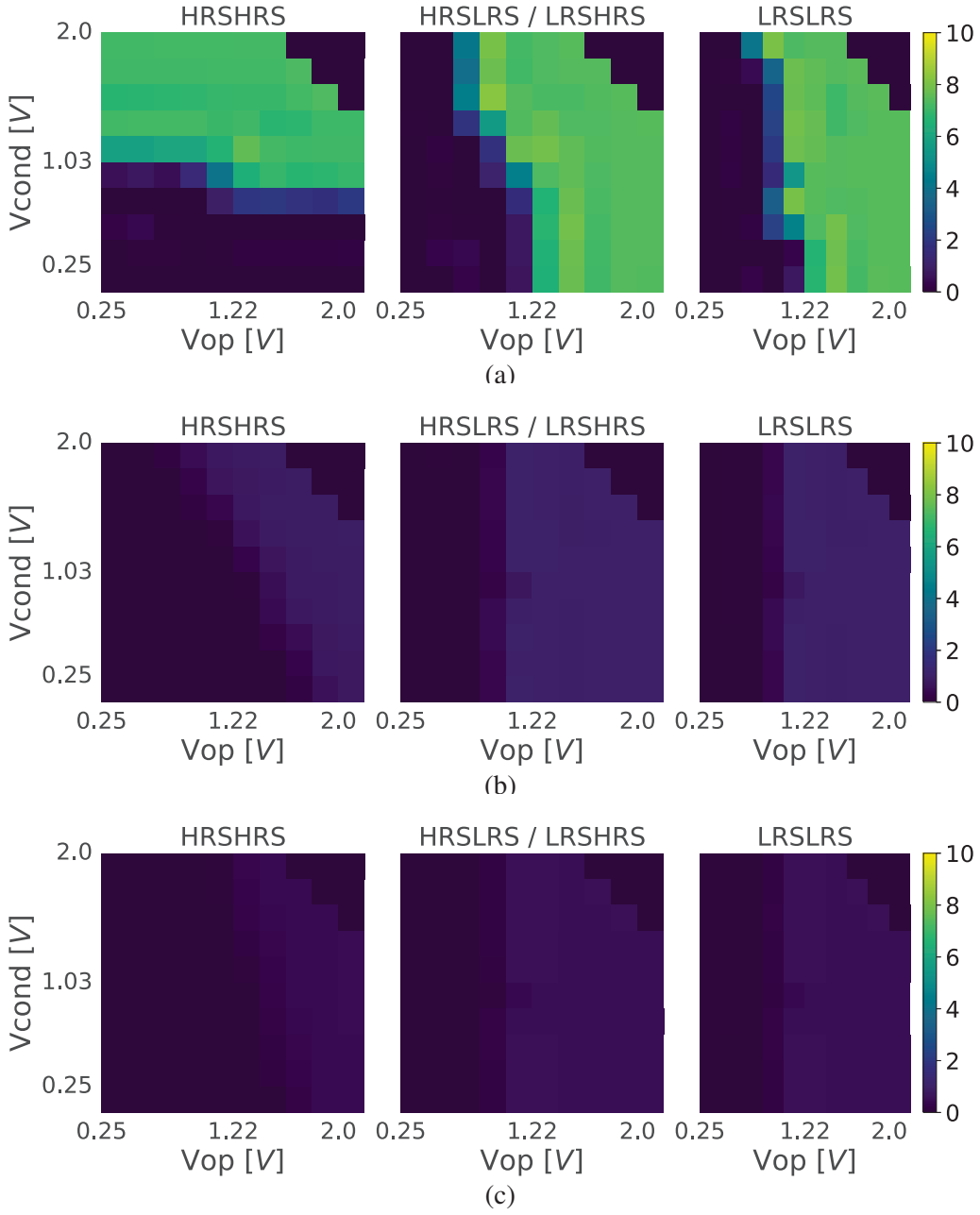


Figure 4.21: Simulated output current change in two-device one-resistor logic operation. The current change is calculated as $\left|\frac{I_{\text{after}} - I_{\text{before}}}{I_{\text{before}}}\right|$, with I_{before},I_{after} being the device readout currents before/after the operation evaluated at $V_{\text{read}} = 0.2\,\text{V}$. The simulations are conducted for R_s (a) = $5\,\text{k}\Omega$, (b) = $50\,\text{k}\Omega$ and (c) = $100\,\text{k}\Omega$.

Table 4.7 additionally lists the probabilities for an input switching during the operations. Both, the SET and RESET single device switching probabilities are required in this logic scheme. As before for the logic OR, the voltage drop across the inputs occurs in the RESET direction. Therefore while high-resistive state inputs might drift to even higher resistances values, only LRS inputs might undergo an unintended switching which corrupts their information. However during the logic NOR operation, the polarity is common for all devices, inputs and output, with a voltage drop in SET direction. As result, the high-resistive state inputs can perform a similar SET switching, yet unintended, as the logic output. Here, the low V_{op} which is applied at the input's active electrode leads to a reduced voltage drop across the inputs and thereby inhibit a false input switching. The calculated probabilities suggest that the input conservation is not as critical as the correct output switching which was already seen before for standard the two-device SET-based operation. As the voltages are applied at both electrodes, ohmic and active, during the two-device one-resistor operation, the effective potential differences across the input devices which cause a false switching are reduced.

Table 4.7: Calculated error probabilities for the simulated two-device one-resistor SET-based operation inputs. During the OR operation only inputs in LRS might switch as the voltage drop occurs in the RESET direction. Vice versa, in the logic NOR a voltage drop across the inputs is in SET polarity, therefore only HRS devices might undergo a switching. The lowest computable probability was 1×10^{-18}.

V_{cond} [V]	V_{op} [V]	Gate	HRS in 00	HRS in 01 / 10	LRS in 01 / 10	LRS in 11
0.25	1.5	OR	-	-	2.5×10^{-16}	1×10^{-18}
1.5	0.25	NOR	7.7×10^{-12}	1×10^{-18}	-	-

A conclusion for the two-device one-resistor SET-based switching is that it can execute a NOR logic gate with a similar reliability as the standard two-device SET-based OR. However, the corresponding OR implementation with a resistor is severely unreliable. Furthermore the two-device one-resistor implementation requires a dedicated circuit design with an additional analog element in each crossbar line. If both operations are heavily used, it can be worth to integrate this resistor in order to execute a logic NOR as two-device one-resistor logic while relying on the standard two-device SET-based OR.

4.2.6 Transistor-specific Logic Operations

As previously discussed, the applied gate voltage, V_{gate}, has a significant impact on the operation result as it limits the total 1T-1R current. For the readout-based logic this was exploit in Section 3.2.5 to implement transistor-specific logic operations in which the transistor gate voltage is used as additional binary input, $GND = 0$ and $V_{gate} = 1$. The switching-based logic operations can be modified in a similar manner. With this, two-device SET-based OR operation turns into a concatenated 1T-1R specific AND, $[(X * A) + (Y * B)]$ with $A, B \in \{HRS, LRS\}$ as ReRAM states and $X, Y \in \{0, V_{DD}\}$ as transistor gate voltages.

The seven unique input combinations, 0-HRS 0-HRS, 0-HRS 0-LRS, 0-LRS 0-LRS, 1-HRS 0-HRS, 1-HRS 0-LRS, 1-LRS 0-HRS and 1-LRS 0-LRS, are measured in the MAD-200 setup. The measurement is performed with the best voltage of Fig. 4.11, $V_{op} = 1.0\,\mathrm{V}$ and $V_{gate} = 1.7\,\mathrm{V}$. The readout currents of both inputs and the output device are depicted as boxplots in Fig. 4.22. For any operation, no input switching is observed. Furthermore, the outputs of 0-HRS 0-HRS, 0-HRS 0-LRS, 0-LRS 0-LRS, 1-HRS 0-HRS, 1-HRS 0-LRS remain in the high-resistive state as expected. Only in the two other combinations, 1-LRS 0-HRS and 1-LRS 0-LRS which should end in a logic 1 output, a current increase is clearly visible. However, only 65 of 100 (0-HRS 1-LRS) and respectively 93 of 100 (1-LRS 0-LRS) outputs achieve post-operation currents above the threshold of $26\,\mu\mathrm{A}$. The output LRS in the transistor-specific operation is not as distinctive as in Fig. 4.11 for the two-device SET-based OR where both input gates are activated. This decrease in reliability might be caused by the current limitation during the switching operation. If both inputs are activated, the current flowing through the output device yields to the sum of both individual 1T-1R currents. If only one input gate is activated, the switching current is limited.

In addition, the same input combinations are also simulated in the variability-aware setup similar as before. The resulting current boxplots for the output and inputs are depicted in Fig. 4.23. The resulting low-resistive state after the operation is higher than the programmed input LRS which is the opposite behavior as seen for the measurements in Fig. 4.22. The simulations use a higher gate voltage which can lead to an increase current conduction and the increased output switching. Based on the simulated voltages, the error probabilities for a correct output switching and a false input RESET are calculated and listed in Table 4.8. The results suggest that the implementation of transistor-specific logic is theoretically possible in a switching-based operation. This operation is more reliable than the standard two-device SET-based switching. Yet, the practical implementation of this transistor-specific logic has demonstrated vulnerable as it requires a careful adjustment to achieve a sufficient SET switching.

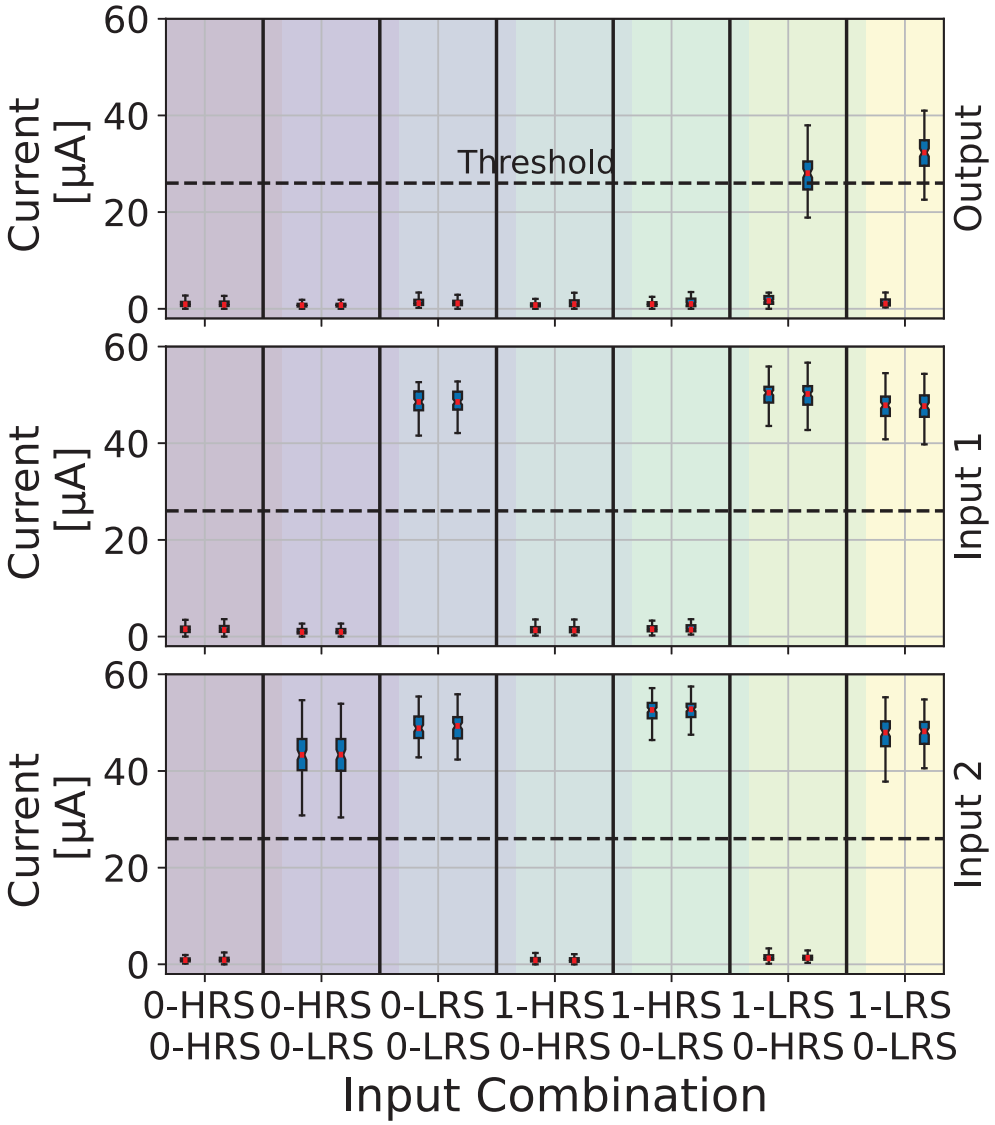


Figure 4.22: Boxplots of the experimental readout currents before (left) and after (right) transistor-specific two-input one-output SET-based logic operations as depicted in Fig. 4.8(b). The operations are performed in a pseudo-crossbar array on the MAD-200 setup. During a logic operation the voltage $V_{\text{op}} = 1.0\,\text{V}$ is applied at the Signal-Lines of both inputs while the Signal-Line of the output device is connected to ground. For an input of 1, the corresponding input Word-Line is set to $V_{\text{gate}} = 1.7\,\text{V}$, otherwise to GND. At the Word-Line of the output device, $1.7\,\text{V}$ are applied in any case. The currents are acquired by applying a readout voltage of $0.25\,\text{V}$ individually to all three devices before and after the logic operation. Each input combination is programmed and measured 100 times in the same device triplet. The threshold line marks the value of $26\,\mu\text{A}$ which is based on the results in Table 3.4.

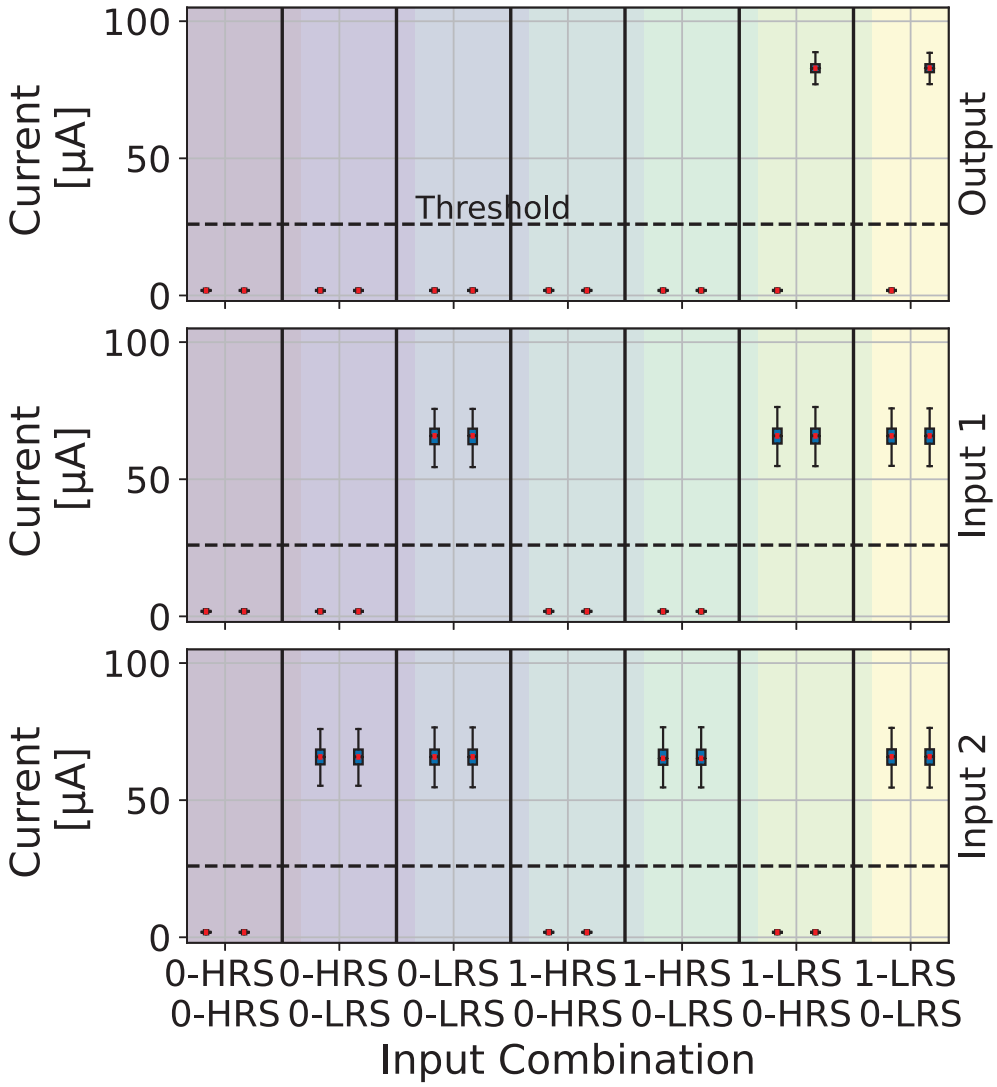


Figure 4.23: Boxplots of the simulated readout current before (left) and after (right) transistor-specific two-input one-output SET-based logic operations as depicted in Fig. 4.8(b). The simulations are performed with 2 000 samples according to Table 2.1, Table 2.2 and with normal distributed parameters with $\sigma = 5\,\%$ and a truncation of 2σ. During a logic operation the voltage $V_{\mathrm{op}} = 1.1\,\mathrm{V}$ is applied at both inputs. For an input of 1, the corresponding transistor gate is set to $V_{\mathrm{gate}} = 5\,\mathrm{V}$, otherwise to GND. At the transistor gate of the output device, 5 V are applied in any case. The currents are acquired by a simulated readout pulse of $0.25\,\mathrm{V}$, applied individually to all three devices before and after the logic operation. The threshold line marks the value of $26\,\mu\mathrm{A}$ which is based on the results in Table 3.4.

Table 4.8: Calculated error probabilities for the inputs and the output in the simulated transistor-specific two-device SET-based operation. The calculations are performed based on the voltage data from Fig. 4.23. During the operation, the voltage drop across the inputs occurs in RESET polarity, therefore only inputs in LRS can undergo an unintended state change. In case of a 0 as logic gate, an input switching is not possible due to the non-conductive transistor channel. The lowest computable probability was 1×10^{-18}.

Input combination	Error probability	
	Output	Input in LRS
0HRS - 0HRS	1×10^{-18}	-
0HRS - 0LRS	1×10^{-18}	-
0LRS - 0LRS	1×10^{-18}	-
1HRS - 0HRS	1.3×10^{-6}	-
1HRS - 0LRS	1.6×10^{-6}	-
1LRS - 0HRS	3.0×10^{-13}	9.3×10^{-5}
1LRS - 0LRS	2.2×10^{-16}	9.3×10^{-5}

4.3 Summary of Stateful Switching Logic

This chapter explored the implementation of memristive stateful logic and its feasibility. The concept relies on a voltage divider with resistive devices as logic inputs which determine the voltage drop across an additional device. This output starts in a pre-defined state and, depending on the input configuration and the resulting voltage, a conditional resistive switching process is induced or inhibited. One advantage of the switching-based implementation is that the occurrence of a switching event determines the logic output and therefore its data is stored within the operation itself. However the execution requires an initialization by reading the initial output state and potentially performing a RESET.

A theoretical analysis has shown that logic implementations based on the resistive RESET-process of filamentary VCM-ReRAM devices demonstrate reliability issues. During the operation, at least one logic input is destroyed. While this itself violates the state conservation and facilitates the need for additional recovery steps, the false behavior produces also a corrupted output state. On the other hand, the RESET process enables the implementation of a two-input NOR gate while the SET process facilitates the creation of an two-input OR gate. However in contrast to the functionally complete NOR, a single OR gate is not logically complete, i.e., it cannot execute all 16 boolean operations. Therefore the SET-based operation requires additional logic components to form a full set of computations. If the RESET-based NOR can be implemented reliably, it is more favorable and versatile for broader applications. In cases that such a RESET-based switching is desired, the voltage required for resetting the device should be generally lower than the voltage needed to initiate the SET process. Filamentary VCM devices typically require lower absolute voltages to initiate the SET process than a RESET. This makes only SET-based logic operations feasible in this device type.

In this chapter, the practical SET-based two-device logic OR gate was demonstrated. While in this operation the state conservation, i.e., a false input switching, is not problematic, ensuring the correct output switching is crucial. The operation shows a strong dependence on the applied voltage, in 1T-1R systems especially the gate voltage which implements a current limitation. A precise voltage tuning is the best method to enable successful operations in stateful logic. The voltage trade-off lies in finding the right balance between ensuring correct output switching and preventing false output switching. If the voltage is too high or too low, it may either cause false switching or fail to trigger the correct switching.

Three potential extensions of the two-input SET-based switching logic were investigated in simulations. At first, the NIMP logic which implements the two-input NIMP, NOT and XOR. Especially the NOT is important as the logic OR has to be combined with a NOT to achieve a logic complete NOR. The NIMP logic is executed by biasing both logic inputs with individual voltages rather than the same potential as in the OR. The investigation found that the optimal ratio of both input voltages is between 0.2 and 0.4. This balances the input conservation with a correct output switching. For a ratio of 0.3, the NIMP logic reliability is comparable to that of a standard two-device OR gate, making it a useful option for logic operations.

The second concept exploited a two-device one-resistor structure which aims to implement both a SET-based NOR and OR gate. While this method maintained a good input state conservation and a reliable NOR, the OR gate output remains unreliable. Therefore it is favorable to rely fully on the functional complete two-device one-resistor NOR or to combine it with a standard two-device OR. However, this presents a trade-off as it requires a resistor as additional analog circuit element, but it also offers the benefit of implementing a NOR directly.

The last extension was the transistor-specific logic which was previously investigated for the non-stateful readout-based operation in Section 3.2.5. The theoretical analysis showed that this extension is possible. However the experimental measurements indicated a critical influence of the bitcell transistor conductivity. For a practical implementation the performance needs to be optimized by adjusting the transistor's width-to-length (W/L) ratio or by applying different gate voltages to the inputs and the output. This effectively controls their current conduction. Unfortunately, it was not possible to measure this approach in the MAD-200 setup due to hardware limitations. All three extensions are different paths to improving the SET-based OR logic, with each approach balancing trade-offs between circuit design and reliability.

In addition to investigating the logic gate behavior, this chapter also proposed a framework dedicated to the evaluation of switching-based logic. This framework offers a tool for investigating switching-based logic primitives even in the absence of extensive experimental data. While the overall trends of the framework align well with a Monte-Carlo benchmark, the quantitative values observed in the simulations require further adjustments to improve the precision.

CHAPTER 5

Crossbar Operations for Logic & Arithmetic Computing-in-Memory Evaluation

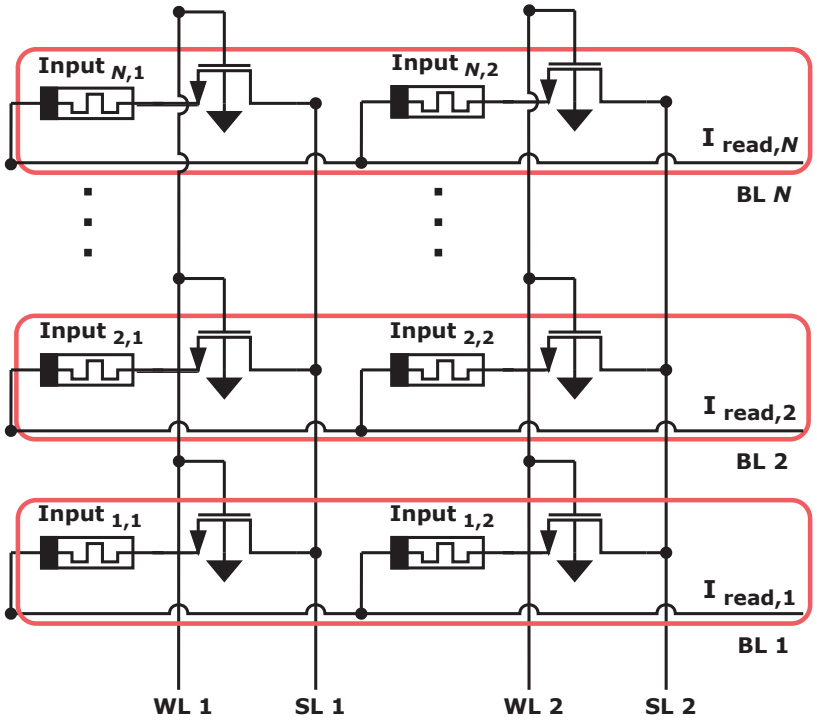
The previous two chapters discussed readout-based and switching-based single logic operations based on resistive 1T-1R devices. This chapter focuses on their evaluation on a crossbar scale and concatenated operations.

5.1 Parallel Multi-Gate Evaluation

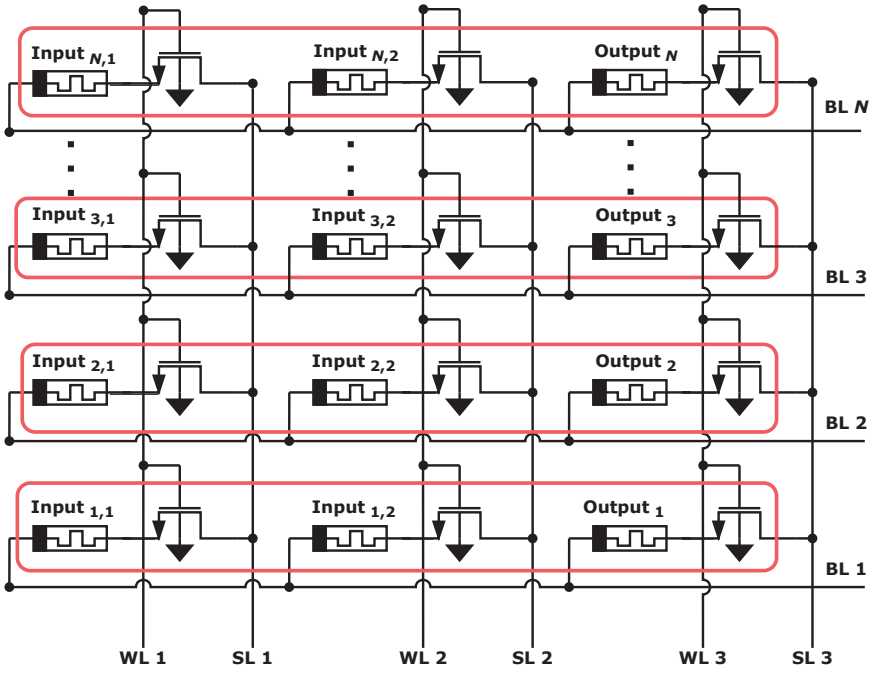
A possible way to increase the operation efficiency is by evaluating multiple logic gates during the same operation cycle. Yet, this parallel evaluation is only possible to execute in certain memristive crossbar structures, for instance the pseudo-crossbar array. An exemplary row-wise parallel two-input evaluation is depicted in Fig. 5.1 for the readout-based and the switching-based operations.

For the readout-based logic this is a common approach which shares similarities with the vector-matrix multiplication [226]. However for the switching-based evaluation a parallel evaluation is more critical. Although it was proposed as solution for optimizing the energy efficiency [11, 154], only 4 parallel gates XOR gates [244] and 8 parallel NAND gates [245] have been demonstrated experimentally in 1T-1R crossbars so far. However the parallel gate evaluation maximizes the efficiency of the crossbar architecture. Therefore these operations have to be studied intensively.

In extension to the previously demonstrated 4-bit parallel XOR and 8-bit parallel NAND operations in literature, the 12-bit parallel switching-based OR operation is presented in the following. This operation utilizes three Signal-Lines/Word-Lines (SL/WL), SL/WL [1, 2, 3], and 12 Bit-Lines (BLs) within the MAD-200 pseudo-crossbar array. The 12 individual logic OR gates are executed as described in Section 4.2.4 where two memristive devices serve as inputs, either as LRS or HRS, while the designated output device is initialized in the HRS. For the measurement, the devices in SL/WL 3 are selected as output devices and programmed into the high-resistive state. The input devices in SL/WL 1 (Input 1) and SL/WL 2 (Input 2) are programmed into the four possible input states, HRS-HRS, HRS-LRS, HRS-HRS and LRS-LRS, in a repetitive manner, each input combination three times. As initial step, the devices are individually programmed into their respective states and readout with $V_{\text{read}} = 0.25\,\text{V}$ and $V_{\text{gate,read}} = 5\,\text{V}$. The switching-based logic operation is then conducted with $V_{\text{op}} = 1.0\,\text{V}$ and $V_{\text{gate,op}} = 1.7\,\text{V}$. This voltage configuration provided the best results in the single gate evaluation in Fig. 4.11. After the operation, each device is readout again.



(a)



(b)

Figure 5.1: Crossbar arrangement for row-wise parallel two-input (a) readout-based and (b) switching-based logic gates. By applying the operation voltages column-wise, i.e., at the vertical Signal-Lines (SL) and Word-Lines (WL), the same logic operation is performed row-wise (highlighted as red boxes), i.e., in each horizontal Bit-Line (BL).

Fig. 5.2 depicts the readout currents before and after one exemplary operation for both, the input and output devices. The parallel operation proves a general feasibility as 11 out of 12 outputs are evaluated correctly. However, the input combination LRS-HRS in Bit-Line 7 fails. Although the output device shifts towards a lower resistance, as indicated by the lighter color, the change is insufficient for it to be classified as LRS afterwards. A similar output current increase is visible in Bit-Line 9. Yet this device should remain in the high-resistive state for the input combination HRS-HRS.

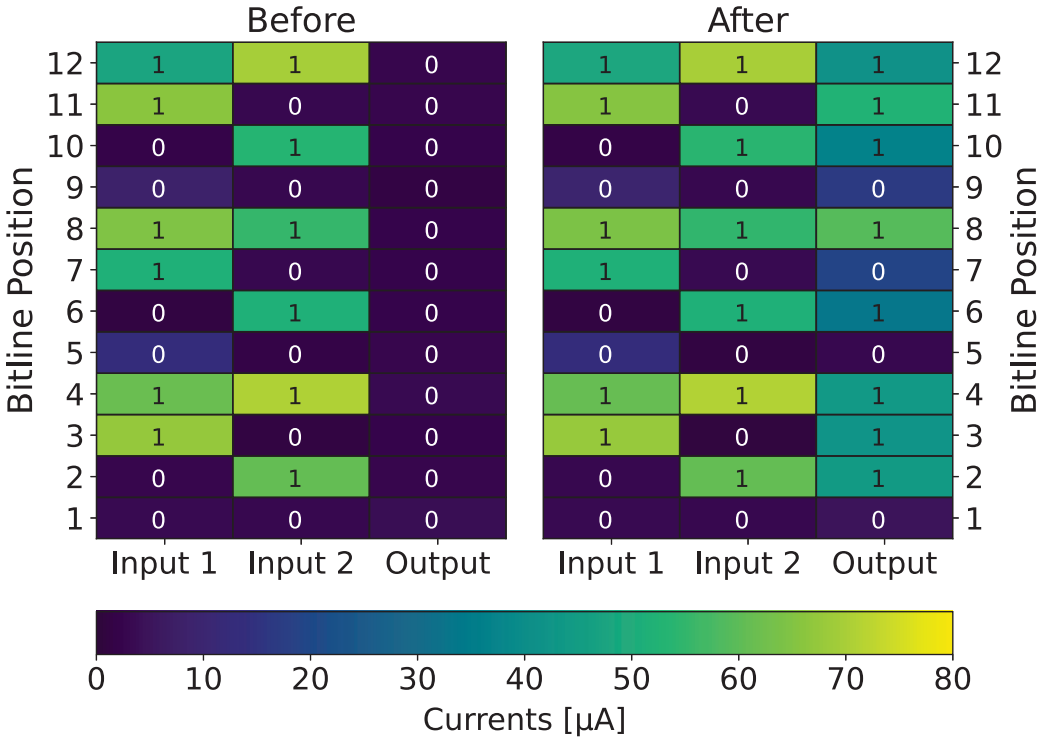


Figure 5.2: Change in the experimental readout currents for a 12-bit parallel switching-based logic operation. Displayed are the readout currents before (left) and after (right) performing the switching-based logic OR parallel in 12 Bit-Lines with V_{op} = 1.0 V and $V_{gate,op}$ = 1.7 V. The labeled states, 0 = HRS and 1 = LRS, are distinguished based on the value of 26 µA which is based on the results in Table 3.4.

To further investigate the switching variability, the same operation was repeated 100 times, with reprogramming all input devices after each operation. The corresponding readout currents for the outputs devices are depicted in Fig. 5.3 and for the input devices in Fig. 5.4. The outputs which are expected to switch to a low-resistance state, Bit-Lines 2–4, 6–8, and 10–12, demonstrate a significantly non-uniform behavior with a large spread in the output current distributions. Additionally, the Bit-Lines where no switching is expected, Bit-Lines 1, 5, and 9, also exhibit a noticeable current increase, especially the previously mentioned Bit-Line 9 which demonstrates the largest deviation. These results indicate a high output variability in the parallel stateful switching operation. However, the input devices remain stable due to being reprogrammed before every operation. While this guarantees accurate input states for the individual operations, it is not a typical use case in memristive crossbars. Reprogramming the inputs for every operation or even verifying their state by a readout would increase the operation costs in energy and latency. This reduces the efficiency of parallel operations.

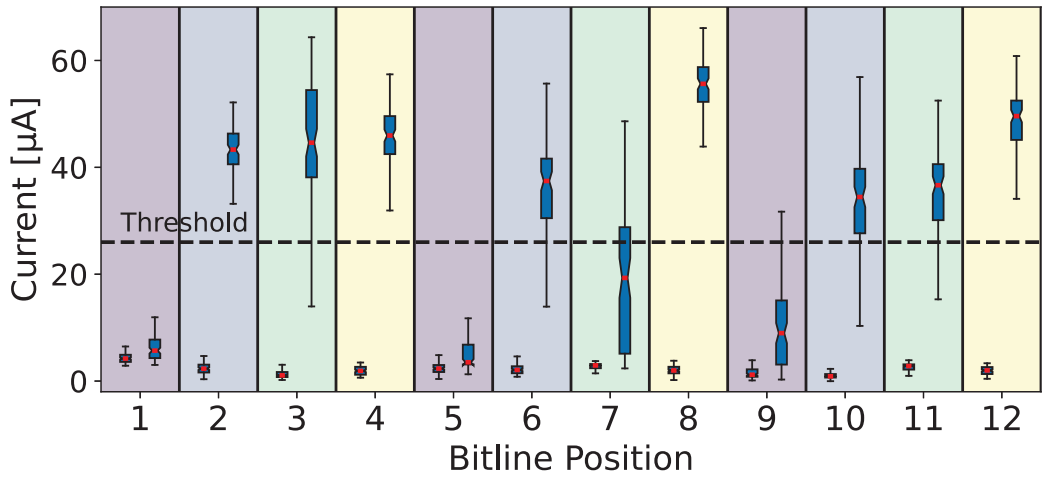


Figure 5.3: Boxplots of the experimental readout currents for the outputs of a 12-bit parallel two-input one-output SET-based logic circuit with 1T-1R bitcells as depicted in Fig. 5.1(b). The boxplots for each Bit-Line position represent the readout currents before (left) and after (right) 100 operations. The input combinations are programmed similar as depicted in Fig. 5.2 in the pattern HRS-HRS, HRS-LRS, LRS-HRS, and LRS-LRS, and the operation is performed with V_{op} = 1.0 V and $V_{gate,op}$ = 1.7 V. The currents are acquired by applying a readout voltage of 0.25 V individually to all devices before and after the logic operation. The threshold line marks the value of 26 µA which is based on the results in Table 3.4.

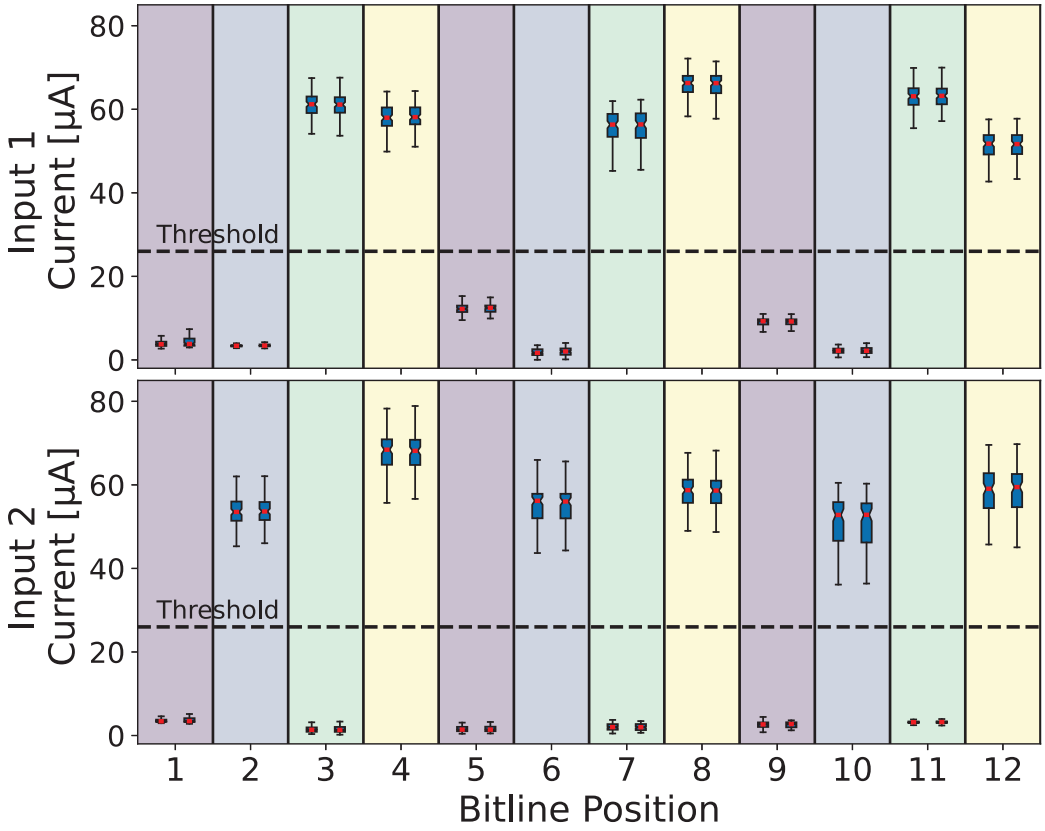


Figure 5.4: Boxplots of the experimental readout currents for the inputs of a 12-bit parallel two-input one-output SET-based operation as depicted in Fig. 5.1(b). The boxplots for each Bit-Line position represent the readout currents before (left) and after (right) 100 operations. The input combinations are programmed similar as depicted in Fig. 5.2 in the pattern HRS-HRS, HRS-LRS, LRS-HRS, and LRS-LRS, and the operation is performed with $V_{\mathrm{op}} = 1.0\,\mathrm{V}$ and $V_{\mathrm{gate,op}} = 1.7\,\mathrm{V}$. The currents are acquired by applying a readout voltage of $0.25\,\mathrm{V}$ individually to all devices before and after the logic operation. The threshold line marks the value of $26\,\mu\mathrm{A}$ which is based on the results in Table 3.4.

To better evaluate the practical conditions, an additional measurement is conducted where the input devices were programmed only once before the 100 operation cycles. The output devices, however, are re-initialized after each operation. The input readout current evolution throughout the operations is depicted in Fig. 5.5. The color representation indicates that repeated operations do not result in any noticeable shift in the input states. Due to the applied voltage polarity in the SET direction, only high-resistive input devices might perform an unintended switching. However, the operation voltages are significantly below the RESET thresholds so an unintended input switching is unlikely to occur.

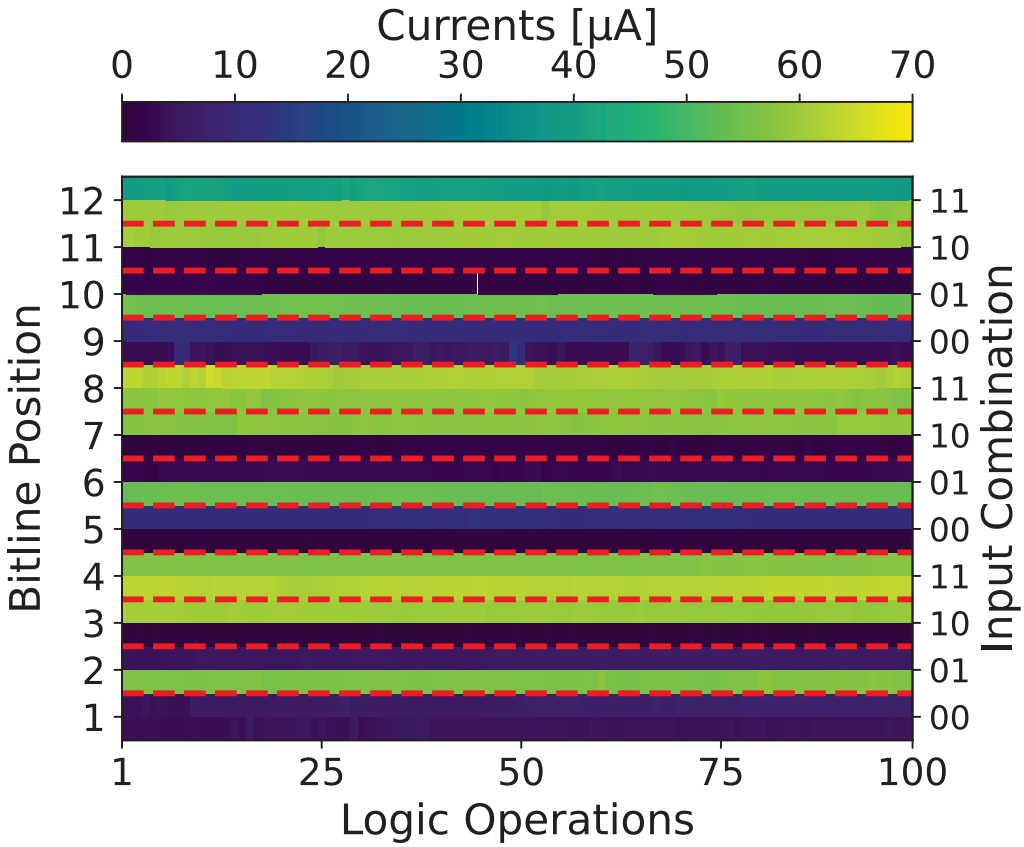


Figure 5.5: Temporal evolution of the experimental readout currents for the logic inputs of a 12-bit parallel two-input one-output SET-based operation without re-programming between the operations. The color-scale denotes the readout current for the two logic inputs of all 12 Bit-Lines, separated by the red horizontal lines, throughout 100 performed operations. The input combinations are programmed similar as depicted in Fig. 5.2 in the pattern HRS-HRS, HRS-LRS, LRS-HRS, and LRS-LRS, and the operation is performed with V_{op} = 1.0 V and $V_{gate,op}$ = 1.7 V. The currents are acquired by applying a readout voltage of 0.25 V individually to all devices after the logic operation.

While ensuring non-disturbed inputs during concatenated parallel operations is beneficial, improving the operation reliability and achieving error-free outputs is more important. Therefore also for this data set, the output currents before and after the operations are depicted in Fig. 5.6. Overall, the observed behavior is similar to that of the measurement with re-programmed inputs in Fig. 5.3. To quantify the operation reliability, the columns SL/WL [1, 2, 3] in Table 5.1 list the corresponding switching rates for the two experimental data sets. While the switching rates are relatively consistent in both data sets, a larger variability is observed within the same input combination across different Bit-Lines. For

example, HRS-LRS (01) is reliably implemented in BL 2 but performs poorly in BLs 6 and 10. This indicates a significant device-to-device variability in the operation, which is challenging to improve. The underlying voltage dilemma is particularly highlighted by the weak switching observed in BL 7 and the prominent unintended switching in BL 9. Increasing the operation voltage could improve the switching performance in cases like BL 7, but it would also support unintended switching as in BL 9. Achieving this trade-off between enhancing desired switching and minimizing errors is the most critical goal for parallel switching-based operations.

Table 5.1: Experimental Switching Rates (SR) in [%] for the 12-bit parallel SET-based switching OR for the measurement data set with input-reprogramming (Re-prog.) in Fig. 5.3 and the two data sets without re-programming (No Re-prog.) in Fig. 5.6 and Fig. 5.7. The rates denote the number of post-operation output devices (out of 100) to be in LRS state. A LRS state is defined as a current value above 26 µA based on the voltage-dependent readout investigation in Table 3.4.

Bit-Line	Inputs	Expected Output	Switching Rates [%]		
			SL/WL [1, 2, 3] Re-prog.	SL/WL [1, 2, 3] No Re-prog.	SL/WL [4, 5, 6] No Re-prog.
1	00	0	0	1	1
2	01	1	100	99	89
3	10	1	92	88	98
4	11	1	99	98	100
5	00	0	1	1	0
6	01	1	88	73	99
7	10	1	32	38	56
8	11	1	100	100	100
9	00	0	4	9	0
10	01	1	79	84	90
11	10	1	87	85	97
12	11	1	100	97	100

In order to further test the device-to-device variability, 100 operations are conducted in three other pairs (SL/WL), SL/WL [4, 5, 6] without re-programming the inputs between the operations. The input pattern is selected as before. Fig. 5.7 depicts the resulting output currents as boxplots and the switching rates for are listed in the last column of Table 5.1. The comparison to Fig. 5.6 reveals several interesting observations. First, the device-to-device variability is very significant, as seen for instance in the behaviors of BL 2 and BL 3. In the new data set, the output current variability of BL 2 is greater, while BL 2 shows a reduced variability. Additionally, the non-switching outputs, BL 5 and BL 9, now behave as expected, with no noticeable current increase. However, BL 1 shows a larger current increase compared to the previous data set. Similar in both data sets is the behavior of BL 7, where a significant portion of the outputs remains in the HRS. While the switching rate for BL 7 has improved to over 50 %, the error rate remains too high for a reliable operation. These findings suggest that either both output devices on BL 7 are damaged in a similar way, or there are underlying issues with this Bit-Line. For instance, the contacting from the Bit-Line to the respective pin might be insufficient or

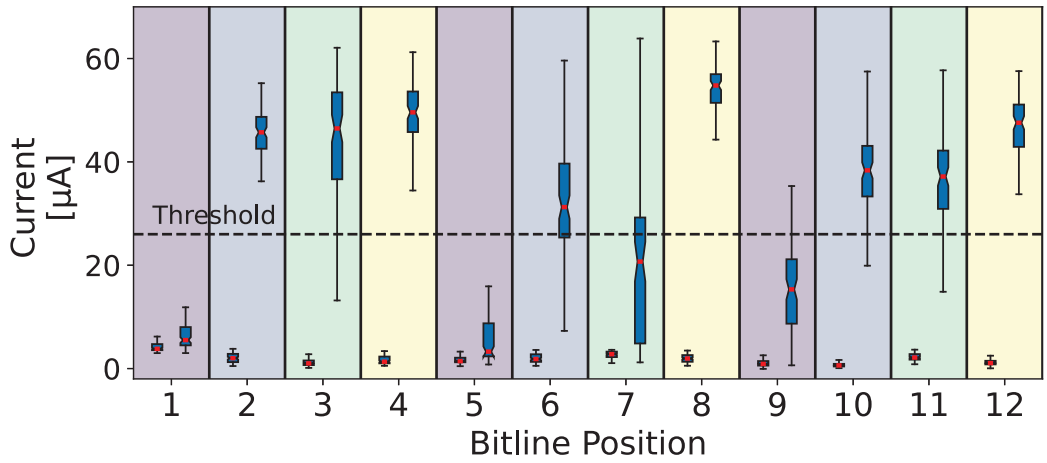


Figure 5.6: Boxplots of the experimental readout currents for the outputs of a 12-bit parallel two-input one-output SET-based logic without re-programming the inputs between the operations. The boxplots for each Bit-Line position represent the readout currents before (left) and after (right) 100 operations. The operations are performed with V_{op} = 1.0 V and $V_{gate,op}$ = 1.7 V, and the currents are acquired by applying a readout voltage of 0.25 V individually to all devices before and after the logic operation. The threshold line marks the value of 26 µA which is based on the results in Table 3.4.

the corresponding line on the PCB is malfunctioning. Overall, the measurements indicate that a strong uniformity in the 1T-1R devices as well as the surrounding circuitry is important in order to realize homogeneous parallel switching-based operations.

In addition to the performed measurements, the stateful logic evaluation was also simulated for 16 parallel gates. In this simulation parasitic elements such as line resistance and coupling capacitance are added for estimating a physical crossbar behavior. Fig. 5.8 depicts the calculated position-dependent error probability based on the proposed framework. It is seen that the simulations strongly overestimate the logic OR operation reliability compared to the above measured data. As in the simulation, all devices undergo the similar device-to-device variability distributions, no single high-error devices can occur. However, this device-to-device variability will dominate the overall reliability. Therefore, the simulations are only valid in cases of a very strong device uniformity. From all simulated gates, only the NIMP 10 combination shows a strong position dependence. This indicates that in this case the switching is very sensitive to voltage fluctuations. In contrast, the two-input parallel OR gates are similarly reliable in all 16 rows.

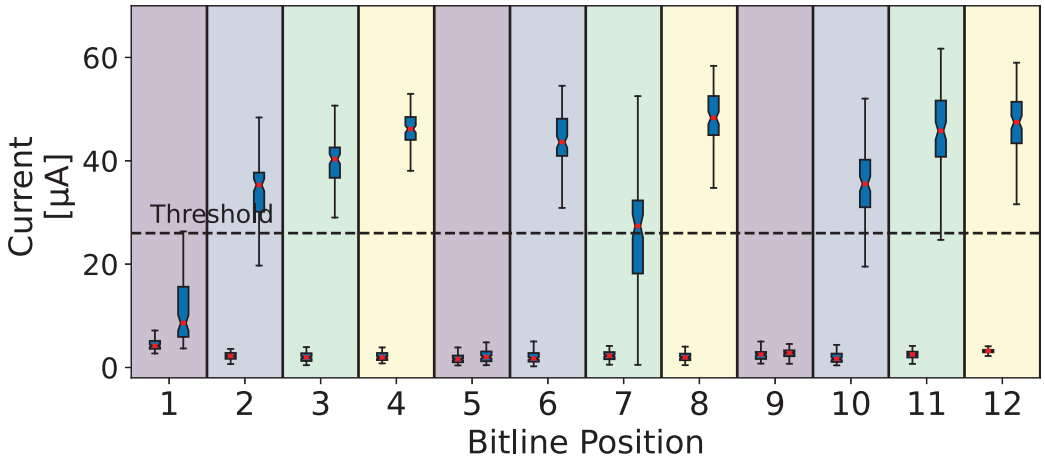


Figure 5.7: Boxplots of the experimental readout currents for the outputs of a 12-bit parallel two-input one-output SET-based logic without re-programming the inputs between the operations. Compared to the data in Fig. 5.6, different Signal-Lines/Word-Lines are activated and therefore different devices serve as inputs and outputs. The boxplots for each Bit-Line position represent the readout currents before (left) and after (right) 100 operations. The operations are performed with $V_{\mathrm{op}} = 1.0\,\mathrm{V}$ and $V_{\mathrm{gate,op}} = 1.7\,\mathrm{V}$, and the currents are acquired by applying a readout voltage of $0.25\,\mathrm{V}$ individually to all devices before and after the logic operation. The threshold line marks the value of $26\,\mu\mathrm{A}$ which is based on the results in Table 3.4.

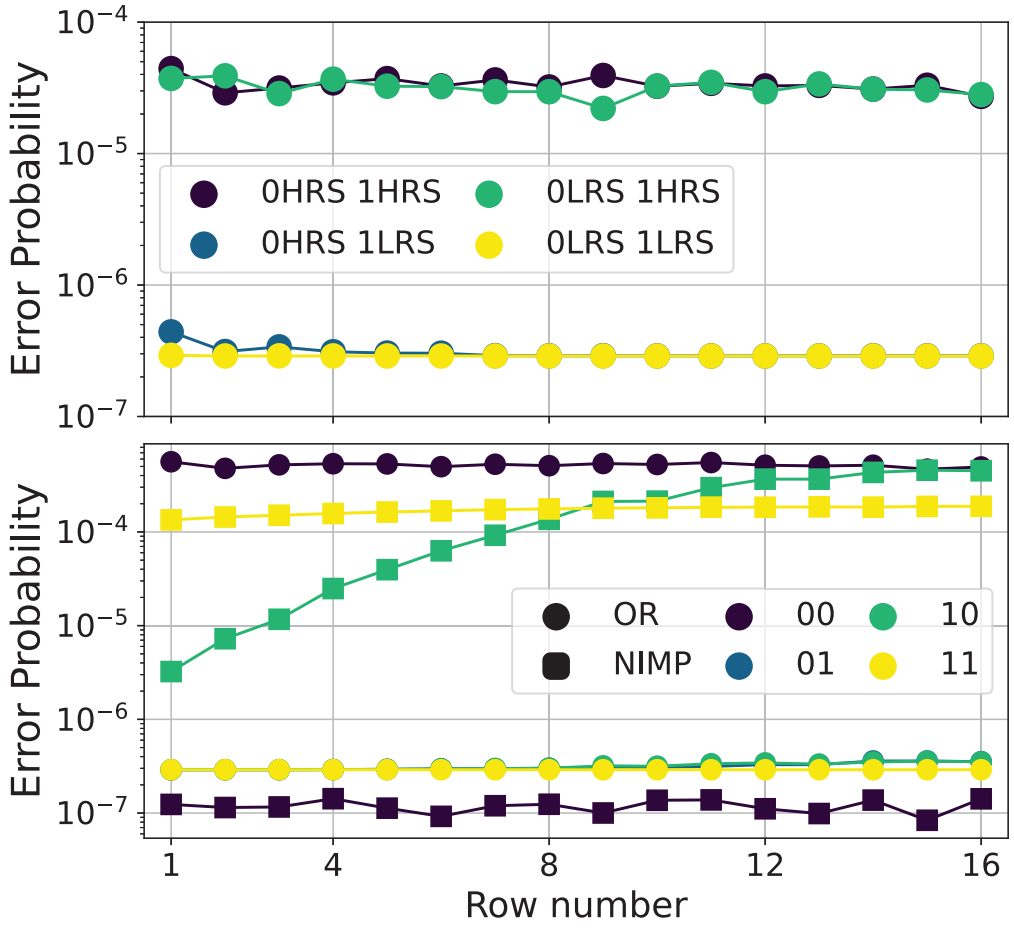


Figure 5.8: Calculated error probabilities for the simulated parallel two-input transistor-specific gates and the stateful logic gates OR and NIMP. The operation is simulated in a 16 x 3 pseudo-crossbar crossbar with parasitic elements per bitcell of $C_{\text{bitcell}} = 33\,\text{aF}$ and $R_{\text{bitcell}} = 9.7\,\Omega$. All other simulation parameters are selected according to Table 2.1 and Table 2.2 with $V_{\text{gate,op}} = 5\,\text{V}$ and $V_{\text{op}} = 1.3\,\text{V}$.

5.2 Arithmetic Operations on Crossbar Level

5.2.1 Error Propagation for Concatenated Operations

The error probability for a single readout-based or switching-based resistive logic operation can be calculated by using the proposed frameworks in Section 3.2.1 and Section 4.2.1. However especially for the switching-based operations, each output is a potential resistive input for the next operation. In case of non-ideal outputs as logic inputs, errors might propagate through the concatenated gates [246].

In order to calculate the error probability for a sequence of concatenated switching-based operations, an error propagation similar to a probability tree is used [197, 247]. This process is exemplary shown in Fig. 5.9 for a two-stage logic OR operation. Firstly, two inputs, $A = 0$ and $B = 0$, are programmed and verified. Afterwards the switching-based OR operation, $A + B = C$, is conducted. Even with the ideal inputs, the output device might undergo a false switching with a probability of P_{00}. A correct result has the opposing probability of $(1 - P_{00})$. The output C is now used in a second step as logic input with the, again ideal, input $D = 1$. Depending of the post-operation state of C, either $CD = 01$ or $CD = 11$ is calculated. While both should result in a logic 1, each operation has an individual error probability, P_{01} and P_{11}, to result in an output of $E = 0$. The total probability for the output E is calculated by first multiplying the probabilities along each path and then summing up those paths which lead to the same output. In this case $P(0) = [P_{01} \times (1 - P_{00})] + [P_{11} \times P_{00}]$ and $P(1) = [(1 - P_{01}) \times (1 - P_{00})] + [(1 - P_{11}) \times P_{00}]$. Although this example only features the switching-based OR, the methodology can be adapted for any kind of switching-based logic function or, if a write-back step is added, also for a readout-based logic. This will be demonstrated in the following section for the example of a Carry-Ripple-Adder.

One possible optimization strategy for the displayed OR gate is to utilize its output asymmetry. Under the assumption of evenly distributed inputs, the OR output should switch in 3 out of 4 possible input cases. Therefore, without knowing the inputs, there is a higher probability that the output should switch rather than remain unchanged. To assist this, the voltage can be biased slightly higher than the balancing point which was calculated in Fig. 4.12. This favors a switching, ensuring that the gate operates correctly in most scenarios. While this approach may increase the risk of a corrupted output in the 00 input case where no switching is needed, this trade-off is acceptable since the 00 case occurs only in 25 % of all cases.

5.2.2 Full-Adder Implementation

As seen before, the readout-based logic evaluation offers the advantage of a lower error probability. However it requires an additional write-back step to permanently store the operation result. In contrast, the switching-based logic inherently combines operation and storage in a single step. This advantage yet comes in hand with a reduced reliability. An interesting approach is therefore to combine both readout-based and switching-based within a single application. This hybrid strategy exploits the strengths of both methods. As example, a combined stateful/non-stateful Carry-Ripple Adder (CRA) is demonstrated in the following, yet the idea can be extended to other logic operations and systems. This application was previously published as part of [206] with portions of the text adapted from that work.

When adding two binary *N*-bit numbers, A and B, a full adder calculates both a SUM and CARRY for each corresponding bit in A and B. While the SUM bit is permanently stored as part of the result, the CARRY bit is only propagated to the following more-significant bit evaluation. Solely the final CARRY bit, representing an overflow, needs to be stored as the most significant bit in the *N+1* bit

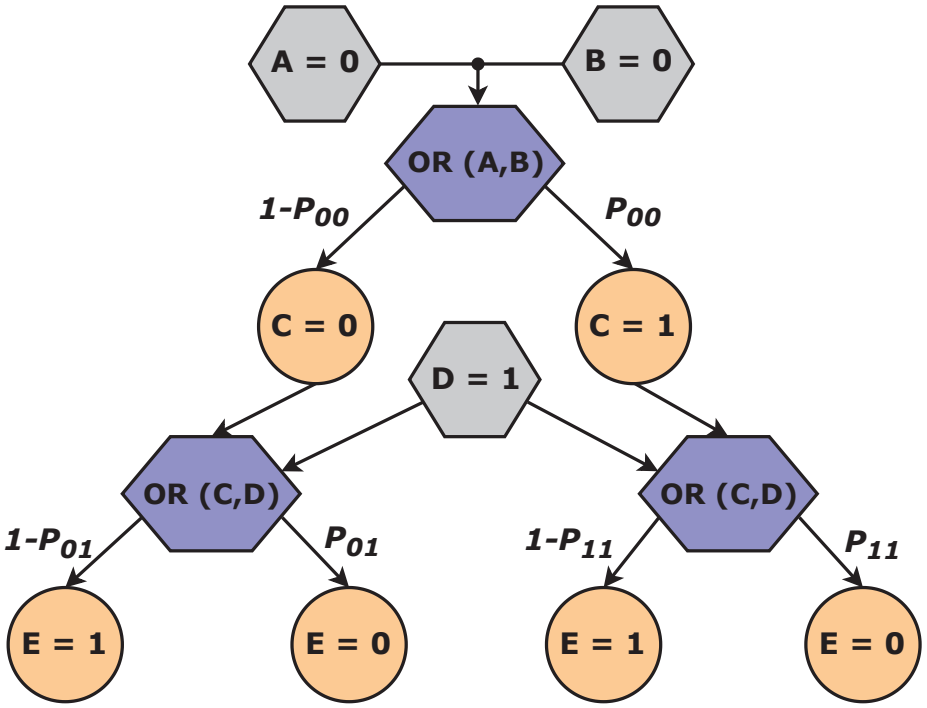


Figure 5.9: Error propagation for concatenated logic operations in a probability tree. A and B are ideal input states while C is the non-ideal first-stage output which is concatenated again with an ideal input D. This results in the non-ideal second-stage output E. The probability for each state of E can be calculated by multiplying along the tree paths and then summing up all paths which result in the same output.

position. There are numerous concepts for implementing binary adders each with varying trade-offs in terms of speed and complexity. The proposed Carry-Ripple Adder here is based on the following representations for the SUM (S) and CARRY (C_{out}) bits

$$C_{\text{out}} = \overline{C_{\text{in}}} * (A * B) + C_{\text{in}} * (A + B) \tag{5.1}$$

$$S = C_{\text{in}} * (\overline{A \oplus B}) + \overline{C_{\text{in}}} * (A \oplus B) \tag{5.2}$$

with C_{in} being the CARRY-In bit and $*, +, \oplus$ the Boolean functions AND, OR and XOR, respectively. This calculation can be splitted in a stateful and a non-stateful part. The stateful evaluation is used for computing the necessary intermediate values which are used in multiple operations. Also the addition SUM (S) is evaluated statefully since the sum is the important addition result. The CARRY (C_{out}) is calculated in a non-stateful manner since most CARRY bits are only required in one following step. This mixed concept combines the inherent advantages of stateful (permanent storage) with non-stateful (more reliable).

The structure of the equations Eq. (5.1) and Eq. (5.2) is important for this splitting. Both equations share a similar form as they consist of the sum of two terms. In each term, either C_{in} or its negated value, $\overline{C_{\text{in}}}$, is multiplied with another factor. This representation is chosen specifically for the implementation in 1T-1R crossbar with the transistor-specific logic as discussed in Section 3.2.5 and Section 4.2.6. In the transistor-specific logic, a volatile gate voltage and a non-volatile ReRAM state are multiplied ($*$) in a 1T-1R bitcell. Two individual of these products can then be added ($+$) by evaluating them in a

parallel manner. Compared to the equations Eq. (5.1) and Eq. (5.2), the common factors C_{in} and $\overline{C_{\text{in}}}$ are chosen as volatile gate voltages and the terms $(A * B)/(A + B)/(A \oplus B)/(\overline{A \oplus B})$ as non-volatile ReRAM states. Consequentially, the operations according to the above representations split into two parts: 1) Obtaining the input-dependent factors $(A * B)/(A + B)/(A \oplus B)/(\overline{A \oplus B})$ (**Pre-Steps**) via stateful switching-based parallel operation and 2) Concatenating them with the CARRY-Inputs, $C_{\text{in}}/\overline{C_{\text{in}}}$ (**Calculation**) via stateful and non-stateful transistor-specific logic operations. The resulting operation flow is depicted in Fig. 5.10.

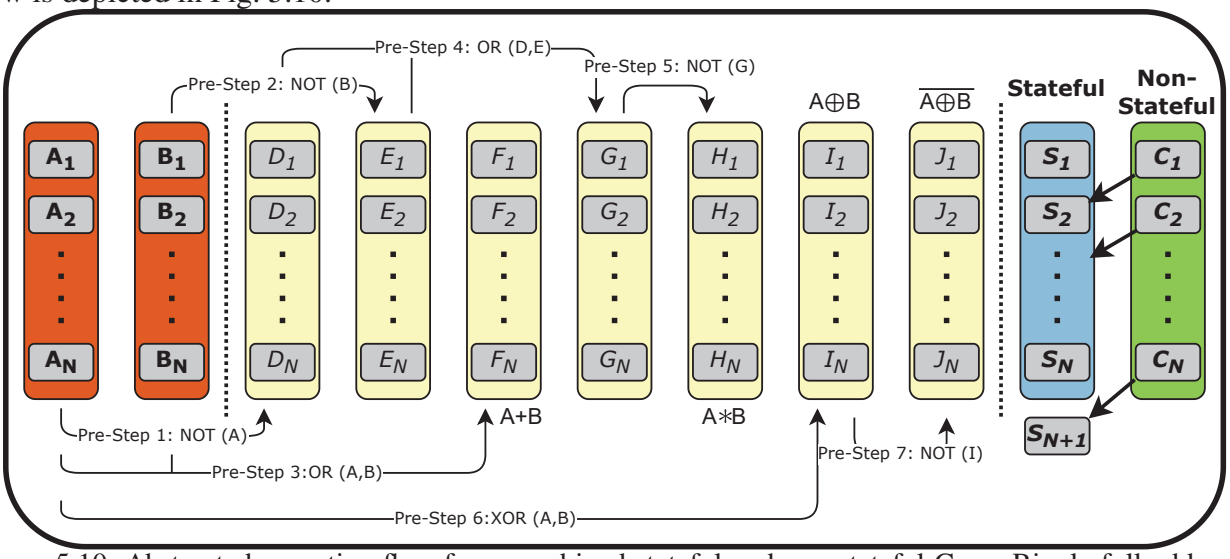


Figure 5.10: Abstracted operation flow for a combined stateful and non-stateful Carry-Ripple full adder. A *stateful* operation refers to the investigated switching-based logic operations from Chapter 4, while *non-stateful* refers to the readout-based logic operations from Chapter 3. A and B represent the N-bit input vectors and D to J the results of the stateful Pre-Step operations. These are used together with the non-stateful CARRY vector, C, to calculate the $N + 1$-bit SUM vector, S by the transistor-specific switching-based operation as proposed in Section 4.2.6. Taken from [206], published under the Creative Commons Attribution License 4.0 [182].

The CRA is tested in a probabilistic simulation. For each addition operation, two N-bit inputs are randomly generated and stored as ideal input values. During each step of the following computation, Pre-Steps and Calculation, the bit-wise results are determined with included error propagation. This propagation is modeled as described in Section 5.2.1 and is based on the simulated simulated error probabilities of parallel evaluated logic gates. Note here that the parallel evaluations are simulated in a different framework than described in Table 2.2. While these simulations exploit the same ReRAM model, JART VCM v1b var, a $22\,\text{nm}$ transistor technology PDK instead of the $180\,\text{nm}$ PDK is used. The achieved $N + 1$ bit output value is then compared to the theoretical expected value. M=1×10^7 additions were simulated each for an input length of $N = [2, 4, 8, 16]$. From the results, the absolute error, f_{abs}, and relative error, f_{rel} per addition are calculated as

$$f_{\text{abs}} = \frac{\sum_i (x_i - y_i)^2}{M} \tag{5.3}$$

$$f_{\text{rel}} = \frac{\sum_i (x_i - y_i)^2}{(M * (2^{N+1} - 2))} \tag{5.4}$$

with x_i being the theoretically expected addition result and y_i the simulated result. The relative error is calculated since an absolute deviation of ± 1 on a 2-bit addition is worse than ± 1 for a 16-bit addition.

The results are depicted Table 5.2. It is seen that while the absolute error increases with the input length, the relative error remains constant.

Table 5.2: Absolute and relative error for the simulated Carry-Ripple Full Adder for different input lengths N. The error calculation is conducted according to Eq. (5.3) and Eq. (5.4).

Number of input bits N	Absolute error f_{abs} per addition	Relative error f_{rel} per addition
2	2.97×10^{-5}	4.94×10^{-6}
4	1.19×10^{-4}	3.96×10^{-6}
8	2.01×10^{-3}	3.94×10^{-6}
16	6.16×10^{-1}	4.70×10^{-6}

The presented example of a carry-ripple full-adder is simplified, yet it illustrates the approach of combining stateful and non-stateful logic operations. Both flavors are logically complete and therefore any arithmetic computation can theoretically be performed through their concatenation. However, this might be not optimal for practical implementations. Replacing all stateful operations with non-stateful gates might lead to a better reliability as the error probability of the non-stateful two-input operations surpasses the same operation in a switching-based implementation. However, in order to perform concatenated crossbar operations, this would require additional write-back steps for every individual result. On the other hand, a fully stateful operation scheme requires no such write-back, yet the stateful single gates demonstrate only a reduce reliability. Therefore for any given application, it is advantageous to determine which results must be stored permanently and then execute them as stateful operations while substituting intermediate steps with faster and more error-tolerant non-stateful operations. This represents an advancement toward reliable and CMOS-competitive Computing-in-Memory architectures.

CHAPTER 6

Conclusion

6.1 Summary

This thesis explores various logic operations based on valence-change-mechanism (VCM)-type filamentary resistive switching devices. From the wide range of gates proposed in the literature, only those were selected for this thesis which fulfill the two criteria: a) being compatible with memristive 1T-1R crossbar architectures and b) preserving the logic state of their inputs. The investigated concepts primarily differ in their representation of the logic outputs, either as volatile current in non-stateful readout-based logic or as a non-volatile resistive state in stateful switching-based logic operation. A special emphasis has been the statistical evaluation of logic gates to identify the key factors influencing their reliability. For this aim, variability-aware circuit simulations are combined with measurements of integrated ReRAM devices.

The fundamental principle of readout-based logic relies on detecting differences in the resistive states of ReRAM devices by applying a small sensing voltage and measuring the resulting device current. The most important factors in the performance of readout-based logic are the resistance ratio between the different resistive states and the magnitude of the sensing voltage. While the available states are limited by the selected materials, the sensing voltage is a free parameter to tune during the operation. Higher sensing voltages can improve the accurate state differentiation by scaling the absolute current difference. However, the voltage has to be adjusted carefully as excessive sensing voltages may induce permanent changes in the device state or a cumulative degradation over time. For the measured HfO_2-based VCM devices, the maximum sensing voltage was identified between $0.25\,\mathrm{V}$ and $0.3\,\mathrm{V}$. Despite selecting the optimal voltage, state classification errors can still occur in cases of overlapping resistive state distributions. In this thesis, multi-input logic gates, such as two-input symmetric gates (e.g., OR, AND, XOR) and three-input gates, i.e., MAJORITY were investigated. While all operations are feasible, they differed in their respective error probability. As the number of inputs increased, the resistive state variability becomes more pronounced. This was especially critical for devices in the low-resistance state. The resulting overlapping state distributions significantly increase the likelihood of logic classification errors.

The switching-based logic is also based on the differences in the resistive state of ReRAM devices, but in this case they are exploited as part of a state-dependent resistive voltage divider. Although often proposed, the investigated switching-based logic gate has not been demonstrated with co-integrated 1T-1R devices so far. Therefore an important achievement of this thesis is to demonstrate the operation on a larger scale. It is observed that the applied voltage plays a critical role to ensure a reliable switching logic. In particular, inducing a correct switching operation is more vulnerable than preventing an unintended switching. This displays the inherent optimization problem in switching-based logic, which tries to improve the likelihood of correct switching. However all measures in this regard can increase the occurrence of unintended switching events. One promising observation was that all operations satisfied the input state conservation. Furthermore, the parallel evaluation of logic gates within a 1T-1R crossbar array was studied. Here, several switching-based logic operations are evaluated as entire column. The main result is the demonstration of 12 parallel stateful two-input OR gates. While the operation is in general possible, the conducted variability study revealed a significant issue regarding its reliability. Especially a non-uniform device behavior will degrade any advantage in the performance for parallel computed logic.

The implementation of transistor-specific logic was tested for both, readout-based and switching-based approaches. In this context, the transistor within the 1T-1R bitcell is utilized not solely as a selector but also as a logic element. Since both approaches depend on differences in resistive states, the transistor's gate voltage is an effective way to modulate the transistor channel resistance and therefore the overall 1T-1R bitcell behavior. However, the investigation in this thesis shows no significant reliability improvement from this approach. Modulating the transistor's channel resistance might bring advantages for specific applications, particularly if the sub-threshold transistor leakage current can be distinguished reliably. In this case, the transistor-specific operation offers an additional logic level. Yet if the leakage current and the HRS distributions overlap, transistor-specific performs only as good as the standard resistive logic gates. Especially, it does not serve as a universal tool for enhancing the limited performance of resistive switching-based logic operations.

Lastly, the combination of both concepts for larger arithmetic calculations is discussed on the example of a carry-ripple full adder. Although this is a simple example, it highlights the advantages of combining both principles to achieve an optimal performance.

The main evaluation criteria in this thesis are the feasibility, i.e., which logic operations can be implemented at all, and the reliability, i.e., how accurate are the operations. It was shown that switching-based two-input logic operation will not be as reliable as comparable readout-based operations. In conclusion, if the operation correctness is the sole evaluation criterion, all resistive logic gates should ideally be implemented as readout-based operations. However, the trade-off is more complicated. A minimal error rate is achieved by implementing logic gates in standard CMOS [248] rather than using resistive devices. Given the 40-year head start of CMOS compared to ReRAM, this is not surprising. Nevertheless, the comparison here is not between a single CMOS gate and a resistive logic gate. Instead, the reference is a resistive 1T-1R crossbar for pure memory purpose. Such an structure consists not solely of a 1T-1R crossbar array but requires at least two peripheral circuits: 1) to selectively program individual 1T-1R bitcells (**write data**), and 2) to apply a sensing voltage and readout the state of a device (**fetch data**). If the ReRAM devices are operated in a binary fashion, the data sensing requires a 1-bit ADC, e.g., a sense amplifier, for converting the single device readout current to a digital signal of '0' or '1'. Therefore combining the readout operation with the execution of a simple logic gate is a low-cost extension.

The switching-based logic gates are more complex to implement as they require larger and dedicated peripheral circuits. Consequently, these operations are not designated for standard applications such as arithmetic-logic units (ALU), graphic processing, or other computationally intensive operations commonly performed by traditional CMOS-based systems. Instead they are better suited for niche scenarios where the unique characteristics of memristive Computing-in-Memory provide specific advantages, for instance standalone systems which perform computations only occasionally but unavoidably.

The advantage of resistive devices over CMOS lies in their ability to store data in a non-volatile manner, eliminating the need for energy-consuming data refreshing. This makes them ideal for energy-constrained applications, where the trade-offs in complexity and reliability are outweighed by the benefits of non-volatility and energy efficiency.

6.2 Outlook

The findings of this thesis can lead to different future research directions.

First, the proposed error probability framework offers a versatile tool that can be adapted to evaluate other types of stateful or non-stateful logic gates. This is particularly interesting for switching-based logic, where determining the reliability is challenging due to limited measurement data. By providing a tool to estimate error probabilities, the framework serves as a preliminary evaluation method to determine whether a given concept is worth pursuing experimentally. Yet, especially the switching probability offers room for improvement as it was only determined in dependence of the applied voltage. Here, other parameters such as the initial resistive state, the ambient temperature or transient effects could be included too. Furthermore in the future, the framework has to be more precisely fitted to experimental data.

Additionally for the switching-based operations, particularly the NIMP, XOR and TMSL, this thesis lacks measurement data due to the complexity of integrating these operations. Also the demonstrated OR gates require a more intensive study of their dependencies on all operation parameters to better understand and optimize their performance. While sufficient data is available for the readout-based operations, the evaluation in this thesis was done manually. Future works need to implement the analog periphery, such as analog-to-digital converters, to fully investigate their interplay with the resistive devices and determine the reliability of the full system.

Although the measurements and simulations conducted in this thesis focus specifically on VCM-filamentary type resistive devices, the underlying methodologies and conclusions for logic operations can be extended to other types of resistive memories. For the readout-based operation, the main criteria are a sufficient resistance ratio to distinguish the resistive states, a narrow state variability and a non-volatile readout. Therefore it would be interesting to see if these operations can also be implemented in phase-change memories, magnetic STT-RAM or other types of oxide-based ReRAM such as the area-dependent VCM devices. Besides a low resistive state variability, the switching-based operation also depends on the variability of the switching voltages and the ratio of $|V_{\mathrm{Reset}}|/|V_{\mathrm{Set}}|$. If the switching voltage variability is too high, even for perfectly adjusted resistive states the output devices might not switch properly or they switch unintended. Furthermore, if the switching voltages, $|V_{\mathrm{Reset}}|$ and $|V_{\mathrm{Set}}|$, are in a similar range, the input cells might perform a RESET/SET process while the designated output cell should perform a SET/RESET vice versa. As this input switching issue is only present in bipolar switching resistive device, a future extension would be to implement stateful switching-based logic gates in unipolar switching memristive technologies, for instance phase-change memories.

Overall, integrated resistive devices represent a promising field with an advanced technology readiness level. This is emphasized by the growing industrial interest in their development. In the field of neuromorphic computing, the advantages of analog-switching resistive devices over conventional CMOS-based digital elements are undeniable as they offer unique opportunities for energy-efficient and biologically inspired processing. However for the logic operations demonstrated in this thesis, a pure CMOS implementation remains the most reliable and efficient solution. This is not surprising, given the decades of intensive research and optimization in CMOS technology. Nevertheless, the areas which are currently dominated by CMOS may require a re-evaluation as emerging memory and computing technologies, including resistive devices, continue to improve. These improvements could lead to potential opportunities for either replacing or extending CMOS in hybrid computing systems.

APPENDIX A

Appendix

A.1 Derivation of Intersection Points for Normal Distributions

Subsequently, the intersection points between two generalized normal distributions are derived. Following Eq. (2.2), the functions are represented as

$$f_{(a,m_1,s_1)}(x) = a \times e^{-(x-m_1)^2/2s_1} \tag{A.1}$$

and

$$f_{(b,m_2,s_2)}(x) = b \times e^{-(x-m_2)^2/2s_2} \tag{A.2}$$

with real coefficients $\{a, b, m_1, s_1, m_2, s_2\} \in \mathbb{R}$ under the restrictions $\{a, b\} \neq 0$ and $s_1 \neq s_2$.
The intersection points $\{x_1, \ldots x_n\}$ are defined as $f_{(a,m_1,s_1)}(x_i) = f_{(b,m_2,s_2)}(x_i); i \in [1, n]$. Therefore the intersection points follow by analytically solving the equations $f_{(a,m_1,s_1)}(x) = f_{(b,m_2,s_2)}(x)$ for x.

$$f_{(a,m_1,s_1)}(x) = f_{(b,m_2,s_2)}(x) \tag{A.3}$$

$$\Leftrightarrow a \times \exp\{-(x-m_1)^2/2s_1\} = b \times \exp\{-(x-m_2)^2/2s_2\}$$

$$\Leftrightarrow \frac{a}{b} = \frac{\exp\{-(x-m_2)^2/2s_2\}}{\exp\{-(x-m_1)^2/2s_1\}} = \exp\{-(x-m_2)^2/2s_2 + (x-m_1)^2/2s_1\}$$

$$\Leftrightarrow \ln\left(\frac{a}{b}\right) = \frac{-2s_1(x-m_2)^2 + 2s_2(x-m_1)^2}{4s_1s_2} \tag{A.4}$$

$$\Leftrightarrow 4s_1s_2 \cdot \ln\left(\frac{a}{b}\right) = -2s_1x^2 + 4s_1m_2x - 2s_1m_2^2 + 2s_2x^2 - 4s_2m_1x + 2s_2m_1^2$$

$$\Leftrightarrow 0 = 2(s_2 - s_1)x^2 + 4(s_1m_2 - s_2m_1)x + 2(s_2m_1^2 - s_1m_2^2) - 4s_1s_2 \ln\left(\frac{a}{b}\right) \tag{A.5}$$

Eq. (A.5) resembles a quadratic equation in the form $px^2 + qx + r = 0$ with real coefficients p, q, r and $p \neq 0$. The two roots of this equation can be easily determined with the quadratic formula:

$$x_{1,2} = \frac{-q \pm \sqrt{q^2 - 4pr}}{2p} \tag{A.6}$$

In the case of Eq. (A.5), the coefficients p, q, r are

$$\begin{aligned} p &= 2(s_2 - s_1) \\ q &= 4(s_1 m_2 - s_2 m_1) \\ r &= 2(s_2 m_1^2 - s_1 m_2^2) - 4 s_1 s_2 \ln\left(\frac{a}{b}\right) \end{aligned}$$

Substituting these coefficients in Eq. (A.6) yiels in

$$\begin{aligned} x_{1,2} = & \frac{-4(s_1 m_2 - s_2 m_1)}{2 \cdot 2(s_2 - s_1)} \\ & \pm \frac{\left([4(s_1 m_2 - s_2 m_1)]^2 - 4 \cdot 2(s_2 - s_1) \cdot \left[2(s_2 m_1^2 - s_1 m_2^2) - 4 s_1 s_2 \ln\left(\frac{a}{b}\right) \right] \right)^{1/2}}{2 \cdot 2(s_2 - s_1)} \\ \Leftrightarrow & \frac{(s_1 m_2 - s_2 m_1)}{(s_1 - s_2)} \pm \frac{1}{4(s_2 - s_1)} \times \Big(16 \big[s_1^2 m_2^2 - 2 s_1 s_2 m_1 m_2 + s_2^2 m_1^2 - s_2^2 m_1^2 \\ & + s_1 s_2 m_1^2 + s_1 s_2 m_2^2 - m_2^2 s_1^2 + 2(s_2 - s_1) \cdot s_1 s_2 \ln\left(\frac{a}{b}\right) \big] \Big)^{1/2} \\ \Leftrightarrow & \frac{(s_1 m_2 - s_2 m_1)}{(s_1 - s_2)} \mp \frac{1}{(s_1 - s_2)} \times \Big(-2 s_1 s_2 m_1 m_2 + s_1 s_2 m_1^2 + s_1 s_2 m_2^2 + \\ & 2(s_2 - s_1) \cdot s_1 s_2 \ln\left(\frac{a}{b}\right) \Big)^{1/2} \end{aligned}$$

Finally this simplifies to the expression

$$x_{1,2} = \frac{1}{(s_1 - s_2)} \left[\mp \Big(s_1 s_2 \big(2(s_1 - s_2) \ln\left(\frac{a}{b}\right) + (m_1 - m_2)^2 \big) \Big)^{1/2} + m_2 s_1 - m_1 s_2 \right] \tag{A.7}$$

Bibliography

[1] *The Nobel Prize in Physics 1956. NobelPrize.org. Nobel Prize Outreach AB 2024*, Accessed: 25/11/2024, URL: https://www.nobelprize.org/prizes/physics/1956/summary/ (cit. on p. 1).

[2] S. Yuvaraja, H. Faber, M. Kumar, N. Xiao, G. I. Maciel García, X. Tang, T. D. Anthopoulos and X. Li, *Three-dimensional integrated metal-oxide transistors*, Nature Electronics **7** (2024) 768, DOI: 10.1038/s41928-024-01205-0 (cit. on p. 1).

[3] Weebit Nano Ltd., 2023, URL: https://www.weebit-nano.com/news/press-releases/weebit-nano-successfully-secures-us40-million-to-accelerate-development-and-commercial-roll-out-of-its-reram/ (visited on 05/12/2023) (cit. on pp. 1, 7).

[4] M. Giordano et al., "CHIMERA: A 0.92 TOPS, 2.2 TOPS/W Edge AI Accelerator with 2 MByte On-Chip Foundry Resistive RAM for Efficient Training and Inference", *2021 Symposium on VLSI Circuits*, 2021 1, DOI: 10.23919/VLSICircuits52068.2021.9492347 (cit. on pp. 1, 7, 15, 18).

[5] Infineon Technologies AG, 2022, URL: https://www.infineon.com/cms/de/about-infineon/press/market-news/2022/INFATV202211-031.html (visited on 05/12/2023) (cit. on pp. 1, 7).

[6] E. Lehtonen and M. Laiho, "Stateful implication logic with memristors", *2009 IEEE/ACM International Symposium on Nanoscale Architectures*, 2009 33, DOI: 10.1109/NANOARCH.2009.5226356 (cit. on pp. 2, 19, 20, 92, 109).

[7] R. Rosezin, E. Linn, C. Kugeler, R. Bruchhaus and R. Waser, *Crossbar Logic Using Bipolar and Complementary Resistive Switches*, IEEE Electron Device Letters **32** (2011) 710, DOI: 10.1109/LED.2011.2127439 (cit. on pp. 2, 19).

[8] S. Kvatinsky, D. Belousov, S. Liman, G. Satat, N. Wald, E. G. Friedman, A. Kolodny and U. C. Weiser, *MAGIC—Memristor-Aided Logic*, IEEE Transactions on Circuits and Systems II: Express Briefs **61** (2014) 895, DOI: 10.1109/TCSII.2014.2357292 (cit. on pp. 2, 19, 92, 98).

[9] R. Ben Hur, N. Wald, N. Talati and S. Kvatinsky, "Simple magic: Synthesis and in-memory Mapping of logic execution for memristor-aided logic", *2017 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2017 225, DOI: 10.1109/ICCAD.2017.8203782 (cit. on p. 2).

[10] H. A. Du Nguyen, L. Xie, M. Taouil, R. Nane, S. Hamdioui and K. Bertels,
On the Implementation of Computation-in-Memory Parallel Adder,
IEEE Transactions on Very Large Scale Integration (VLSI) Systems **25** (2017) 2206,
DOI: 10.1109/TVLSI.2017.2690571 (cit. on pp. 2, 19).

[11] A. Haj-Ali, R. Ben-Hur, N. Wald, R. Ronen and S. Kvatinsky,
IMAGING: In-Memory AlGorithms for Image processiNG,
IEEE Transactions on Circuits and Systems I: Regular Papers **65** (2018) 4258,
DOI: 10.1109/TCSI.2018.2846699 (cit. on pp. 2, 127).

[12] S. Singh, C. K. Jha, A. Bende, P. L. Thangkhiew, V. Rana, S. Patkar, R. Drechsler and F. Merchant,
Should We Even Optimize for Execution Energy? Rethinking Mapping for MAGIC Design Style,
IEEE Embedded Systems Letters **15** (2023) 230, DOI: 10.1109/LES.2023.3298740
(cit. on p. 2).

[13] C. K. Jha, K. Qayyum, K. Çağlar Coşkun, S. Singh, M. Hassan, R. Leupers, F. Merchant and R. Drechsler, *veriSIMPLER: An Automated Formal Verification Methodology for SIMPLER MAGIC Design Style Based In-Memory Computing*,
IEEE Transactions on Circuits and Systems I: Regular Papers **71** (2024) 4169,
DOI: 10.1109/TCSI.2024.3424682 (cit. on p. 2).

[14] M. Fritscher et al.,
A flexible and fast digital twin for RRAM systems applied for training resilient neural networks,
Scientific Reports **14** (2024) 23695, DOI: 10.1038/s41598-024-73439-z (cit. on p. 2).

[15] R. Landauer, *Information is a physical entity*,
Physica A: Statistical Mechanics and its Applications **263** (1999), Proceedings of the 20th IUPAP International Conference on Statistical Physics 63,
DOI: https://doi.org/10.1016/S0378-4371(98)00513-5 (cit. on p. 5).

[16] IRDS, *Beyond CMOS and Emerging Materials Integration*,
International Roadmap for Devices and Systems - 2022 Edition (2022),
URL: https://irds.ieee.org/images/files/pdf/2022/2022IRDS_BC.pdf
(cit. on pp. 6, 7, 9).

[17] J. S. Meena, S. M. Sze, U. Chand and T.-Y. Tseng,
Overview of emerging nonvolatile memory technologies,
Nanoscale Research Letters **9** (2014) 526, DOI: 10.1186/1556-276X-9-526 (cit. on p. 6).

[18] R. Dittmann, S. Menzel and R. Waser,
Nanoionic memristive phenomena in metal oxides: the valence change mechanism,
Advances in Physics **70** (2021) 155, DOI: 10.1080/00018732.2022.2084006
(cit. on pp. 5–7, 10).

[19] R. Waser and M. Aono, *Nanoionics-based resistive switching memories*,
Nature Materials **6** (2007) 833, DOI: 10.1038/nmat2023 (cit. on p. 5).

[20] I. Valov, R. Waser, J. R. Jameson and M. N. Kozicki,
Electrochemical metallization memories—fundamentals, applications, prospects,
Nanotechnology **22** (2011) 254003, DOI: 10.1088/0957-4484/22/25/254003 (cit. on p. 5).

[21] H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. T. Chen and M.-J. Tsai, *Metal–Oxide RRAM*, Proceedings of the IEEE **100** (2012) 1951, DOI: 10.1109/JPROC.2012.2190369 (cit. on p. 6).

[22] G. Bersuker and D. Gilmer, "Metal oxide resistive random access memory (RRAM) technology", *Advances in Non-volatile Memory and Storage Technology*, ed. by Y. Nishi, Woodhead Publishing, 2014 288, ISBN: 978-0-85709-803-0, DOI: https://doi.org/10.1533/9780857098092.2.288 (cit. on pp. 6, 7).

[23] R. Waser, R. Dittmann, G. Staikov and K. Szot, *Redox-Based Resistive Switching Memories – Nanoionic Mechanisms, Prospects, and Challenges*, Advanced Materials **21** (2009) 2632, DOI: https://doi.org/10.1002/adma.200900375 (cit. on pp. 6, 7).

[24] J. J. Yang, F. Miao, M. D. Pickett, D. A. A. Ohlberg, D. R. Stewart, C. N. Lau and R. S. Williams, *The mechanism of electroforming of metal oxide memristive switches*, Nanotechnology **20** (2009) 215201, DOI: 10.1088/0957-4484/20/21/215201 (cit. on pp. 7, 16).

[25] R. Waser, R. Bruchhaus and S. Menzel, "Redox-based Resistive Switching Memories", *Nanoelectronics and Information Technology*, ed. by R. Waser, 3rd ed., Wiley-VCH, 2012 683 (cit. on pp. 7, 8).

[26] H. Wu, X. H. Wang, B. Gao, N. Deng, Z. Lu, B. Haukness, G. Bronner and H. Qian, *Resistive Random Access Memory for Future Information Processing System*, Proceedings of the IEEE **105** (2017) 1770, DOI: 10.1109/JPROC.2017.2684830 (cit. on pp. 7, 9).

[27] M. Rao et al., *Thousands of conductance levels in memristors integrated on CMOS*, Nature **615** (2023) 823, DOI: 10.1038/s41586-023-05759-5 (cit. on pp. 7, 10, 12).

[28] F. Cüppers, S. Menzel, C. Bengel, A. Hardtdegen, M. von Witzleben, U. Böttger, R. Waser and S. Hoffmann-Eifert, *Exploiting the switching dynamics of HfO_2-based ReRAM devices for reliable analog memristive behavior*, APL Materials **7** (2019) 091105, DOI: 10.1063/1.5108654 (cit. on pp. 7, 10, 11, 25, 31).

[29] T. Ninomiya, S. Muraoka, Z. Wei, R. Yasuhara, K. Katayama and T. Takagi, *Improvement of Data Retention During Long-Term Use by Suppressing Conductive Filament Expansion in TaO_x Bipolar-ReRAM*, IEEE Electron Device Letters **34** (2013) 762, DOI: 10.1109/LED.2013.2258653 (cit. on pp. 7, 13).

[30] S. Wiefels, U. Böttger, S. Menzel, D. J. Wouters and R. Waser, "Statistical Modeling and Understanding of HRS Retention in 2.5 Mb HfO_2 based ReRAM", *2020 IEEE International Memory Workshop (IMW)*, 2020 1, DOI: 10.1109/IMW48823.2020.9108123 (cit. on pp. 7, 13, 21).

[31] G. Molas et al., "High temperature stability embedded ReRAM for 2x nm node and beyond", *2022 IEEE International Memory Workshop (IMW)*, 2022 1, DOI: 10.1109/IMW52921.2022.9779293 (cit. on pp. 7, 13).

[32] L. Grenouillet et al., "16kbit 1T1R OxRAM arrays embedded in 28nm FDSOI technology demonstrating low BER, high endurance, and compatibility with core logic transistors", *2021 IEEE International Memory Workshop (IMW)*, 2021 1, DOI: 10.1109/IMW51353.2021.9439607 (cit. on pp. 7, 13, 15, 16).

[33] N. Kopperberg, S. Wiefels, K. Hofmann, J. Otterstedt, D. J. Wouters, R. Waser and S. Menzel, *Endurance of 2 Mbit Based BEOL Integrated ReRAM*, IEEE Access **10** (2022) 122696, DOI: 10.1109/ACCESS.2022.3223657 (cit. on pp. 7, 13).

[34] S. Menzel, M. Waters, A. Marchewka, U. Böttger, R. Dittmann and R. Waser, *Origin of the Ultra-nonlinear Switching Kinetics in Oxide-Based Resistive Switches*, Advanced Functional Materials **21** (2011) 4487, DOI: https://doi.org/10.1002/adfm.201101117 (cit. on pp. 9, 10, 31).

[35] L. Chua, *Memristor-The missing circuit element*, IEEE Transactions on Circuit Theory **18** (1971) 507, DOI: 10.1109/TCT.1971.1083337 (cit. on p. 9).

[36] D. B. Strukov, G. S. Snider, D. R. Stewart and R. S. Williams, *The missing memristor found*, Nature **453** (2008) 80, DOI: 10.1038/nature06932 (cit. on p. 9).

[37] A. Ascoli, R. Tetzlaff and S. Menzel, *Exploring the Dynamics of Real-World Memristors on the Basis of Circuit Theoretic Model Predictions*, IEEE Circuits and Systems Magazine **18** (2018) 48, DOI: 10.1109/MCAS.2018.2821760 (cit. on p. 9).

[38] S. Vongehr and X. Meng, *The Missing Memristor has Not been Found*, Scientific Reports **5** (2015) 11657, DOI: 10.1038/srep11657 (cit. on p. 9).

[39] J. H. Yoon, Y.-W. Song, W. Ham, J.-M. Park and J.-Y. Kwon, *A review on device requirements of resistive random access memory (RRAM)-based neuromorphic computing*, APL Materials **11** (2023) 090701, DOI: 10.1063/5.0149393 (cit. on p. 9).

[40] H. Hassan, M. Anis and M. Elmasry, *MOS current mode circuits: analysis, design, and variability*, IEEE Transactions on Very Large Scale Integration (VLSI) Systems **13** (2005) 885, DOI: 10.1109/TVLSI.2005.853609 (cit. on p. 9).

[41] A. Rahimi, L. Benini and R. K. Gupta, *Variability Mitigation in Nanometer CMOS Integrated Systems: A Survey of Techniques From Circuits to Software*, Proceedings of the IEEE **104** (2016) 1410, DOI: 10.1109/JPROC.2016.2518864 (cit. on p. 9).

[42] S. Yu, X. Guan and H.-S. P. Wong, "On the stochastic nature of resistive switching in metal oxide RRAM: Physical modeling, monte carlo simulation, and experimental characterization", *2011 International Electron Devices Meeting*, 2011 17.3.1, DOI: 10.1109/IEDM.2011.6131572 (cit. on p. 9).

[43] R. Degraeve, A. Fantini, N. Raghavan, L. Goux, S. Clima, B. Govoreanu, A. Belmonte, D. Linten and M. Jurczak, *Causes and consequences of the stochastic aspect of filamentary RRAM*, Microelectronic Engineering **147** (2015), Insulating Films on Semiconductors 2015 171, DOI: https://doi.org/10.1016/j.mee.2015.04.025 (cit. on pp. 9, 10).

[44] E. Abbaspour, S. Menzel and C. Jungemann,
Studying the Switching Variability in Redox-Based Resistive Switching Devices,
J. Comput. Electron. **19** (2020) 1426, DOI: 10.1007/s10825-020-01537-y
(cit. on pp. 9, 10).

[45] J. L. Rieck, F. V. E. Hensling and R. Dittmann,
Trade-off between variability and retention of memristive epitaxial $SrTiO_3$ devices,
APL Materials **9** (2021) 021110, DOI: 10.1063/5.0035707 (cit. on p. 9).

[46] B. Lin, B. Gao, Y. Pang, P. Yao, D. Wu, H. He, J. Tang, H. Qian and H. Wu, "A High-Speed and High-Reliability TRNG Based on Analog RRAM for IoT Security Application",
2019 IEEE International Electron Devices Meeting (IEDM), 2019 14.8.1,
DOI: 10.1109/IEDM19573.2019.8993486 (cit. on p. 10).

[47] F. Moro et al., "Hardware calibrated learning to compensate heterogeneity in analog RRAM-based Spiking Neural Networks",
2022 IEEE International Symposium on Circuits and Systems (ISCAS), 2022 380,
DOI: 10.1109/ISCAS48785.2022.9937820 (cit. on p. 10).

[48] M. Zhao, B. Gao, Y. Xi, F. Xu, H. Wu and H. Qian,
Endurance and Retention Degradation of Intermediate Levels in Filamentary Analog RRAM,
IEEE Journal of the Electron Devices Society **7** (2019) 1239,
DOI: 10.1109/JEDS.2019.2943017 (cit. on pp. 10, 13).

[49] H. Wu, M. Zhao, Y. Liu, P. Yao, Y. Xi, X. Li, W. Wu, Q. Zhang, J. Tang, B. Gao and H. Qian, "Reliability Perspective on Neuromorphic Computing Based on Analog RRAM",
2019 IEEE International Reliability Physics Symposium (IRPS), 2019 1,
DOI: 10.1109/IRPS.2019.8720609 (cit. on p. 10).

[50] M. Lanza et al., *Recommended Methods to Study Resistive Switching Devices*,
Advanced Electronic Materials **5** (2019) 1800143,
DOI: https://doi.org/10.1002/aelm.201800143 (cit. on p. 10).

[51] T. Hennen, A. Elias, J. Nodin, G. Molas, R. Waser, D. Wouters and D. Bedau,
A high throughput generative vector autoregression model for stochastic synapses,
Frontiers in Neuroscience **16** (2022) 941753, DOI: 10.3389/fnins.2022.941753
(cit. on pp. 10, 11).

[52] S. Ambrogio, S. Balatti, A. Cubeta, A. Calderoni, N. Ramaswamy and D. Ielmini,
Statistical Fluctuations in HfO_x Resistive-Switching Memory: Part I - Set/Reset Variability,
IEEE Transactions on Electron Devices **61** (2014) 2912, DOI: 10.1109/TED.2014.2330200
(cit. on pp. 10, 11, 21, 98).

[53] A. Grossi et al., *Impact of Intercell and Intracell Variability on Forming and Switching Parameters in RRAM Arrays*, IEEE Transactions on Electron Devices **62** (2015) 2502,
DOI: 10.1109/TED.2015.2442412 (cit. on pp. 10, 16).

[54] G. Piccolboni, G. Molas, D. Garbin, E. Vianello, O. Cueto, C. Cagli, B. Traore, B. De Salvo, G. Ghibaudo and L. Perniola,
Investigation of Cycle-to-Cycle Variability in HfO_2-Based OxRAM,
IEEE Electron Device Letters **37** (2016) 721, DOI: 10.1109/LED.2016.2553370
(cit. on pp. 10, 21).

[55] A. Grossi, C. Zambelli, P. Olivo, E. Nowak, G. Molas, J. F. Nodin and L. Perniola, *Cell-to-Cell Fundamental Variability Limits Investigation in OxRRAM Arrays*, IEEE Electron Device Letters **39** (2018) 27, DOI: 10.1109/LED.2017.2774604 (cit. on pp. 10, 11, 49, 94).

[56] A. Hardtdegen, C. La Torre, F. Cüppers, S. Menzel, R. Waser and S. Hoffmann-Eifert, *Improved Switching Stability and the Effect of an Internal Series Resistor in HfO_2/TiO_x Bilayer ReRAM Cells*, IEEE Transactions on Electron Devices **65** (2018) 3229, DOI: 10.1109/TED.2018.2849872 (cit. on p. 10).

[57] D. Maldonado, S. Aldana, M. González, F. Jiménez-Molinos, M. Ibáñez, D. Barrera, F. Campabadal and J. Roldán, *Variability estimation in resistive switching devices, a numerical and kinetic Monte Carlo perspective*, Microelectronic Engineering **257** (2022) 111736, DOI: https://doi.org/10.1016/j.mee.2022.111736 (cit. on p. 10).

[58] C. Funck, S. Wiefels, C. Bengel, K. Schnieders, S. Hoffmann-Eifert, R. Dittmann and S. Menzel, *A Physical Description of the Variability in Single-ReRAM Devices and Hardware-Based Neuronal Networks*, Advanced Intelligent Systems **5** (2023) 2300129, DOI: https://doi.org/10.1002/aisy.202300129 (cit. on p. 10).

[59] J. P. Strachan, A. C. Torrezan, G. Medeiros-Ribeiro and R. S. Williams, *Measuring the switching dynamics and energy efficiency of tantalum oxide memristors*, Nanotechnology **22** (2011) 505402, DOI: 10.1088/0957-4484/22/50/505402 (cit. on p. 10).

[60] K. Fleck, U. Böttger, R. Waser and S. Menzel, *Interrelation of Sweep and Pulse Analysis of the SET Process in $SrTiO_3$ Resistive Switching Memories*, IEEE Electron Device Letters **35** (2014) 924, DOI: 10.1109/LED.2014.2340016 (cit. on p. 10).

[61] L. Goux, Y.-Y. Chen, L. Pantisano, X.-P. Wang, G. Groeseneken, M. Jurczak and D. J. Wouters, *On the Gradual Unipolar and Bipolar Resistive Switching of $TiN/HfO_2/Pt$ Memory Systems*, Electrochemical and Solid-State Letters **13** (2010) G54, DOI: 10.1149/1.3373529 (cit. on pp. 10, 11).

[62] J. P. Strachan, A. C. Torrezan, F. Miao, M. D. Pickett, J. J. Yang, W. Yi, G. Medeiros-Ribeiro and R. S. Williams, *State Dynamics and Modeling of Tantalum Oxide Memristors*, IEEE Transactions on Electron Devices **60** (2013) 2194, DOI: 10.1109/TED.2013.2264476 (cit. on pp. 10, 11).

[63] S. Menzel, U. Böttger, M. Wimmer and M. Salinga, *Physics of the Switching Kinetics in Resistive Memories*, Advanced Functional Materials **25** (2015) 6306, DOI: https://doi.org/10.1002/adfm.201500825 (cit. on p. 10).

[64] S. Ambrogio, S. Balatti, V. Milo, R. Carboni, Z.-Q. Wang, A. Calderoni, N. Ramaswamy and D. Ielmini, *Neuromorphic Learning and Recognition With One-Transistor-One-Resistor Synapses and Bistable Metal Oxide RRAM*, IEEE Transactions on Electron Devices **63** (2016) 1508, DOI: 10.1109/TED.2016.2526647 (cit. on p. 10).

[65] Y. Nishi, S. Menzel, K. Fleck, U. Böttger and R. Waser,
Origin of the SET Kinetics of the Resistive Switching in Tantalum Oxide Thin Films,
IEEE Electron Device Letters **35** (2014) 259, DOI: 10.1109/LED.2013.2294868
(cit. on pp. 10, 94).

[66] K. Fleck, C. La Torre, N. Aslam, S. Hoffmann-Eifert, U. Böttger and S. Menzel,
Uniting Gradual and Abrupt set Processes in Resistive Switching Oxides,
Phys. Rev. Appl. **6** (6 2016) 064015, DOI: 10.1103/PhysRevApplied.6.064015
(cit. on p. 10).

[67] L. Zhao, H.-Y. Chen, S.-C. Wu, Z. Jiang, S. Yu, T.-H. Hou, H.-S. Wong and Y. Nishi,
Multi-level control of conductive nano-filament evolution in HfO_2 ReRAM by pulse-train operations, Nanoscale **6** (2014), DOI: 10.1039/c4nr00500g (cit. on pp. 11, 49).

[68] G. Medeiros-Ribeiro, F. Perner, R. Carter, H. Abdalla, M. D. Pickett and R. S. Williams,
Lognormal switching times for titanium dioxide bipolar memristors: origin and resolution,
Nanotechnology **22** (2011) 095702, DOI: 10.1088/0957-4484/22/9/095702 (cit. on p. 11).

[69] M. Hu, Y. Wang, Q. Qiu, Y. Chen and H. Li,
"The stochastic modeling of TiO_2 memristor and its usage in neuromorphic system design",
2014 19th Asia and South Pacific Design Automation Conference (ASP-DAC), 2014 831,
DOI: 10.1109/ASPDAC.2014.6742993 (cit. on p. 11).

[70] C. Wenger, F. Zahari, M. K. Mahadevaiah, E. Pérez, I. Beckers, H. Kohlstedt and M. Ziegler,
Inherent Stochastic Learning in CMOS-Integrated HfO_2 Arrays for Neuromorphic Computing,
IEEE Electron Device Letters **40** (2019) 639, DOI: 10.1109/LED.2019.2900867
(cit. on p. 11).

[71] F. Zahari, E. Perez, M. K. Mahadevaiah, H. Kohlstedt, C. Wenger and M. Ziegler, *Analogue pattern recognition with stochastic switching binary CMOS-integrated memristive devices*,
Scientific Reports **10** (2020) 14450, DOI: 10.1038/s41598-020-71334-x (cit. on p. 11).

[72] C. La Torre, K. Fleck, S. Starschich, E. Linn, R. Waser and S. Menzel,
Dependence of the SET switching variability on the initial state in HfO_x-based ReRAM,
physica status solidi (a) **213** (2016) 316,
DOI: https://doi.org/10.1002/pssa.201532375 (cit. on p. 11).

[73] Y. Nishi, K. Fleck, U. Böttger, R. Waser and S. Menzel,
Effect of RESET Voltage on Distribution of SET Switching Time of Bipolar Resistive Switching in a Tantalum Oxide Thin Film, IEEE Transactions on Electron Devices **62** (2015) 1561,
DOI: 10.1109/TED.2015.2411748 (cit. on p. 11).

[74] G. Piccolboni et al., "Investigation of the potentialities of Vertical Resistive RAM (VRRAM) for neuromorphic applications", *2015 IEEE International Electron Devices Meeting (IEDM)*,
2015 17.2.1, DOI: 10.1109/IEDM.2015.7409717 (cit. on p. 11).

[75] J. McPherson, J.-Y. Kim, A. Shanware and H. Mogul,
Thermochemical description of dielectric breakdown in high dielectric constant materials,
Applied Physics Letters **82** (2003) 2121, DOI: 10.1063/1.1565180 (cit. on p. 11).

[76] S. Sun C. and Lu, F. Jin, W. Mo, J. Song and K. Dong, *The Resistive Switching Characteristics of $TiN/HfO_2/Ag$ RRAM Devices with Bidirectional Current Compliance*, Journal of Electronic Materials **48** (2019) 2992, DOI: 10.1007/s11664-019-07069-x (cit. on p. 11).

[77] S. Ambrogio, S. Balatti, A. Cubeta, A. Calderoni, N. Ramaswamy and D. Ielmini, *Statistical Fluctuations in HfOx Resistive-Switching Memory: Part II—Random Telegraph Noise*, IEEE Transactions on Electron Devices **61** (2014) 2920, DOI: 10.1109/TED.2014.2330202 (cit. on pp. 12, 49).

[78] F. M. Puglisi, N. Zagni, L. Larcher and P. Pavan, *Random Telegraph Noise in Resistive Random Access Memories: Compact Modeling and Advanced Circuit Design*, IEEE Transactions on Electron Devices **65** (2018) 2964, DOI: 10.1109/TED.2018.2833208 (cit. on p. 12).

[79] N. Raghavan, R. Degraeve, L. Goux, A. Fantini, D. Wouters, G. Groeseneken and M. Jurczak, "RTN insight to filamentary instability and disturb immunity in ultra-low power switching HfO_x and AlO_x RRAM", *2013 Symposium on VLSI Technology*, 2013 T164 (cit. on p. 12).

[80] P. Huang, D. B. Zhu, C. Liu, Z. Zhou, Z. Dong, H. Jiang, W. S. Shen, L. F. Liu, X. Y. Liu and J. F. Kang, "RTN based oxygen vacancy probing method for Ox-RRAM reliability characterization and its application in tail bits", *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017 21.4.1, DOI: 10.1109/IEDM.2017.8268435 (cit. on p. 12).

[81] S. Wiefels, C. Bengel, N. Kopperberg, K. Zhang, R. Waser and S. Menzel, *HRS Instability in Oxide-Based Bipolar Resistive Switching Cells*, IEEE Transactions on Electron Devices **67** (2020) 4208, DOI: 10.1109/TED.2020.3018096 (cit. on pp. 12, 25, 49, 63).

[82] T. Figliolia, P. Julian, G. Tognetti and A. G. Andreou, "A true Random Number Generator using RTN noise and a sigma delta converter", *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2016 17, DOI: 10.1109/ISCAS.2016.7527159 (cit. on p. 12).

[83] P.-H. Tseng, M.-H. Lee, Y.-H. Lin, H.-L. Lung, K.-C. Wang and C.-Y. Lu, *ReRAM-Based Pseudo-True Random Number Generator With High Throughput and Unpredictability Characteristics*, IEEE Transactions on Electron Devices **68** (2021) 1593, DOI: 10.1109/TED.2021.3057028 (cit. on p. 12).

[84] S. Rafiq, J. Hazra, M. Liehr, K. Beckmann, M. Abedin, J. S. Pannu, S. K. Jha and N. C. Cady, *Investigation of ReRAM Variability on Flow-Based Edge Detection Computing Using HfO_2-Based ReRAM Arrays*, IEEE Transactions on Circuits and Systems I: Regular Papers **68** (2021) 2900, DOI: 10.1109/TCSI.2021.3072210 (cit. on pp. 12, 21).

[85] J. Solanki, K. Beckmann, J. Pelton, N. Cady and M. Liehr, "Effect of Resistance variability in Vector Matrix Multiplication operations of 1T1R ReRAM crossbar arrays using an Embedded test platform", *2023 IEEE 32nd Microelectronics Design & Test Symposium (MDTS)*, 2023 1, DOI: 10.1109/MDTS58049.2023.10168152 (cit. on p. 12).

[86] E. Pérez, A. Grossi, C. Zambelli, P. Olivo, R. Roelofs and C. Wenger, *Reduction of the Cell-to-Cell Variability in $Hf_{1-x}Al_xO_y$ Based RRAM Arrays by Using Program Algorithms*, IEEE Electron Device Letters **38** (2017) 175, DOI: 10.1109/LED.2016.2646758 (cit. on pp. 12, 32).

[87] S. Balatti, S. Ambrogio, A. Cubeta, A. Calderoni, N. Ramaswamy and D. Ielmini, "Voltage-dependent random telegraph noise (RTN) in HfO_x resistive RAM", *2014 IEEE International Reliability Physics Symposium*, 2014 MY.4.1, DOI: 10.1109/IRPS.2014.6861159 (cit. on p. 12).

[88] F. M. Puglisi, L. Larcher, A. Padovani and P. Pavan, *A Complete Statistical Investigation of RTN in HfO_2-Based RRAM in High Resistive State*, IEEE Transactions on Electron Devices **62** (2015) 2606, DOI: 10.1109/TED.2015.2439812 (cit. on pp. 12, 49).

[89] K. Schnieders, C. Funck, F. Cüppers, S. Aussen, T. Kempen, A. Sarantopoulos, R. Dittmann, S. Menzel, V. Rana, S. Hoffmann-Eifert and S. Wiefels, *Effect of electron conduction on the read noise characteristics in ReRAM devices*, APL Materials **10** (2022) 101114, DOI: 10.1063/5.0109787 (cit. on p. 12).

[90] N. Kopperberg, S. Wiefels, S. Liberda, R. Waser and S. Menzel, *A Consistent Model for Short-Term Instability and Long-Term Retention in Filamentary Oxide-Based Memristive Devices*, ACS Applied Materials & Interfaces **13** (2021) 58066, DOI: 10.1021/acsami.1c14667 (cit. on p. 13).

[91] Z. Wei et al., "Demonstration of high-density ReRAM ensuring 10-year retention at 85°C based on a newly developed reliability model", *2011 International Electron Devices Meeting*, 2011 31.4.1, DOI: 10.1109/IEDM.2011.6131650 (cit. on p. 13).

[92] M. Zhao, H. Wu, B. Gao, Q. Zhang, W. Wu, S. Wang, Y. Xi, D. Wu, N. Deng, S. Yu, H.-Y. Chen and H. Qian, "Investigation of statistical retention of filamentary analog RRAM for neuromophic computing", *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017 39.4.1, DOI: 10.1109/IEDM.2017.8268522 (cit. on p. 13).

[93] Y. Y. Chen et al., "Improvement of data retention in HfO_2/Hf 1T1R RRAM cell under low operating current", *2013 IEEE International Electron Devices Meeting*, 2013 10.1.1, DOI: 10.1109/IEDM.2013.6724598 (cit. on p. 13).

[94] J. Sandrini et al., "OxRAM for embedded solutions on advanced node: scaling perspectives considering statistical reliability and design constraints", *2019 IEEE International Electron Devices Meeting (IEDM)*, 2019 30.5.1, DOI: 10.1109/IEDM19573.2019.8993484 (cit. on p. 13).

[95] C. Nail et al., "Understanding RRAM endurance, retention and window margin trade-off using experimental results and simulations", *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016 4.5.1, DOI: 10.1109/IEDM.2016.7838346 (cit. on p. 13).

[96] B. Govoreanu et al.,
"Performance and reliability of Ultra-Thin HfO_2-based RRAM (UTO-RRAM)",
2013 5th IEEE International Memory Workshop, 2013 48,
DOI: 10.1109/IMW.2013.6582095 (cit. on p. 13).

[97] Y. Chen, *ReRAM: History, Status, and Future*,
IEEE Transactions on Electron Devices **67** (2020) 1420, DOI: 10.1109/TED.2019.2961505
(cit. on pp. 14, 18).

[98] H. Jin, B. Lei, H. Liu, X. Liao, Z. Duan, C. Ye and Y. Zhang,
A Compilation Tool for Computation Offloading in ReRAM-Based CIM Architectures,
ACM Trans. Archit. Code Optim. **20** (2023), DOI: 10.1145/3617686 (cit. on p. 14).

[99] C. Bengel, L. Dixius, R. Waser, D. J. Wouters and S. Menzel,
it - Information Technology **65** (2023) 3, DOI: doi:10.1515/itit-2023-0018 (cit. on p. 14).

[100] P.-F. Chiu, M.-F. Chang, C.-W. Wu, C.-H. Chuang, S.-S. Sheu, Y.-S. Chen and M.-J. Tsai,
Low Store Energy, Low VDDmin, 8T2R Nonvolatile Latch and SRAM With Vertical-Stacked Resistive Memory (Memristor) Devices for Low Power Mobile Applications,
IEEE Journal of Solid-State Circuits **47** (2012) 1483, DOI: 10.1109/JSSC.2012.2192661
(cit. on p. 14).

[101] C.-P. Lo et al., "A ReRAM-based single-NVM nonvolatile flip-flop with reduced stress-time and write-power against wide distribution in write-time by using self-write-termination scheme for nonvolatile processors in IoT era",
2016 IEEE International Electron Devices Meeting (IEDM), 2016 16.3.1,
DOI: 10.1109/IEDM.2016.7838430 (cit. on p. 14).

[102] O. Golonzka et al., "Non-Volatile RRAM Embedded into 22FFL FinFET Technology",
2019 Symposium on VLSI Technology, 2019 T230, DOI: 10.23919/VLSIT.2019.8776570
(cit. on pp. 14, 15).

[103] R. Nair and J. Yoon,
"7 - The evolving role of storage-class memory in servers and large systems",
Semiconductor Memories and Systems, ed. by A. Redaelli and F. Pellizzer,
Woodhead Publishing Series in Electronic and Optical Materials, Woodhead Publishing, 2022
217, ISBN: 978-0-12-820758-1,
DOI: https://doi.org/10.1016/B978-0-12-820758-1.00002-9 (cit. on p. 14).

[104] F. Zahoor, T. Z. Azni Zulkifli and F. A. Khanday,
Resistive Random Access Memory (RRAM): an Overview of Materials, Switching Mechanism, Performance, Multilevel Cell (mlc) Storage, Modeling, and Applications,
Nanoscale Research Letters **15** (2020) 90, DOI: 10.1186/s11671-020-03299-9
(cit. on pp. 14, 19).

[105] A. W. Burks, H. H. Goldstine and J. von Neumann,
"Preliminary Discussion of the Logical Design of an Electronic Computing Instrument",
Papers of John von Neumann on computing and computer theory,
ed. by J. Von Neumann, W. Aspray and A. W. Burks, Cambridge, MA, USA: MIT Press, 1987
97, ISBN: 978-0-262220309,
URL: https://archive.org/details/papersofjohnvonn00vonn (cit. on p. 14).

[106] S. Mittal and S. Nag, *A survey of encoding techniques for reducing data-movement energy*, Journal of Systems Architecture **97** (2019) 373, DOI: https://doi.org/10.1016/j.sysarc.2018.11.001 (cit. on p. 14).

[107] P. Stanley-Marbell, V. C. Cabezas and R. P. Luijten, "Pinned to the walls — Impact of packaging and application properties on the memory and power walls", *IEEE/ACM International Symposium on Low Power Electronics and Design*, 2011 51, DOI: 10.1109/ISLPED.2011.5993603 (cit. on p. 14).

[108] A. Kagi, J. Goodman and D. Burger, "Memory Bandwidth Limitations of Future Microprocessors", *23rd Annual International Symposium on Computer Architecture (ISCA'96)*, 1996 78, DOI: 10.1145/232973.232983 (cit. on p. 14).

[109] W. A. Wulf and S. A. McKee, *Hitting the Memory Wall: Implications of the Obvious*, SIGARCH Comput. Archit. News **23** (1995) 20, DOI: 10.1145/216585.216588 (cit. on p. 14).

[110] J. L. Hennessy and D. A. Patterson, *Computer Architecture, Fifth Edition: A Quantitative Approach*, 5th, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011, ISBN: 012383872X (cit. on p. 14).

[111] *Volume of data/information created, captured, copied and consumed worldwide from 2010 to 202, with forecasts from 2021 to 2025 (in zettabytes)*, https://www.statista.com/statistics/871513/worldwide-data-created/, [Accessed 08-01-2024] (cit. on p. 14).

[112] N. Jones, *How to stop data centres from gobbling up the world's electricity*, Nature **561** (2018) 163 (cit. on p. 14).

[113] *AI and compute — openai.com*, https://openai.com/research/ai-and-compute, [Accessed 08-01-2024] (cit. on p. 14).

[114] F. Gül, *Addressing the sneak-path problem in crossbar RRAM devices using memristor-based one Schottky diode-one resistor array*, Results in Physics **12** (2019) 1091, DOI: https://doi.org/10.1016/j.rinp.2018.12.092 (cit. on p. 14).

[115] L. Shi, G. Zheng, B. B. Tian, B. Dkhil and C.-G. Duan, *Research progress of the solutions to the sneak path issue in memristor crossbar arrays*, Nanoscale Advances **2** (2020), DOI: 10.1039/D0NA00100G (cit. on pp. 14, 15).

[116] L. Zhang, S. Cosemans, D. J. Wouters, G. Groeseneken, M. Jurczak and B. Govoreanu, *Cell Variability Impact on the One-Selector One-Resistor Cross-Point Array Performance*, IEEE Transactions on Electron Devices **62** (2015) 3490, DOI: 10.1109/TED.2015.2439958 (cit. on p. 15).

[117] S. Kim, J. Zhou and W. D. Lu, *Crossbar RRAM Arrays: Selector Device Requirements During Write Operation*, IEEE Transactions on Electron Devices **61** (2014) 2820, DOI: 10.1109/TED.2014.2327514 (cit. on p. 15).

[118] D. A. Robayo et al.,
"Reliability and Variability of 1S1R OxRAM-OTS for High Density Crossbar Integration",
2019 IEEE International Electron Devices Meeting (IEDM), 2019 35.3.1,
DOI: 10.1109/IEDM19573.2019.8993439 (cit. on p. 15).

[119] X. Xu, J. Yu, T. Gong, J. Yang, J. Yin, D. Nian Dong, Q. Luo, J. Liu, Z. Yu, Q. Liu, H. Lv and M. Liu, "First Demonstration of OxRRAM Integration on 14nm FinFet Platform and Scaling Potential Analysis towards Sub-10nm Node",
2020 IEEE International Electron Devices Meeting (IEDM), 2020 24.3.1,
DOI: 10.1109/IEDM13553.2020.9371971 (cit. on p. 15).

[120] W. Wan et al., *A compute-in-memory chip based on resistive random-access memory*,
Nature **608** (2022) 504, DOI: 10.1038/s41586-022-04992-8 (cit. on p. 15).

[121] C.-C. Chou, Z.-J. Lin, C.-A. Lai, C.-I. Su, P.-L. Tseng, W.-C. Chen, W.-C. Tsai, W.-T. Chu, T.-C. Ong, H. Chuang, Y.-D. Chih and T.-Y. J. Chang,
"A 22nm 96KX144 RRAM Macro with a Self-Tracking Reference and a Low Ripple Charge Pump to Achieve a Configurable Read Window and a Wide Operating Voltage Range",
2020 IEEE Symposium on VLSI Circuits, 2020 1,
DOI: 10.1109/VLSICircuits18222.2020.9163014 (cit. on p. 15).

[122] C. Peters, F. Adler, K. Hofmann and J. Otterstedt,
"Reliability of 28nm embedded RRAM for consumer and industrial products",
2022 IEEE International Memory Workshop (IMW), 2022 1,
DOI: 10.1109/IMW52921.2022.9779300 (cit. on p. 15).

[123] T.-Y. Liu et al., "A 130.7mm2 2-layer 32Gb ReRAM memory device in 24nm technology",
2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers, 2013 210,
DOI: 10.1109/ISSCC.2013.6487703 (cit. on p. 15).

[124] X. Xu, J. Yu, T. Gong, J. Yang, J. Yin, D. Nian Dong, Q. Luo, J. Liu, Z. Yu, Q. Liu, H. Lv and M. Liu, "First Demonstration of OxRRAM Integration on 14nm FinFet Platform and Scaling Potential Analysis towards Sub-10nm Node",
2020 IEEE International Electron Devices Meeting (IEDM), 2020 24.3.1,
DOI: 10.1109/IEDM13553.2020.9371971 (cit. on pp. 15, 16).

[125] S. Pandit, C. Mandal and A. Patra, CRC Press, 2014 2, ISBN: 978-1-31521-610-2,
DOI: https://doi.org/10.1201/9781315216102 (cit. on p. 16).

[126] P. Drennan and C. McAndrew, "Understanding MOSFET mismatch for analog design",
Proceedings of the IEEE 2002 Custom Integrated Circuits Conference, 2002 449,
DOI: 10.1109/CICC.2002.1012872 (cit. on p. 16).

[127] P. Siegl, R. Buchty and M. Berekovic,
"Data-Centric Computing Frontiers: A Survey On Processing-In-Memory",
Proceedings of the Second International Symposium on Memory Systems, MEMSYS '16,
Alexandria, VA, USA: Association for Computing Machinery, 2016 295,
ISBN: 9781450343053, DOI: 10.1145/2989081.2989087 (cit. on p. 18).

[128] S. Mittal, G. Verma, B. Kaushik and F. A. Khanday,
A survey of SRAM-based in-memory computing techniques and applications,
Journal of Systems Architecture **119** (2021) 102276,
DOI: https://doi.org/10.1016/j.sysarc.2021.102276 (cit. on p. 18).

[129] G. Singh, L. Chelini, S. Corda, A. J. Awan, S. Stuijk, R. Jordans, H. Corporaal and A.-J. Boonstra, *Near-memory computing: Past, present, and future*, Microprocessors and Microsystems **71** (2019) 102868, DOI: https://doi.org/10.1016/j.micpro.2019.102868 (cit. on p. 18).

[130] S. Srinivasa, A. K. Ramanathan, X. Li, W.-H. Chen, S. K. Gupta, M.-F. Chang, S. Ghosh, J. Sampson and V. Narayanan, *ROBIN: Monolithic-3D SRAM for Enhanced Robustness with In-Memory Computation Support*, IEEE Transactions on Circuits and Systems I: Regular Papers **66** (2019) 2533, DOI: 10.1109/TCSI.2019.2897497 (cit. on p. 18).

[131] A. Agrawal, A. Jaiswal, C. Lee and K. Roy, *X-SRAM: Enabling In-Memory Boolean Computations in CMOS Static Random Access Memories*, IEEE Transactions on Circuits and Systems I: Regular Papers **65** (2018) 4219, DOI: 10.1109/TCSI.2018.2848999 (cit. on pp. 18, 19).

[132] J. Wang, X. Wang, C. Eckert, A. Subramaniyan, R. Das, D. Blaauw and D. Sylvester, *A 28-nm Compute SRAM With Bit-Serial Logic/Arithmetic Operations for Programmable In-Memory Vector Computing*, IEEE Journal of Solid-State Circuits **55** (2020) 76, DOI: 10.1109/JSSC.2019.2939682 (cit. on pp. 18, 19).

[133] F. Cai, J. M. Correll, S. H. Lee, Y. Lim, V. Bothra, Z. Zhang, M. P. Flynn and W. D. Lu, *A fully integrated reprogrammable memristor–CMOS system for efficient multiply–accumulate operations*, Nature Electronics **2** (2019) 290, DOI: 10.1038/s41928-019-0270-x (cit. on p. 18).

[134] C. Li, C. E. Graves, X. Sheng, D. Miller, M. Foltin, G. Pedretti and J. P. Strachan, *Analog content-addressable memories with memristors*, Nature Communications **11** (2020) 1638, DOI: 10.1038/s41467-020-15254-4 (cit. on p. 18).

[135] I. Giannopoulos, A. Singh, M. Le Gallo, V. P. Jonnalagadda, S. Hamdioui and A. Sebastian, *In-Memory Database Query*, Advanced Intelligent Systems **2** (2020) 2000141, DOI: https://doi.org/10.1002/aisy.202000141 (cit. on p. 18).

[136] M. A. Lebdeh, U. Reinsalu†, H. A. D. Nguyen, S. Wong and S. Hamdioui, "Memristive Device Based Circuits for Computation-in-Memory Architectures", *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2019 1, DOI: 10.1109/ISCAS.2019.8702542 (cit. on p. 18).

[137] H. A. D. Nguyen, J. Yu, M. A. Lebdeh, M. Taouil, S. Hamdioui and F. Catthoor, *A Classification of Memory-Centric Computing*, J. Emerg. Technol. Comput. Syst. **16** (2020), DOI: 10.1145/3365837 (cit. on pp. 18, 19).

[138] D. Ielmini and G. Pedretti, *Device and Circuit Architectures for In-Memory Computing*, Advanced Intelligent Systems **2** (2020) 2000040, DOI: https://doi.org/10.1002/aisy.202000040 (cit. on p. 18).

[139] G. Pedretti and D. Ielmini, *In-Memory Computing with Resistive Memory Circuits: Status and Outlook*, Electronics **10** (2021), DOI: 10.3390/electronics10091063 (cit. on p. 18).

[140] Z. Wang et al.,
Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing,
Nature Materials **16** (2017) 101, DOI: 10.1038/nmat4756 (cit. on p. 18).

[141] S. Siegel, T. Ziegler, Y. Bouhadjar, T. Tetzlaff, R. Waser, R. Dittmann and D. Wouters,
"Demonstration of Neuromorphic Sequence Learning on a Memristive Array",
Proceedings of the 2023 Annual Neuro-Inspired Computational Elements Conference,
NICE '23, San Antonio, TX, USA: Association for Computing Machinery, 2023 108,
ISBN: 9781450399470, DOI: 10.1145/3584954.3585000 (cit. on p. 18).

[142] P.-E. Gaillardon, L. Amarú, A. Siemon, E. Linn, R. Waser, A. Chattopadhyay and
G. De Micheli, "The Programmable Logic-in-Memory (PLiM) computer",
2016 Design, Automation & Test in Europe Conference & Exhibition (DATE), 2016 427
(cit. on p. 19).

[143] S. K. Kingra, V. Parmar, C.-C. Chang, B. Hudec, T.-H. Hou and M. Suri,
SLIM: Simultaneous Logic-in-Memory Computing Exploiting Bilayer Analog OxRAM Devices,
Scientific Reports **10** (2020) 2567, DOI: 10.1038/s41598-020-59121-0 (cit. on p. 19).

[144] G. Papandroulidakis, I. Vourkas, A. Abusleme, G. C. Sirakoulis and A. Rubio,
Crossbar-Based Memristive Logic-in-Memory Architecture,
IEEE Transactions on Nanotechnology **16** (2017) 491, DOI: 10.1109/TNANO.2017.2691713
(cit. on p. 19).

[145] G. Santoro, G. Turvani and M. Graziano,
New Logic-In-Memory Paradigms: An Architectural and Technological Perspective,
Micromachines **10** (2019), DOI: 10.3390/mi10060368 (cit. on p. 19).

[146] A. Coluccio, M. Vacca and G. Turvani,
Logic-in-Memory Computation: Is It Worth It? A Binary Neural Network Case Study,
Journal of Low Power Electronics and Applications **10** (2020),
DOI: 10.3390/jlpea10010007 (cit. on p. 19).

[147] W. Kautz, *Cellular Logic-in-Memory Arrays*,
IEEE Transactions on Computers **C-18** (1969) 719, DOI: 10.1109/T-C.1969.222754
(cit. on p. 19).

[148] M. Cofano, M. Vacca, G. Santoro, G. Causapruno, G. Turvani and M. Graziano,
Exploiting the Logic-In-Memory paradigm for speeding-up data-intensive algorithms,
Integration **66** (2019) 153, DOI: https://doi.org/10.1016/j.vlsi.2019.02.007
(cit. on p. 19).

[149] X. Si, W.-S. Khwa, J.-J. Chen, J.-F. Li, X. Sun, R. Liu, S. Yu, H. Yamauchi, Q. Li and
M.-F. Chang, *A Dual-Split 6T SRAM-Based Computing-in-Memory Unit-Macro With Fully
Parallel Product-Sum Operation for Binarized DNN Edge Processors*,
IEEE Transactions on Circuits and Systems I: Regular Papers **66** (2019) 4172,
DOI: 10.1109/TCSI.2019.2928043 (cit. on p. 19).

[150] A. Gebregiorgis, H. A. Du Nguyen, J. Yu, R. Bishnoi, M. Taouil, F. Catthoor and S. Hamdioui,
A Survey on Memory-Centric Computer Architectures,
J. Emerg. Technol. Comput. Syst. **18** (2022), DOI: 10.1145/3544974 (cit. on p. 19).

[151] L. Xie, H. Du Nguyen, J. Yu, A. Kaichouhi, M. Taouil, M. AlFailakawi and S. Hamdioui, "Scouting Logic: A Novel Memristor-Based Logic Design for Resistive Computing", *2017 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2017 176, DOI: 10.1109/ISVLSI.2017.39 (cit. on pp. 19, 33, 55).

[152] L. Guckert and E. E. Swartzlander, *MAD Gates—Memristor Logic Design Using Driver Circuitry*, IEEE Transactions on Circuits and Systems II: Express Briefs **64** (2017) 171, DOI: 10.1109/TCSII.2016.2551554 (cit. on pp. 19, 20).

[153] A. Siemon, S. Menzel, R. Waser and E. Linn, *A Complementary Resistive Switch-Based Crossbar Array Adder*, IEEE Journal on Emerging and Selected Topics in Circuits and Systems **5** (2015) 64, DOI: 10.1109/JETCAS.2015.2398217 (cit. on p. 19).

[154] R. Ben Hur and S. Kvatinsky, "Memristive memory processing unit (MPU) controller for in-memory processing", *2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE)*, 2016 1, DOI: 10.1109/ICSEE.2016.7806045 (cit. on pp. 19, 127).

[155] L. Cheng, Y. Li, K.-S. Yin, S.-Y. Hu, Y.-T. Su, M.-M. Jin, Z.-R. Wang, T.-C. Chang and X.-S. Miao, *Functional Demonstration of a Memristive Arithmetic Logic Unit (MemALU) for In-Memory Computing*, Advanced Functional Materials **29** (2019) 1905660, DOI: https://doi.org/10.1002/adfm.201905660 (cit. on p. 19).

[156] D. Fey, *Using the multi-bit feature of memristors for register files in signed-digit arithmetic units*, Semiconductor Science and Technology **29** (2014) 104008, DOI: 10.1088/0268-1242/29/10/104008 (cit. on p. 19).

[157] C. Bengel, F. Liu, Z. Chen, X. Zhao, R. Waser, H. Schmidt, N. Du and S. Menzel, *Ternary Łukasiewicz logic using memristive devices*, Neuromorphic Computing and Engineering **3** (2023) 044001, DOI: 10.1088/2634-4386/acfbf3 (cit. on p. 19).

[158] Z. Sun, S. Kvatinsky, X. Si, A. Mehonic, Y. Cai and R. Huang, *A full spectrum of computing-in-memory technologies*, Nature Electronics **6** (2023) 823, DOI: 10.1038/s41928-023-01053-4 (cit. on p. 19).

[159] E. Linn, R. Rosezin, C. Kügeler and R. Waser, *Complementary resistive switches for passive nanocrossbar memories*, Nature Materials **9** (2010) 403, DOI: 10.1038/nmat2748 (cit. on pp. 19, 86, 92).

[160] J. Reuben, *Rediscovering Majority Logic in the Post-CMOS Era: A Perspective from In-Memory Computing*, Journal of Low Power Electronics and Applications **10** (2020), DOI: 10.3390/jlpea10030028 (cit. on pp. 19, 69).

[161] S. Shirinzadeh, M. Soeken, P.-E. Gaillardon and R. Drechsler, "Fast logic synthesis for RRAM-based in-memory computing using Majority-Inverter Graphs", *2016 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2016 948 (cit. on pp. 19, 69).

[162] C. E. Graves et al.,
Memristor TCAMs Accelerate Regular Expression Matching for Network Intrusion Detection,
IEEE Transactions on Nanotechnology **18** (2019) 963, DOI: 10.1109/TNANO.2019.2936239
(cit. on p. 20).

[163] D. Reiser, M. Reichenbach, T. Rizzi, A. Baroni, M. Fritscher, C. Wenger, C. Zambelli and D. Bertozzi,
"Technology-Aware Drift Resilience Analysis of RRAM Crossbar Array Configurations",
2023 21st IEEE Interregional NEWCAS Conference (NEWCAS), 2023 1,
DOI: 10.1109/NEWCAS57931.2023.10198076 (cit. on p. 21).

[164] Y. Fang, Z. Yu, Z. Wang, T. Zhang, Y. Yang, Y. Cai and R. Huang,
Improvement of HfOx-Based RRAM Device Variation by Inserting ALD TiN Buffer Layer,
IEEE Electron Device Letters **39** (2018) 819, DOI: 10.1109/LED.2018.2831698
(cit. on p. 21).

[165] A. Fantini, L. Goux, R. Degraeve, D. Wouters, N. Raghavan, G. Kar, A. Belmonte, Y.-Y. Chen, B. Govoreanu and M. Jurczak, "Intrinsic switching variability in HfO2 RRAM",
2013 5th IEEE International Memory Workshop, 2013 30,
DOI: 10.1109/IMW.2013.6582090 (cit. on p. 21).

[166] E. Pérez, D. Maldonado, C. Acal, J. Ruiz-Castro, F. Alonso, A. Aguilera, F. Jiménez-Molinos, C. Wenger and J. Roldán, *Analysis of the statistics of device-to-device and cycle-to-cycle variability in TiN/Ti/Al:HfO2/TiN RRAMs*, Microelectronic Engineering **214** (2019) 104,
DOI: https://doi.org/10.1016/j.mee.2019.05.004 (cit. on p. 21).

[167] G. Pedretti, E. Ambrosi and D. Ielmini, "Conductance variations and their impact on the precision of in-memory computing with resistive switching memory (RRAM)",
2021 IEEE International Reliability Physics Symposium (IRPS), 2021 1,
DOI: 10.1109/IRPS46558.2021.9405130 (cit. on p. 21).

[168] P. Huang, Y. C. Xiang, Y. D. Zhao, C. Liu, B. Gao, H. Q. Wu, H. Qian, X. Y. Liu and J. F. Kang, "Analytic Model for Statistical State Instability and Retention Behaviors of Filamentary Analog RRAM Array and Its Applications in Design of Neural Network",
2018 IEEE International Electron Devices Meeting (IEDM), 2018 40.4.1,
DOI: 10.1109/IEDM.2018.8614567 (cit. on p. 21).

[169] A. Grossi, E. Nowak, C. Zambelli, C. Pellissier, S. Bernasconi, G. Cibrario, K. El Hajjam, R. Crochemore, J. Nodin, P. Olivo and L. Perniola,
"Fundamental variability limits of filament-based RRAM",
2016 IEEE International Electron Devices Meeting (IEDM), 2016 4.7.1,
DOI: 10.1109/IEDM.2016.7838348 (cit. on pp. 21, 26).

[170] C. Bengel, A. Siemon, F. Cüppers, S. Hoffmann-Eifert, A. Hardtdegen, M. von Witzleben, L. Hellmich, R. Waser and S. Menzel, *Variability-Aware Modeling of Filamentary Oxide-Based Bipolar Resistive Switching Cells Using SPICE Level Compact Models*,
IEEE Transactions on Circuits and Systems I: Regular Papers **67** (2020) 4618,
DOI: 10.1109/TCSI.2020.3018502 (cit. on pp. 21, 24, 25).

[171] F. M. Puglisi, T. Zanotti and P. Pavan,
Unimore Resistive Random Access Memory (RRAM) Verilog-A Model, 2019,
DOI: doi:/10.21981/15GF-KX29 (cit. on pp. 21, 23).

[172] G. Tanaka and R. Nakane, *Simulation platform for pattern recognition based on reservoir computing with memristor networks*, Scientific Reports **12** (2022) 9868, DOI: 10.1038/s41598-022-13687-z (cit. on p. 21).

[173] H. F. Inman and E. L. Bradley, *The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities*, Communications in Statistics - Theory and Methods **18** (1989) 3851, DOI: 10.1080/03610928908830127 (cit. on p. 22).

[174] A. Grossi, M. Coppetta, S. Aresu, A. Kux, T. Kern and R. Strenz, "28nm Data Memory with Embedded RRAM Technology in Automotive Microcontrollers", *2023 IEEE International Memory Workshop (IMW)*, 2023 1, DOI: 10.1109/IMW56887.2023.10145951 (cit. on pp. 22, 33, 86).

[175] P. Virtanen et al., *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nature Methods **17** (2020) 261, DOI: 10.1038/s41592-019-0686-2 (cit. on p. 23).

[176] X. Yin, J. Goudriaan, E. A. Lantinga, J. Vos and H. J. Spiertz, *A Flexible Sigmoid Function of Determinate Growth*, Annals of Botany **91** (2003) 361, ISSN: 03057364, 10958290, URL: http://www.jstor.org/stable/42795598 (visited on 30/11/2023) (cit. on p. 23).

[177] L. Roper, *Using Sigmoid and Double-Sigmoid Functions for Earth-States Transitions. 1* (2000), URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=718b9fdd2ed27d0193179ec0eb01a2ac622d25e3 (cit. on p. 23).

[178] X. Guan, S. Yu and H.-S. P. Wong, *A SPICE Compact Model of Metal Oxide Resistive Switching Memory With Variations*, IEEE Electron Device Letters **33** (2012) 1405, DOI: 10.1109/LED.2012.2210856 (cit. on p. 23).

[179] Z. Jiang and H.-S. P. Wong, *Stanford University Resistive-Switching Random Access Memory (RRAM) Verilog-A Model*, 2014, DOI: doi:/10.4231/D37H1DN48 (cit. on p. 23).

[180] A. Siemon, S. Menzel, A. Marchewka, Y. Nishi, R. Waser and E. Linn, "Simulation of TaOx-based complementary resistive switches by a physics-based memristive model", 2014 1420, ISBN: 978-1-4799-3432-4, DOI: 10.1109/ISCAS.2014.6865411 (cit. on p. 23).

[181] S. Menzel and C. Bengel, *JART – Jülich Aachen Resistive Switching Tools*, 2019, URL: http://www.emrl.de/JART.html (cit. on pp. 23, 25).

[182] *CreativeCommons CC BY 4.0 License*, Accessed: 06/05/2024, URL: https://creativecommons.org/licenses/by/4.0/ (cit. on pp. 24, 139).

[183] A. R. Genreith-Schriever and R. A. De Souza, *Field-enhanced ion transport in solids: Reexamination with molecular dynamics simulations*, Phys. Rev. B **94** (22 2016) 224304, DOI: 10.1103/PhysRevB.94.224304 (cit. on p. 24).

[184] C. R. Bengel, *Variability-aware compact modeling of valence-change-mechanism based devices for computation-in-memory*, Veröffentlicht auf dem Publikationsserver der RWTH Aachen University; Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, 2023, Dissertation: Rheinisch-Westfälische Technische Hochschule Aachen, 2023 1 Online, Diagramme, DOI: 10.18154/RWTH-2023-08986 (cit. on p. 25).

[185] *EUROPRACTICE IC Service*, Accessed: 09/07/2024, URL: https://europractice-ic.com/technologies/asics/cea-leti/ (cit. on p. 26).

[186] R. J. Baker, *CMOS : : circuit design, layout, and simulation*, Third edition, IEEE Press series on microelectronic systems, Hoboken, NJ: John Wiley & Sons, Inc., 2010 143, ISBN: 9780470891179 (cit. on pp. 27, 51).

[187] X. Liu, C. Bengel, F. Cüppers, O. Solfronk, B. Zhang, S. Hoffmann-Eifert, S. Menzel, R. Waser and S. Wiefels, *Effect of Transistor Transfer Characteristics on the Programming Process in 1T1R Configuration*, IEEE Transactions on Electron Devices **71** (2024) 2423, DOI: 10.1109/TED.2024.3370536 (cit. on pp. 27, 51).

[188] *AIXMATRIX Test System*, Accessed: 26/06/2024, URL: https://www.aixacct.com/en/testsysteme/paralleltestsysteme/aixmatrix/#matrix-mode-2 (cit. on p. 28).

[189] W. Shim, Y. Luo, J.-S. Seo and S. Yu, *Investigation of Read Disturb and Bipolar Read Scheme on Multilevel RRAM-Based Deep Learning Inference Engine*, IEEE Transactions on Electron Devices **67** (2020) 2318, DOI: 10.1109/TED.2020.2985013 (cit. on pp. 31, 50).

[190] L. Brackmann, A. Jafari, C. Bengel, M. Mayahinia, R. Waser, D. Wouters, S. Menzel and M. Tahoori, "A failure analysis framework of ReRAM In-Memory Logic operations", *2022 IEEE International Test Conference in Asia (ITC-Asia)*, 2022 67, DOI: 10.1109/ITCAsia55616.2022.00022 (cit. on pp. 33, 47, 93).

[191] M. Mayahinia, A. Jafari and M. B. Tahoori, "Voltage Tuning for Reliable Computation in Emerging Resistive Memories", *2022 IEEE 40th VLSI Test Symposium (VTS)*, 2022 1, DOI: 10.1109/VTS52500.2021.9794233 (cit. on p. 33).

[192] R. L. Wadsack, *Fault Modeling and Logic Simulation of CMOS and MOS Integrated Circuits*, Bell System Technical Journal **57** (1978) 1449, DOI: https://doi.org/10.1002/j.1538-7305.1978.tb02106.x (cit. on p. 33).

[193] S. Mukherjee, C. Weaver, J. Emer, S. Reinhardt and T. Austin, "A systematic methodology to compute the architectural vulnerability factors for a high-performance microprocessor", *Proceedings. 36th Annual IEEE/ACM International Symposium on Microarchitecture, 2003. MICRO-36.* 2003 29, DOI: 10.1109/MICRO.2003.1253181 (cit. on p. 33).

[194] C.-Y. Chen, H.-C. Shih, C.-W. Wu, C.-H. Lin, P.-F. Chiu, S.-S. Sheu and F. T. Chen, *RRAM Defect Modeling and Failure Analysis Based on March Test and a Novel Squeeze-Search Scheme*, IEEE Transactions on Computers **64** (2015) 180, DOI: 10.1109/TC.2014.12 (cit. on p. 33).

[195] P. Liu, Z. You, J. Kuang, Z. Hu, H. Duan and W. Wang, *Efficient March test algorithm for 1T1R cross-bar with complete fault coverage*, Electronics Letters **52** (2016) 1520, DOI: https://doi.org/10.1049/el.2016.1693 (cit. on p. 33).

[196] P. Liu, Z. You, J. Kuang, M. Elimu, S. Cai and W. Wang, *Logic operation-based Design for Testability method and parallel test algorithm for 1T1R crossbar*, Electronics Letters **53** (2017) 1631, DOI: https://doi.org/10.1049/el.2017.2424 (cit. on p. 33).

[197] X. Zhu, Z. Li, H. Long, H. Liu, Y. Wang and H. Xu, "Unsafe Writing Impacts on the Stateful Memristor Gates", *2020 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2020 1, DOI: 10.1109/ISCAS45731.2020.9180672 (cit. on pp. 33, 93, 94, 137).

[198] F. Parveen, Z. He, S. Angizi and D. Fan, "HielM: Highly flexible in-memory computing using STT MRAM", *2018 23rd Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2018 361, DOI: 10.1109/ASPDAC.2018.8297350 (cit. on p. 33).

[199] M. Escudero, I. Vourkas, A. Rubio and F. Moll, "Variability-Tolerant Memristor-based Ratioed Logic in Crossbar Array", *Proceedings of the 14th IEEE/ACM International Symposium on Nanoscale Architectures*, NANOARCH '18, Athens, Greece: Association for Computing Machinery, 2018 13, ISBN: 9781450358156, DOI: 10.1145/3232195.3232213 (cit. on pp. 33, 92).

[200] J. Yu, H. A. Du Nguyen, M. Abu Lebdeh, M. Taouil and S. Hamdioui, "Enhanced Scouting Logic: A Robust Memristive Logic Design Scheme", *2019 IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH)*, 2019 1, DOI: 10.1109/NANOARCH47378.2019.181296 (cit. on p. 33).

[201] M. Escudero, I. Vourkas, A. Rubio and F. Moll, *Memristive Logic in Crossbar Memory Arrays: Variability-Aware Design for Higher Reliability*, IEEE Transactions on Nanotechnology **18** (2019) 635, DOI: 10.1109/TNANO.2019.2923731 (cit. on pp. 33, 94).

[202] L. Qu, X. Cui and X. Cui, "A Testability Enhancement Method for the Memristor Ratioed Logic Circuits", *2020 IEEE 29th Asian Test Symposium (ATS)*, 2020 1, DOI: 10.1109/ATS49688.2020.9301537 (cit. on p. 33).

[203] W.-H. Chen, W.-J. Lin, L.-Y. Lai, S. Li, C.-H. Hsu, H.-T. Lin, H.-Y. Lee, J.-W. Su, Y. Xie, S.-S. Sheu and M.-F. Chang, "A 16Mb dual-mode ReRAM macro with sub-14ns computing-in-memory and memory functions enabled by self-write termination scheme", *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017 28.2.1, DOI: 10.1109/IEDM.2017.8268468 (cit. on pp. 33, 55).

[204] L. Brackmann, T. Ziegler, D. J. Wouters and S. Menzel, *Experimental Verification and Evaluation of Non-Stateful Logic Gates in Resistive RAM*, IEEE Transactions on Circuits and Systems I: Regular Papers **72** (2025) 2029, DOI: 10.1109/TCSI.2024.3486376 (cit. on pp. 33, 69).

[205] H. Padberg, A. Regev, G. Piccolboni, A. Bricalli, G. Molas, J. F. Nodin and S. Kvatinsky, *Experimental Demonstration of Non-Stateful In-Memory Logic With 1T1R OxRAM Valence Change Mechanism Memristors*, IEEE Transactions on Circuits and Systems II: Express Briefs **71** (2024) 395, DOI: 10.1109/TCSII.2023.3302235 (cit. on p. 33).

[206] L. Brackmann, T. Ziegler, A. Jafari, D. J. Wouters, M. B. Tahoori and S. Menzel, *Improved Arithmetic Performance by Combining Stateful and Non-Stateful Logic in Resistive Random Access Memory 1T–1R Crossbars*, Advanced Intelligent Systems **6** (2024) 2300579, DOI: https://doi.org/10.1002/aisy.202300579 (cit. on pp. 33, 77, 92, 100, 108, 137, 139).

[207] J.-K. Park, S.-Y. Kim, J.-M. Baek, D.-J. Seo, J.-H. Chun and K.-W. Kwon, "Analysis of resistance variations and variance-aware read circuit for cross-point ReRAM", *2013 5th IEEE International Memory Workshop*, 2013 112, DOI: 10.1109/IMW.2013.6582111 (cit. on p. 34).

[208] V. G. Karpov and D. Niraula, *OFF State Conduction in Filamentary RRAM*, IEEE Electron Device Letters **40** (2019) 550, DOI: 10.1109/LED.2019.2899258 (cit. on p. 43).

[209] A. Fantini, D. J. Wouters, R. Degraeve, L. Goux, L. Pantisano, G. Kar, Y.-Y. Chen, B. Govoreanu, J. A. Kittl, L. Altimime and M. Jurczak, "Intrinsic Switching Behavior in HfO2 RRAM by Fast Electrical Measurements on Novel 2R Test Structures", *2012 4th IEEE International Memory Workshop*, 2012 1, DOI: 10.1109/IMW.2012.6213646 (cit. on p. 43).

[210] J. Kim, T. Ahmed, H. Nili, J. Yang, D. S. Jeong, P. Beckett, S. Sriram, D. C. Ranasinghe and O. Kavehei, *A Physical Unclonable Function With Redox-Based Nanoionic Resistive Memory*, IEEE Transactions on Information Forensics and Security **13** (2018) 437, DOI: 10.1109/TIFS.2017.2756562 (cit. on p. 43).

[211] A. Chen, *A Comprehensive Crossbar Array Model With Solutions for Line Resistance and Nonlinear Device Characteristics*, IEEE Transactions on Electron Devices **60** (2013) 1318, DOI: 10.1109/TED.2013.2246791 (cit. on pp. 46, 47).

[212] T. Zanotti, F. M. Puglisi and P. Pavan, "Circuit Reliability Analysis of RRAM-based Logic-in-Memory Crossbar Architectures Including Line Parasitic Effects, Variability, and Random Telegraph Noise", *2020 IEEE International Reliability Physics Symposium (IRPS)*, 2020 1, DOI: 10.1109/IRPS45951.2020.9128343 (cit. on p. 46).

[213] K. Tang and E. Friedman, "Lumped versus distributed RC and RLC interconnect impedances", *Proceedings of the 43rd IEEE Midwest Symposium on Circuits and Systems (Cat.No.CH37144)*, vol. 1, 2000 136, DOI: 10.1109/MWSCAS.2000.951604 (cit. on p. 46).

[214] S. A. Thomas, S. K. Vohra, R. Kumar, R. Sharma and D. M. Das, "Analysis of Parasitics on CMOS based Memristor Crossbar Array for Neuromorphic Systems", *2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2021 309, DOI: 10.1109/MWSCAS47672.2021.9531867 (cit. on pp. 47, 48).

[215] J. Chen, S. Yang, H. Wu, G. Indiveri and M. Payvand, *Scaling Limits of Memristor-Based Routers for Asynchronous Neuromorphic Systems*, IEEE Transactions on Circuits and Systems II: Express Briefs **71** (2024) 1576, DOI: 10.1109/TCSII.2023.3343292 (cit. on p. 47).

[216] Y. Li, W. Chen, W. Lu and R. Jha, “Impact of coupling capacitance on read operation of RRAM devices in 1D1R crossbar architectures”,
2014 IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS), 2014 989, DOI: 10.1109/MWSCAS.2014.6908583 (cit. on p. 48).

[217] F. M. Puglisi, L. Larcher, P. Pavan, A. Padovani and G. Bersuker,
“Instability of HfO2 RRAM devices: Comparing RTN and cycling variability”,
2014 IEEE International Reliability Physics Symposium, 2014 MY.5.1,
DOI: 10.1109/IRPS.2014.6861160 (cit. on p. 49).

[218] P. Stasner, N. Kopperberg, K. Schnieders, T. Hennen, S. Wiefels, S. Menzel, R. Waser and D. J. Wouters, *Reliability effects of lateral filament confinement by nano-scaling the oxide in memristive devices*, Nanoscale Horizons **9** (5 2024) 764, DOI: 10.1039/D3NH00520H (cit. on p. 49).

[219] C. Bengel, J. Mohr, S. Wiefels, A. Singh, A. Gebregiorgis, R. Bishnoi, S. Hamdioui, R. Waser, D. Wouters and S. Menzel,
Reliability aspects of binary vector-matrix-multiplications using ReRAM devices,
Neuromorphic Computing and Engineering **2** (2022), DOI: 10.1088/2634-4386/ac6d04 (cit. on pp. 50, 72).

[220] R. Degraeve et al., “Dynamic ‘hour glass’ model for SET and RESET in HfO2 RRAM”,
2012 Symposium on VLSI Technology (VLSIT), 2012 75,
DOI: 10.1109/VLSIT.2012.6242468 (cit. on p. 50).

[221] C. Nail et al., “Understanding RRAM endurance, retention and window margin trade-off using experimental results and simulations”,
2016 IEEE International Electron Devices Meeting (IEDM), 2016 4.5.1,
DOI: 10.1109/IEDM.2016.7838346 (cit. on pp. 51, 102).

[222] M. Azzaz et al., “Endurance/Retention Trade Off in HfOx and TaOx Based RRAM”,
2016 IEEE 8th International Memory Workshop (IMW), 2016 1,
DOI: 10.1109/IMW.2016.7495268 (cit. on pp. 51, 102).

[223] G. Vinuesa, H. García, E. Pérez, C. Wenger, I. Íñiguez-de-la-Torre, T. González, S. Dueñas and H. Castán, *On the Asymmetry of Resistive Switching Transitions*, Electronics **13** (2024),
DOI: 10.3390/electronics13132639 (cit. on p. 51).

[224] C. Bengel, K. Zhang, J. Mohr, T. Ziegler, S. Wiefels, R. Waser, D. Wouters and S. Menzel,
Tailor-made synaptic dynamics based on memristive devices,
Frontiers in electronic materials **3** (2023) 1061269, DOI: 10.3389/femat.2023.1061269 (cit. on p. 51).

[225] Z.-R. Wang, Y.-T. Su, Y. Li, Y.-X. Zhou, T.-J. Chu, K.-C. Chang, T.-C. Chang, T.-M. Tsai, S. M. Sze and X.-S. Miao,
Functionally Complete Boolean Logic in 1T1R Resistive Random Access Memory,
IEEE Electron Device Letters **38** (2017) 179, DOI: 10.1109/LED.2016.2645946 (cit. on p. 77).

[226] T. Ziegler, R. Waser, D. J. Wouters and S. Menzel,
In-Memory Binary Vector–Matrix Multiplication Based on Complementary Resistive Switches,
Advanced Intelligent Systems **2** (2020) 2000134,
DOI: https://doi.org/10.1002/aisy.202000134 (cit. on pp. 77, 127).

[227] W. Gosney, *Subthreshold drain leakage currents in MOS field-effect transistors*, IEEE Transactions on Electron Devices **19** (1972) 213, DOI: 10.1109/T-ED.1972.17399 (cit. on p. 77).

[228] R. Rao, A. Srivastava, D. Blaauw and D. Sylvester, *Statistical analysis of subthreshold leakage current for VLSI circuits*, IEEE Transactions on Very Large Scale Integration (VLSI) Systems **12** (2004) 131, DOI: 10.1109/TVLSI.2003.821549 (cit. on p. 77).

[229] J. Borghetti, G. S. Snider, P. J. Kuekes, J. J. Yang, D. R. Stewart and R. S. Williams, *'Memristive' switches enable 'stateful' logic operations via material implication*, Nature **464** (2010) 873, DOI: 10.1038/nature08940 (cit. on p. 92).

[230] S. Gupta, M. Imani and T. Rosing, "FELIX: Fast and Energy-Efficient Logic in Memory", *2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2018 1, DOI: 10.1145/3240765.3240811 (cit. on pp. 92, 116).

[231] P. Huang, J. Kang, Y. Zhao, S. Chen, R. Han, Z. Zhou, Z. Chen, W. Ma, M. Li, L. Liu and X. Liu, *Reconfigurable Nonvolatile Logic Operations in Resistance Switching Crossbar Array for Large-Scale Circuits*, Advanced Materials **28** (2016) 9758, DOI: https://doi.org/10.1002/adma.201602418 (cit. on pp. 92, 115, 116, 118).

[232] R. Rahimi Disfani, N. TaheriNejad and M. Valinataj, "Operational Conditions Analysis for Memristive Stateful Logics - A Study on IMPLY and TMSL", *2022 20th IEEE Interregional NEWCAS Conference (NEWCAS)*, 2022 480, DOI: 10.1109/NEWCAS52662.2022.9841969 (cit. on pp. 92, 95, 118).

[233] D. D. Weller, M. Hefenbrock, M. S. Golanbari, M. Beigl, J. Aghassi-Hagmann and M. B. Tahoori, *Bayesian Optimized Mixture Importance Sampling for High-Sigma Failure Rate Estimation*, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **39** (2020) 2772, DOI: 10.1109/TCAD.2019.2961321 (cit. on p. 93).

[234] D. J. Wouters, L. Brackmann, A. Jafari, C. Bengel, M. Mayahinia, R. Waser, S. Menzel and M. Tahoori, "Reliability of Computing-In-Memory Concepts Based on Memristive Arrays", *2022 International Electron Devices Meeting (IEDM)*, 2022 5.3.1, DOI: 10.1109/IEDM45625.2022.10019423 (cit. on pp. 93, 112).

[235] M. Shirasawa, M. E. Dlamini and Y. Kamakura, "Kinetic Monte Carlo simulation for switching probability of ReRAM", *2016 IEEE International Meeting for Future of Electron Devices, Kansai (IMFEDK)*, 2016 1, DOI: 10.1109/IMFEDK.2016.7521667 (cit. on pp. 93, 94).

[236] Z. Chai, W. Zhang, R. Degraeve, S. Clima, F. Hatem, J. F. Zhang, P. Freitas, J. Marsland, A. Fantini, D. Garbin, L. Goux and G. S. Kar, *Dependence of Switching Probability on Operation Conditions in GexSe1–x Ovonic Threshold Switching Selectors*, IEEE Electron Device Letters **40** (2019) 1269, DOI: 10.1109/LED.2019.2924270 (cit. on pp. 93, 94).

[237] C. Bengel, S. Wiefels, V. Rana, H.-Y. Chen, Q.-T. Zhao, R. Waser, H. Padberg, F. Xi and S. Menzel, “Experimental and Theoretical Analysis of Stateful Logic in Passive and Active Crossbar Arrays for Computation-in-Memory”, *2022 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2022 2792, DOI: 10.1109/ISCAS48785.2022.9937565 (cit. on p. 93).

[238] M. Maestro-Izquierdo, J. Martin-Martinez, A. Crespo Yepes, M. Escudero, R. Rodriguez, M. Nafria, X. Aymerich and A. Rubio, *Experimental Verification of Memristor-Based Material Implication NAND Operation*, IEEE Transactions on Emerging Topics in Computing **7** (2019) 545, DOI: 10.1109/TETC.2017.2760929 (cit. on p. 95).

[239] B. Hoffer, V. Rana, S. Menzel, R. Waser and S. Kvatinsky, *Experimental Demonstration of Memristor-Aided Logic (MAGIC) Using Valence Change Memory (VCM)*, IEEE Transactions on Electron Devices **67** (2020) 3115, DOI: 10.1109/TED.2020.3001247 (cit. on pp. 98, 115, 116).

[240] B. Hoffer, N. Wainstein, C. M. Neumann, E. Pop, E. Yalon and S. Kvatinsky, *Stateful Logic Using Phase Change Memory*, IEEE Journal on Exploratory Solid-State Computational Devices and Circuits **8** (2022) 77, DOI: 10.1109/JXCDC.2022.3219731 (cit. on p. 98).

[241] B. Hoffer and S. Kvatinsky, “Performing Stateful Logic Using Spin-Orbit Torque (SOT) MRAM”, *2022 IEEE 22nd International Conference on Nanotechnology (NANO)*, 2022 571, DOI: 10.1109/NANO54668.2022.9928606 (cit. on p. 98).

[242] A. Bende et al., “Experimental Validation of Memristor-Aided Logic Using 1T1R TaOx RRAM Crossbar Array”, *2024 37th International Conference on VLSI Design and 2024 23rd International Conference on Embedded Systems (VLSID)*, 2024 565, DOI: 10.1109/VLSID60093.2024.00100 (cit. on pp. 108, 115).

[243] Y. S. Kim, M. W. Son and K. M. Kim, *Memristive Stateful Logic for Edge Boolean Computers*, Advanced Intelligent Systems **3** (2021) 2000278, DOI: https://doi.org/10.1002/aisy.202000278 (cit. on p. 109).

[244] W. Shen, P. Huang, M. Fan, R. Han, Z. Zhou, B. Gao, H. Wu, H. Qian, L. Liu, X. Liu, X. Zhang and J. Kang, *Stateful Logic Operations in One-Transistor-One- Resistor Resistive Random Access Memory Array*, IEEE Electron Device Letters **40** (2019) 1538, DOI: 10.1109/LED.2019.2931947 (cit. on pp. 118, 127).

[245] X. Zhang, F. Wei, X. Liu and X. Cui, “Design and Implementation of Full Adder in One-Transistor-One-Resistor RRAM Array”, *2021 IEEE 14th International Conference on ASIC (ASICON)*, 2021 1, DOI: 10.1109/ASICON52560.2021.9620360 (cit. on pp. 118, 127).

[246] P. Inglese, E. I. Vatajelu and G. Di Natale, “On the Limitations of Concatenating Boolean Operations in Memristive-Based Logic-In-Memory Solutions”, *2021 16th International Conference on Design & Technology of Integrated Systems in Nanoscale Era (DTIS)*, 2021 1, DOI: 10.1109/DTIS53253.2021.9505096 (cit. on p. 137).

[247] D. T. Franco, M. C. Vasconcelos, L. Naviner and J.-F. Naviner,
Signal probability for reliability evaluation of logic circuits,
Microelectronics Reliability **48** (2008), 19th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF 2008) 1586,
DOI: https://doi.org/10.1016/j.microrel.2008.07.002 (cit. on p. 137).

[248] S. Vemuru and S. Vittala,
"Reliability Analysis of Nano-CMOS Logic Gates Under Process Induced Variations",
NAECON 2018 - IEEE National Aerospace and Electronics Conference, 2018 350,
DOI: 10.1109/NAECON.2018.8556702 (cit. on p. 142).

List of Figures

2.1 Flowchart for the taxonomy of solid-state memories 6
2.2 VCM switching mechanism and *I*-*V* characteristics 8
2.3 Circuit diagram of an $M \times N$ passive crossbar array 15
2.4 Possible active 1T-1R crossbar arrangements 17
2.5 JART VCM v1b Equivalent Circuit Diagram 24
2.6 Schematic circuit of MAD-200 test chip measurement setup 27
2.7 Circuit Layout of Elmos $350\,\mathrm{nm}$ test chip 28

3.1 Fundamental readout operation in a 1T-1R bitcell 31
3.2 Distribution overlap for single device readout 34
3.3 Experimental voltage condition testing for 1T-1R device programming 36
3.4 Experimental example of current evaluation for non-stateful single device readout . . 37
3.5 Experimental current means for non-stateful single device readout 37
3.6 Histogram and CDF plot for experimental non-stateful single device readout 38
3.7 Threshold-dependent error probability for non-stateful single device readout 39
3.8 Histogram and CDF plot for simulated non-stateful single device readout 41
3.9 Experimental voltage dependency of single device readout currents 42
3.10 Simulated non-linearity of ReRAM equivalent resistances 44
3.11 Experimental non-linearity in single device equivalent resistances 45
3.12 Experimental voltage-dependent single device current histograms 46
3.13 Simulated LRS readout currents with included parasitic RC elements 47
3.14 Experimental position-dependent single device readout currents 48
3.15 Experimental poxplots for single device read current variability 49
3.16 Experimental readout current variability for selected HRS states 50
3.17 Experimental CDF plots for single device read disturb 52
3.18 Experimental time-current data for Elmos readout measurement in SET and RESET direction 53
3.19 Experimental current data for readout measurement in SET and RESET direction . . 54
3.20 Experimental current means for non-stateful 2-Input readout 56
3.21 Current means for simulated non-stateful 2-Input readout 57
3.22 Equivalent circuit diagram for two-device readout 59
3.23 Resistance ratio dependent equivalent two-device readout current 59
3.24 Two-device readout current margins 61
3.25 Experimental voltage-dependent two-device readout currents 62
3.26 Experimental voltage-time correlation between two-device readout currents 63

3.27 Experimental temporal evolution of the equivalent bitcell resistance for two-device readout . . . 64
3.28 Experimental two-device readout current distribution fitting . . . 65
3.29 Simulated two-device voltage-dependent readout currents . . . 66
3.30 Experimental absolute position dependency for two-device readout currents . . . 67
3.31 Experimental relative position dependency for two-device readout currents . . . 68
3.32 Experimental current means and histograms for non-stateful three-device readout . . . 70
3.33 Experimental current means for non-stateful five-device readout . . . 72
3.34 Current means for simulated non-stateful three-device readout . . . 74
3.35 Current means for simulated non-stateful five-device readout . . . 75
3.36 Transistor-specific single device readout operation . . . 78
3.37 Experimental time-current data for single transistor-specific logic in Elmos . . . 78
3.38 Experimental current means for single device transistor-specific readout . . . 79
3.39 Simulated current means for single device transistor-specific readout . . . 81
3.40 Experimental current means for non-stateful transistor-specific two-device readout . . . 83
3.41 Current histograms and fitted distributions for non-stateful transistor-specific two-device readout . . . 84
3.42 Current means for non-stateful two-device readout . . . 85

4.1 Fundamental switching-based logic with one input and one output . . . 90
4.2 Primitive gates for switching-based logic operations . . . 92
4.3 Simulated voltage-dependent single-device switching probability . . . 93
4.4 Simulated resistive voltage divider output voltages . . . 94
4.5 Resistance-dependent conditional voltage divider . . . 96
4.6 Simulated transient voltages in conditional voltage divider . . . 97
4.7 Simulated voltage-dependent two-device RESET-based switching . . . 99
4.8 Primitive 1R and 1T-1R gate for SET-based switching logic operations . . . 100
4.9 Simulated voltage-dependent two-device SET-based switching . . . 101
4.10 Experimental Device-to-Device variability of a two-device SET-based switching . . . 103
4.11 Experimental voltage-dependent two-device SET-based switching output change . . . 105
4.12 Calculated error probabilities of a simulated two-device SET-based switching . . . 107
4.13 Experimental transient voltage signals of a two-device SET-based switching . . . 108
4.14 Resistance state dependencies of a two-device SET-based switching . . . 109
4.15 Experimental voltage-dependent two-device SET-based switching input change . . . 110
4.16 Experimental Device-to-Device variability of two-device SET-based switching inputs 111
4.17 Calculated input error probabilities of a simulated two-device SET-based switching . 113
4.18 Switching-based error framework benchmarking . . . 114
4.19 Primitive gate for extended SET-based switching logic operations . . . 115
4.20 Calculated error probabilities and benchmarking of a simulated two-device SET-based NIMP switching . . . 117
4.21 Simulated output current change in two-device one-resistor logic operation . . . 119
4.22 Experimental transistor-specific two-device SET-based switching . . . 122
4.23 Simulated transistor-specific two-device SET-based switching . . . 123

5.1 Crossbar arrangement for parallel gate evaluation . . . 128
5.2 Experimental readout current change in parallel stateful logic operation . . . 129

5.3 Experimental position dependence of parallel two-device SET-based switching outputs 130
5.4 Experimental position dependence of parallel two-device SET-based switching inputs 131
5.5 Input disturb of parallel two-device SET-based switching inputs 132
5.6 Experimental position dependence of parallel two-device SET-based outputs for non-reprogrammed inputs 134
5.7 Experimental position dependence of parallel two-device SET-based inputs for non-reprogrammed inputs 135
5.8 Simulated error probabilities for parallel stateful gates 136
5.9 Error propagation for concatenated logic operations 138
5.10 Abstracted operation flow for a combined stateful and non-stateful Carry-Ripple full adder 139

List of Tables

2.1 ReRAM JART v1b simulation parameters 25
2.2 Transistor simulation parameters for XFAB PDK 26

3.1 Logic truth table for a one-Input logic IDENTITY / NOT 35
3.2 Voltage settings used for programming MAD-200 1T-1R devices 35
3.3 Variability parameters of simulated 1T-1R current distribution. 40
3.4 Experimental voltage-dependent singe device error probabilities 43
3.5 Logic truth table for the Two-Input logic functions OR, AND and XOR 55
3.6 Error probabilities for experimental 2-Input OR, AND and XOR 57
3.7 Error probabilities for simulated 2-Input OR, AND and XOR 58
3.8 Maximum error probabilities for simulated and experimental 2-Input OR, AND and XOR 58
3.9 Theoretical current calculation for two-device readout 59
3.10 Logic truth table for the 3-Input Majority function 69
3.11 Error probabilities for measured three-device non-stateful Majority (MAJ) 71
3.12 Experimental maximum error probabilities for measured 2-bit readout 71
3.13 Experimental error probabilities for the five-device non-stateful Majority (MAJ) . . . 73
3.14 Experimental maximum error probabilities for measured $2\,1/2$-bit/6-state readout . . 73
3.15 Maximum error probabilities for simulated 2-bit readout 75
3.16 Experimental maximum error probabilities for simulated $2\,1/2$-bit/6-state readout . . 76
3.17 Logic truth table for the single device transistor-specific 1T-1R AND operation . . . 77
3.18 Error probabilities for the experimental single device transistor-specific readout . . . 80
3.19 Logic truth table for the two-device transistor-specific 1T-1R operations 82
3.20 Maximum error probabilities for simulated and measured transistor-specific 2-Input OR and AND 84

4.1 Logic truth table for a switching-based operation with one input and one output . . . 90
4.2 Logic truth table for two-input switching operations 96
4.3 Device-dependent experimental switching rates (SR) for a SET-based OR 102
4.4 Voltage-Dependent experimental switching rates (SR) for a SET-based OR 104
4.5 Logic truth table for a switching-based two-input logic NIMP, NOT and XOR 115
4.6 Calculated error probabilities for the simulated two-device one-resistor SET-based operations outputs 118
4.7 Calculated error probabilities for the simulated two-device one-resistor SET-based operations inputs 120
4.8 Calculated error probabilities for the transistor specific two-device SET-based operations 124

5.1 Position-dependent experimental switching rates (SR) for a 12-bit parallel SET-based OR 133
5.2 Absolute and relative error for the simulated Carry-Ripple Full Adder for different input lengths N 140

Acronyms

AE active electrode. 7

C2C cycle-to-cycle. 10

CIM Computing-In-Memory. 7

D2D device-to-device. 10

HRS high-resistive state. 7

LRS low-resistive state. 7

MIEC Mixed-Ion-Electron Conductor. 6

MIM Metal-Insulator-Metal. 6

MOSFET Metal-Oxide-Semiconductor Field-Effect Transistor. 1, 5

OE ohmic electrode. 7

RAM Random Access Memory. 5

ReRAM Redox-based Random Access Memory. 5

VCM Valence Change Mechanism. 5

Supervised Students

Bachelor Theses

Jinyun Park, “Modellierung der Schaltvariabilität von resistiv schaltenden Bauelementen für lesebasierte Logikoperationen”, Materialwissenschaften, March 2025.

List of Publications

Articles in Peer Reviewed Journals

F. Liu, **L. Brackmann**, X. Zhao, N. Du, R. Waser and S. Menzel (2025), "Memristive Ternary Łukasiewicz Logic based on Reading-based Ratioed Resistive States (3R)," in Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 383, no. 2288, DOI: 10.1098/rsta.2023.0397

L. Brackmann, T. Ziegler, D. J. Wouters and S. Menzel (2025), "Experimental Verification and Evaluation of Non-Stateful Logic Gates in Resistive RAM," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 72, no. 5, pp. 2029-2038, DOI: 10.1109/TCSI.2024.3486376

T. Hennen, **L. Brackmann**, T. Ziegler, S. Siegel, S. Menzel, R. Waser, D. J. Wouters and D. Bedau (2024), "Synaptogen: A Cross-Domain Generative Device Model for Large-Scale Neuromorphic Circuit Design," in IEEE Transactions on Electron Devices, vol. 71, no. 9, pp. 5345-5353, Sept. 2024, DOI: 10.1109/TED.2024.3427616.

L. Brackmann, T. Ziegler, A. Jafari, D. J. Wouters, M. Tahoori and S. Menzel (2023), "Improved Arithmetic Performance by Combining Stateful and Non-Stateful Logic in Resistive Random Access Memory 1T–1R Crossbars," in Advanced Intelligent Systems, vol. 6, no. 3, March 2024, 2300579, DOI: 10.1002/aisy.202300579.

Articles in Conference Proceedings

T. Ziegler, **L. Brackmann**, T. Hennen, C. Bengel, S. Menzel and D. J. Wouters (2023), "Eliminating Capacitive Sneak Paths in Associative Capacitive Networks based on Complementary Resistive Switches for In-Memory Computing," in 2023 IEEE International Memory Workshop (IMW), Monterey, CA, USA, 2023, pp. 1-4, DOI: 10.1109/IMW56887.2023.10145947.

L. Brackmann, T. Ziegler, A. Jafari, D. J. Wouters, M. Tahoori and S. Menzel (2023), "Design Limitations in Oxide-Based Memristive Ternary Content Addressable Memories," in 2023 IEEE International Symposium on Circuits and Systems (ISCAS), Monterey, CA, USA, 2023, pp. 1-5, DOI: 10.1109/ISCAS46773.2023.10181488.

D. J. Wouters, **L. Brackmann**, A. Jafari, C. Bengel, M. Mayahinia, R. Waser, S. Menzel and M. B. Tahoori (2022), "Reliability of Computing-In-Memory Concepts Based on Memristive Arrays," in 2022 International Electron Devices Meeting (IEDM), San Francisco, CA, USA, 2022, pp. 5.3.1-5.3.4, DOI: 10.1109/IEDM45625.2022.10019423.

L. Brackmann, A. Jafari, C. Bengel, M. Mayahinia, R. Waser, D. J. Wouters, S. Menzel and M. B. Tahoori (2022), "A failure analysis framework of ReRAM In-Memory Logic operations," in 2022 IEEE International Test Conference in Asia (ITC-Asia), Taipei, Taiwan, 2022, pp. 67-72, DOI: 10.1109/ITCAsia55616.2022.00022.

Conference Talks and Posters

L. Brackmann, T. Ziegler, M. Najjar, S. Menzel and D. J. Wouters (2024), "Improved Stateful Logic Designs based on Memristive 1T-1R Arrays," at MEMRISYS 2024, Poster, Seoul, Republic of Korea, November 10th - 13th, 2024.

L. Brackmann, T. Ziegler, M. Najjar and D. J. Wouters (2024), "Performing reliable Logic Computing-in-Memory based on resistive switching devices," at International Conference on Neuromorphic Computing and Engineering (ICNCE 2024), Poster, Aachen, Germany, June 3rd - 6th, 2024.

F. Liu, X. Zhao, **L. Brackmann** and S. Menzel (2024), "Ternary Łukasiewicz Logic for Computing in Memory with Resistive Random Access Memory (RRAM)," at International Conference on Neuromorphic Computing and Engineering (ICNCE 2024), Poster, Aachen, Germany, June 3rd - 6th, 2024.

T. Hennen, **L. Brackmann**, T. Ziegler, S. Siegel, S. Menzel, R. Waser, D. J. Wouters and D. Bedau (2024), "A generative device model for large-scale neuromorphic circuit design," at International Conference on Neuromorphic Computing and Engineering (ICNCE 2024), Poster, Aachen, Germany, June 3rd - 6th, 2024.