Harnessing Stochasticity and Negative Differential Resistance for Unconventional Computation

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

vorgelegt von

Tyler Aaron Hennen

M.Sc.

aus Minnesota (USA)

Berichter: Univ.-Prof Dr.-Ing. Rainer Waser
Univ.-Prof Dr.-Ing. Tobias Gemmeke

Tag der mündlichen Prüfung: 11.07.2023

Abstract

Recently, there has been a resurgence of interest in materials with unusual electronic properties such as strong nonlinearity, hysteresis, and memory. This interest is due in part to the end of Moore scaling as well as the emergence of novel computing architectures. Currently, computational performance is limited by the memory bottleneck, as physical memory is not fast or large enough to feed the central processing unit (CPU) pipeline. One alternative is to introduce a new tier of memory that must be substantially faster and more scalable than existing Flash storage. Another approach is to develop schemes that take advantage of in-memory computation, as in the brain-inspired concepts of neuromorphic computing (NC). To reach their full potential, each of these strategies rely on the ability of new classes of memory technologies to exploit physical mechanisms yet to be fully harnessed on an industrial level.

This dissertation contains an investigation of two such nascent nanotechnologies in the category of resistive switching (RS). The first, redox-based resistive random access memory (ReRAM), is capable of mimicking biological synapses by allowing storage of large numbers of interconnected and continuously adaptable resistance values. The second technology is based on Cr-doped V_2O_3, a correlated-electron material for which electronic control of Mott insulator-to-metal transitions potentially offers a fast and durable way to emulate certain dynamical behaviors of neurons. Here, we apply reimagined methods and analysis of electrical measurement to these synaptic and neuronal devices. The newly acquired data sheds further light on the nature of the resistance transitions and is used to design physically validated device models for embedding in large-scale neuromorphic simulations.

The measurement circuitry developed here addresses long-standing challenges in the external stabilization of device test structures, and allows (I, V) switching curves to be captured eight orders of magnitude faster than with commercially available equipment while causing significantly less electrical stress to the measured devices. Applying the measurement system, we introduce a new stochastic device model for solid-state synapses that is trained on a mass quantity of statistical measurement data of ReRAM. This model enables extremely fast ($> 10^8$ OPS) and accurate simulations of large synaptic arrays ($> 10^9$ cells) and provides a powerful new tool for statistical analysis of resistive switching data. Next, we identify an electro-thermal mechanism behind the negative differential resistance (NDR) and excitable dynamics observed in $(V_{1-x}Cr_x)_2O_3$ nanodevices. We show fast volatile switching (< 10 ns), high switching endurance ($> 10^{12}$ cycles), and favorable scaling characteristics in

this promising Mott insulating material. A coexisting non-volatile (NV) mechanism is investigated and the conditional occurrence of filamentation in the devices is linked to circuit instabilities, with wider implications for NV switching in other RS materials. The $(V_{1-x}Cr_x)_2O_3$ study culminates in a physical model that covers the scaling behavior and threshold adaptability, and is closely fit to observed oscillatory data.

Kurzfassung

In jüngster Zeit hat das Interesse an Materialien mit ungewöhnlichen elektronischen Eigenschaften wie starker Nichtlinearität, Hysterese und Speicherfähigkeit wieder zugenommen. Dieses Interesse ist zum Teil auf das Ende der Moore'schen Skalierung sowie auf die Entwicklung neuartiger Computerarchitekturen zurückzuführen. Derzeit wird die Rechenleistung durch den Speicherengpass begrenzt, da der physische Speicher nicht schnell oder groß genug ist, um die Pipeline der Zentraleinheit (CPU) zu speisen. Eine Alternative ist die Einführung einer neuen Speicherebene, die wesentlich schneller und skalierbarer sein muss als der vorhandene Flash-Speicher. Ein anderer Ansatz besteht darin, Systeme zu entwickeln, die die Vorteile des In-Memory-Computings nutzen, wie bei den vom Gehirn inspirierten Konzepten des neuromorphen Computings (NC). Um ihr volles Potenzial auszuschöpfen, ist jede dieser Strategien auf die Fähigkeit neuer Klassen von Speichertechnologien angewiesen, physikalische Mechanismen auszunutzen, die auf industrieller Ebene noch nicht vollständig genutzt werden.

In dieser Dissertation werden zwei solcher aufstrebenden Nanotechnologien in der Kategorie des resistiven Schaltens (RS) untersucht. Die erste Technologie, der redoxbasierte resistive Direktzugriffsspeicher (ReRAM), ist in der Lage, biologische Synapsen zu imitieren, indem er die Speicherung einer großen Anzahl vernetzter und kontinuierlich anpassbarer Widerstandswerte ermöglicht. Die zweite Technologie basiert auf Cr-dotiertem V_2O_3, einem Material mit korrelierten Elektronen, bei dem die elektronische Kontrolle von Mott-Isolator-zu-Metall-Übergängen eine schnelle und dauerhafte Möglichkeit zur Nachahmung des dynamischen Verhaltens von Neuronen bieten könnte. Hier wenden wir neuartige Methoden und Analysen elektrischer Messungen auf diese synaptischen und neuronalen Bauteile an. Die neu gewonnenen Daten werfen ein weiteres Licht auf die Natur der Widerstandsübergänge und werden verwendet, um physikalisch validierte Bauteilmodelle für die Einbettung in groß angelegte neuromorphe Simulationen zu entwerfen.

Die hier entwickelte Messschaltung löst die seit langem bestehenden Probleme bei der externen Stabilisierung von Teststrukturen und ermöglicht die Erfassung von (I, V)-Schaltkurven um acht Größenordnungen schneller als mit handelsüblichen Geräten, wobei die gemessenen Bauteile deutlich weniger elektrisch beansprucht werden. Mittels Anwendung des Messsystems stellen wir ein neues stochastisches Modell für Festkörpersynapsen vor, das auf einer großen Menge statistischer Messdaten von ReRAM trainiert wurde. Dieses Modell ermöglicht extrem schnelle ($> 10^8$ OPS) und genaue Simulationen von großen synaptischen Arrays ($> 10^9$ Zellen) und bietet ein leistungsfähiges neues Werkzeug für die statistische Analyse von resistiven Schaltdaten. Danach identifizieren wir einen elektrothermischen Mechanismus hinter dem negativen differentiellen Widerstand und der

neuronalen Dynamik, die in $(V_{1-x}Cr_x)_2O_3$-Nanobauteilen beobachtet werden. Wir zeigen schnelle flüchtige Schaltvorgänge (< 10 ns), hohe Schaltausdauer (> 10^{12} Zyklen) und günstige Skalierungseigenschaften in diesem vielversprechenden Mott-Isolatormaterial. Ein koexistierender nicht-flüchtiger Mechanismus wird untersucht und das bedingte Auftreten von Filamentierung in den Bauelementen wird mit Schaltungsinstabilitäten in Verbindung gebracht, was weitreichende Auswirkungen auf nicht-flüchtiges Schalten in anderen RS-Materialien hat. Aus der $(V_{1-x}Cr_x)_2O_3$-Studie resultiert ein physikalisches Modell, welches das Skalierungsverhalten und die Anpassungsfähigkeit des Schwellenwerts abdeckt und eng mit den beobachteten oszillierenden Daten übereinstimmt.

Acknowledgements

This dissertation was written during my doctoral research at the Institut für Werkstoffe der Elektrotechnik 2 (IWE2) at RWTH Aachen University, in collaboration with Western Digital Corporation (WDC). This involved relocating across the planet, an experience that expanded my mind in ways I never could have predicted. At this point, I would like to acknowledge the contributions of many people who helped make this work possible.

I am grateful to Prof. Rainer Waser for his persistence in recruiting me, for the opportunity to work in the IWE2, and for the support in the meantime. Also, thanks to Prof. Tobias Gemmeke for agreeing to co-examine the thesis. I thank Dirk J. Wouters for advising from IWE2, for reviewing the thesis, and for many helpful discussions over the years.

I am immeasurably thankful to Daniel Bedau, advisor of the projects from the WDC side. Ten years ago, Daniel planted the seed of my interest in doing a PhD in a foreign country, and later made me aware of IWE2. During my time here he maintained the collaboration, however painstakingly, and continued to support and discuss the work in detail on a weekly basis. The industrial resources he contributed really elevated what was possible to do here.

Thanks Martina Heins for always being pleasantly willing to help and for being dynamic in unusual circumstances, and thanks Ulrich Böttger for extending my contract an absurd number of times.

I thank Jonathan A. J. Rupp and Johannes Mohr, colleagues who grew and optimized the VCrOx thin films and participated in the investigations.

Thanks to Erik Wichmann, professional Hiwi and the only one within 30 km^2 who can operate a soldering iron, for working on the electronics, the PCB designs, and diagrams in the unholy TikZ language. Thanks to Eugen Gebert for helping build the temperature stage. Thanks to Alejandro García for helpful contributions to our bespoke instrumentation library. Thanks to Markus Beschow for adapting the COMSOL model.

I am ridiculously grateful to Anne Siemon, Marcel Schie, Camilla La Torre, Moritz von Witzleben, and Carsten Funck, without whom I never would have survived the logistics of living in Germany.

Thanks, Carlos Rosário, for being the only one who ever read any of my papers.

Thanks to our external collaborators, Peter Warnicke, Etienne Janod, Danylo Babich, Julien Tranchant, and Hongchu Du, for discussions and characterizations of the VCrOx films. Thank you Jeffrey Lille, Oleksandr Mosendz, Jean-François Nodin, and Gabriel Molas for fabrication of the ReRAM devices. Also thank you to the rest of our collaborators at WDC, Alexander Elias, Jim Rainer, Hans Richter, Derek Stewart, Joyeeta Nag, Wen Ma, and Michael Grobis.

Thanks to my very important teachers and mentors, Louis Cliff Grosberg and Prof. E. Dan Dahlberg, and thanks Prof. Eric Fullerton for investing in my early PhD career.

Thanks to Thomas Pössinger for the great illustrations. Thanks Udo Evertz for keeping the network up and giving me admin rights. Thanks Erhard Halfmann for roasting the coffee that fueled my writing.

Finally, a big thank you to my mom and dad, who I hold fully responsible for my existence. I hope this giant unintelligible book makes you proud.

Contents

List of Abbreviations

RS	**R**esistive **S**witching
ReRAM	**Re**sistive **R**andom **A**ccess **M**emory
CeRAM	**C**orrelated **e**lectron **R**andom **A**ccess **M**emory
NV	**N**on-**V**olatile
TS	**T**hreshold **S**witching
IMT	**I**nsulator to **M**etal **T**ransition
NDR	**N**egative **D**ifferential **R**esistance
PCB	**P**rinted **C**ircuit **B**oard
CLA	**C**urrent **L**imiting **A**mplifier
DUT	**D**evice **U**nder **T**est
CtC	**C**ycle-**t**o-**C**ycle
DtD	**D**evice-**t**o-**D**evice
ML	**M**achine **L**earning
HRS	**H**igh **R**esistance **S**tate
LRS	**L**ow **R**esistance **S**tate
IRS	**I**ntermediate **R**esistance **S**tate
CMOS	**C**omplementary **M**etal-**O**xide-**S**emiconductor
BJT	**B**ipolar **J**unction **T**ransistor
FET	**F**ield **E**ffect **T**ransistor
SPA	**S**emiconductor **P**arameter **A**nalyzer
SMU	**S**ource **M**easure **U**nit
CC	**C**ompliance **C**urrent
VCM	**V**alence **C**hange **M**echanism
MIM	**M**etal-**I**nsulator-**M**etal
TMO	**T**ransition **M**etal **O**xide
DOS	**D**ensity **O**f **S**tates
TEM	**T**ransmission **E**lectron **M**icroscopy
SEM	**S**canning **E**lectron **M**icroscopy
XAS	**X**-ray **A**bsorption **S**pectroscopy
SNR	**S**ignal to **N**oise **R**atio
CPU	**C**entral **P**rocessing **U**nit
GPU	**G**raphics **P**rocessing **U**nit
ODE	**O**rdinary **D**ifferential **E**quation
PDE	**P**artial **D**ifferential **E**quation
DAE	**D**ifferential-**A**lgebraic **E**quations

List of Symbols

Symbol	Name	Unit
I	Current	A
I_d	Device current	A
V	Voltage	V
V_a	Applied voltage	V
V_d	Device voltage	V
V_S	SET voltage	V
V_R	RESET voltage	V
P	Power	W
t_ox	Oxide/device thickness	m
w_ox	Oxide/device width	m
a	Conduction parameter	S
b	Conduction parameter	K
c	Conduction parameter	$\sqrt{\mathrm{V}}$
α	Conduction parameter	$\mathrm{S\,m^{-1}}$
γ	Conduction parameter	$\sqrt{\mathrm{m/V}}$
T	Device temperature	K
T_0	Ambient temperature	K
R_s	Series resistance	Ω
R_d	Device resistance	Ω
R_H	Resistance of a high resistance state	Ω
R_L	Resistance of a low resistance state	Ω
ρ	Electrical resistivity	$\Omega\,\mathrm{m}$
σ	Electrical conductivity	$\mathrm{S\,m^{-1}}$
κ	Thermal conductivity	$\mathrm{W\,m^{-1}\,K^{-1}}$
C_p	Parallel capacitance	F
R_th	Thermal resistance	$\mathrm{K\,W^{-1}}$
C_th	Thermal capacitance	$\mathrm{J\,K^{-1}}$
t	Time	s
k_B	Boltzmann constant	$\mathrm{J\,K^{-1}}$
μ_0	Vacuum permeability	$\mathrm{H\,m^{-1}}$

Introduction

For over 75 years, the notion of a "computer" has been largely synonymous with the general purpose digital architecture attributed to the Hungarian-American polymath, John von Neumann [1]. Miniaturization, integration, and optimization of semiconductor-based electronic devices has since propelled the advancement of computer hardware in this classic architecture to the incredible level of today. However, fundamental inefficiencies of the von Neumann paradigm are becoming relevant for modern workloads, as a larger share of the available energy is spent shuttling data back and forth between storage and locations where it can be processed [2, 3]. Despite the overwhelming success of the technological strategy of recent decades, physical and economic constraints are making the same approach increasingly challenging to follow into the future [4].

Along with the rise of machine learning (ML), modern trends in computing have likewise emphasized neuromorphic architectures that implement brain-inspired algorithms directly on the hardware level [5–7]. These approaches, through in-memory computation and massive parallelism, excel in new classes of computational problems and offer promising advantages with respect to power consumption and error resiliency. While complementary metal-oxide-semiconductor (CMOS)-based neuromorphic computing (NC) implementations have made substantial progress recently [8, 9], new materials and physical mechanisms may ultimately provide better opportunities for energy efficiency and scaling [7, 10–15].

A functionality sought after in NC applications is an ability to mimic the biological building blocks of the brain; neurons and the synaptic connections between them. To cover different electronic aspects of this behavior, many schemes make use of resistive switching (RS), which refers to a broad class of related phenomena wherein the resistance of two-terminal devices can be controlled via electrical stimuli [16]. Several solid-state memory technologies in this category, such as phase-change memory (PCM), magnetic random access memory (MRAM), resistive RAM (ReRAM), and correlated electron RAM (CeRAM), are emerging as candidate components, each exploiting different material properties [11, 17–19].

Among the device candidates, ReRAM is attractive for its simplicity of materials and device structure, providing the necessary CMOS compatibility and scalability [20]. ReRAMs, also controversially called memristors [21–31], are essentially two-terminal nanoscale electrochemical cells, whose variable resistance state is based on manipulation of the point-defect configuration in an

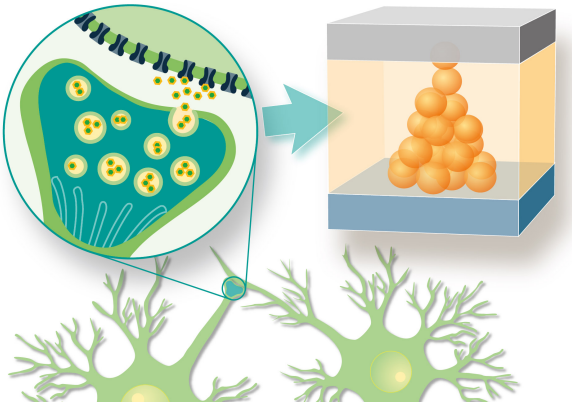

Figure 1: In loose analogy to biological synapses, two-terminal solid-state nanodevices such as ReRAM can store the strength of network connections as electrical resistance states. The devices, consisting simply of patterned metal-insulator-metal material stacks, have an adjustable resistance level determined by the ionic configuration inside the insulating layer.

oxide material (depicted in Fig. 1). This redox-based switching mechanism is intrinsically analog, allowing a range of stable resistance levels to be stored and adjusted through application of bipolar voltage stimuli. However, various non-idealities such as stochasticity, nonlinearity, and noise are prominent features of these devices that, for better or for worse, critically impact the performance of systems composed of them [32, 33].

Another promising possibility currently being explored is to use electronically controlled insulator-to-metal transitions (IMTs) arising due to electron-electron correlation effects [34]. Cr-doped V_2O_3 is a well-known example of a Mott insulating material supporting a particular kind of IMT called a Mott transition, which could prove to be a robust basis for future volatile or non-volatile memories [35]. At the same time, similar volatile switching phenomena commonly occur in nanodevices due to electro-thermal feedback effects, which involve either simple thermal activation, temperature-induced IMTs, or both [36, 37]. These types of devices continue to attract attention due to their interesting dynamical behavior in circuits [38–40]. The effects, which include spiking and oscillations, are viewed increasingly in the context of neural networks [41, 42].

This dissertation addresses remaining challenges in electrical characterization and modeling of synaptic and neuronal devices, as required by next generation memory and computing architectures. It is divided into four chapters. We begin in Chapter 1 by providing background information on selected volatile and non-volatile RS phenomena while introducing important concepts in electrical measurements that pertain generally to the rest of the work.

Emerging devices are based on only partially understood mechanisms, and may exhibit strong non-linearity, negative differential resistance (NDR), oscillations, stochasticity, and memory effects. In assessing the electrical capabilities of RS devices, it is important to consider not only the device material properties but also the effects of feedback, instability, excess electrical stress, and the general role of the driving circuitry on measurement data. Chapter 2 describes the design of two new measurement circuits specialized in characterization of RS. We demonstrate extraction of important statistical information from mass collection of (I, V) cycles at eight orders of magnitude faster rates than possible using commercially available equipment. This is done while also capturing important transient events during switching, which usually escape detection. The developed circuits are then employed for electrical measurements in the rest of the work.

Modern ML models have reached an astonishingly large and ever-increasing size, with recent examples exceeding a hundred billion weights [43]. Before comparable hardware implementations based on solid-state memories can become a reality, large-scale network designs need to first be evaluated by computer simulations. Training, validation, and optimization of such networks involves a huge number of simulated devices, voltage pulses, and current readouts. Within this process, it is important to accurately consider the constraints of the underlying hardware in detail. Therefore, lightweight, fast, and accurate stochastic simulations of the individual synaptic devices are a key requirement. In Chapter 3, we present a new data-driven approach for modeling large collections of stochastic synapses. We created a hierarchical statistical model based on vector autoregression and probability density transformation, which replicates the measured stochastic behavior of memory cells very closely. We provide high-performance, parallelized implementations for both CPUs and GPUs and demonstrate simulation of over one billion cells simultaneously with throughputs exceeding one hundred million weight updates per second.

In Chapter 4, nano-devices based on the Mott-insulating material $(V_{1-x}Cr_x)_2O_3$ are investigated. The electronic structure of the material is probed with synchrotron X-ray measurements, and the electrical operation of corresponding nano-patterned devices is extensively characterized. Among the results are a volatile NDR effect with below ten nanosecond switching speeds, endurance over a trillion cycles, and oscillatory modes in the tens and hundreds of megahertz range [44, 45]. The device operation clearly involves an electro-thermal mechanism as a volatile precursor to a separate non-volatile switching mechanism, which is likely of nano-ionic origin. New insights are inferred from careful electrical measurements of the various states that can be set up in the devices by using current-limiting feedback. These insights address the question of how the conductivity inside the switching material is spatially redistributed as a result of NV switching operations, and specifically how the degree of filamentation is connected to the material properties and to the measurement conditions. Finally, physics-based differential equation models are developed to fit the measured dependence

of time, temperature, and length scale on device operation, which enable simulations of neuromorphic systems based on these devices [46].

Chapter 1

Fundamentals

This chapter introduces the different memory technologies studied in this dissertation, mainly in the context of electrical measurement; the principal experimental tool used in the work. Section 1.1 first covers ReRAM based on the electro-migration of oxygen vacancies, and Section 1.2 discusses the correlated electron material $(V_{1-x}Cr_x)_2O_3$. Section 1.3 is concerned with the challenges in electrical measurement of these devices and how they relate to current practices. Finally, in Section 1.4 we take a step back to look at the ubiquitous effect of device Joule heating, providing illustrative examples useful for understanding simple measurement situations where coupling with the impedance of an external measurement circuit has unexpected consequences.

1.1 Redox-based resistive switching mechanisms

All insulating layers suffer a loss of their insulating properties if a sufficient voltage is forced across them for a sufficient amount of time. The mechanisms of this so-called dielectric breakdown effect are various [47], but commonly involve an instability where the driving force of the physical processes responsible for the increasing material conductivity is subject to positive feedback. This runaway process can rapidly lead to structural or chemical changes in the insulator, eventually forming a conductive pathway through the material which may persist long after the electrical stimulus is removed.

Left unimpeded, the destructive effects of breakdown are considered an irreversible phenomenon, leaving no possibility to recover an insulating state. This breakdown effect has been thoroughly investigated in the field of oxide reliability, where its role in the failure of thin gate oxides make it an effect to be avoided. However, the field of resistive switching (RS) takes a different perspective on the same class of effects. The basic objective of RS is to interrupt the process before it causes permanent and irreversible damage to the material, so that the device resistance state can be repeatedly and reliably modulated by applying voltage signals with alternating polarity, duration, or amplitude. Provided that such devices can be (re)programmed to stable resistance levels, they can then be employed as binary or multi-level non-volatile memories, or as weights in an artificial neural network.

Non-volatile RS effects occur in a wilderness of different materials, and essentially every thin insulating layer between two electrodes can show persistant and reversible resistance transitions. This amusingly includes beer, whiskey, bananas, honey, leaves, and milk [48–52]. Broadly speaking, a class of RS can be distinguished by the involvement of mobile defects in the material that locally increase its conductivity. Different spatial configurations of the defects with different total resistance levels can be set up through migration of the defects under the applied field as well as temperature and/or concentration gradients. A detailed classification of these effects has been made according to the insulating and electrode materials used and to the type of defects responsible for the switching effect [16], though the switching mechanisms and modes are not all mutually exclusive and can occur under different types of electrical control and in statistically overlapping regimes [53–56].

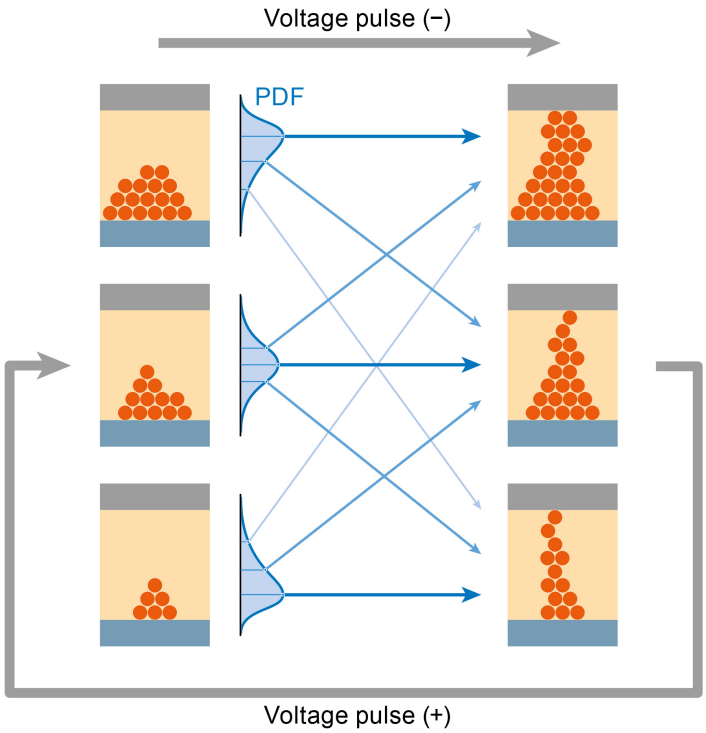

Figure 1.1: Resistance states in VCM-based ReRAM devices depend on the configuration of oxygen vacancy defects (drawn as orange circles) in the oxide layer. Application of voltage pulses can produce new states, which exhibit a probabilistic dependence on past states, leading to long-range correlations. Starting with effectively infinite state possibilities (represented by the three cell states on the left) an applied voltage pulse brings about a set of transition probabilities to many possible subsequent states (right).

For the specific case of the valence change mechanism (VCM) that is employed in the ReRAM investigated in this work, the non-volatile effect is mediated by oxygen vacancy defects in a binary transition metal oxide (TMO)

resistive layer, which act as mobile electron donors to increase the local material conductivity [20]. The oxygen vacancies, either already present in the deposited oxide (e.g. TaOx or HfOx) or generated through redox reactions at the oxide/electrode interface, then drift under the applied electric field, with their mobility activated by Joule heating. Under a constant applied voltage, a rising concentration of oxygen vacancies locally increases the driving force for their production and/or migration (namely, the electric field and temperature), causing the defects to assemble and form one or more filamentary conductive paths through the film [57–60].

The basic requirement for an electronic device serving as an artificial synapse is to moderate the flow of electrical signals through connections in a network. Left undisturbed, the device ideally maintains a fixed weight, or dependence between the voltage across the two device terminals, V, and the resulting current through the device. Further, for learning there must be some means of affecting the weight in a durable way. Bipolar ReRAMs have an adjustable (potentially nonlinear) and non-volatile resistance state, which is based on the size and shape of a conducting filament that partially or fully bridges the insulating gap of the oxide material. Simplistically, when V exceeds certain threshold levels, the resistance state begins to transition toward lower or higher values depending on the voltage polarity, which corresponds to growth and shrinkage of the conducting filament. When the filament only partially bridges the insulating gap, conduction may be limited for example by tunneling through a Schottky barrier of a material interface, leading to a relatively high resistance levels [20, 61]. As the filament grows and gradually bridges the gap, the resistance decreases as conduction transitions towards the ohmic type.

There are a number of technologically promising features of VCM-based ReRAMs that make them attractive for next generation memory and computing architectures [12]. The simple device structure can be fabricated using a wide variety of CMOS-compatible materials [62, 63]. ReRAMs also have good write endurance [64], support high read/write speeds [65–68], and have strong potential for 3D integration [69]. Importantly, the effect can be scaled to nanometer dimensions as only a small number of ions need to participate in the switching process in principle [70–72].

However, several drawbacks currently stand in the way of widespread application of these devices. These stem from the fact that filamentary states in the material are electrically delicate, prone to instability, and the transitions between the states are strongly stochastic (see Fig. 1.1). Therefore, a central challenge for ReRAM devices is dealing with the intrinsically random nature of their switching processes, which leads to large variability in the programmed resistance states and switching parameters [73, 74]. Achieving an acceptable level of control of ReRAM devices will require an in-depth understanding of the statistical processes at play, as well as an optimization of the active materials together with the control circuitry.

1.2 Correlated electrons and the vanadium oxide system

Mott insulators, named after the 1977 Nobel leureate Sir Nevill Francis Mott, are a class of materials with unusual electronic properties arising due to strong electron-electron correlation effects. Simplistic models used to calculate the electronic band structure of these materials neglect the electron-electron interactions and predict a conductive ground state. However, an insulating state in fact arises because, essentially, Coloumb repulsion suppresses the transfer of electrons from one atomic site to another. This phenomenon is treated approximately in the single band Hubbard model, where a Hamiltonian describes a competition between the forces of kinetic energy and the on-site repulsive potential [75]. In Mott insulators, the electron correlation effect causes a splitting of the electron bands as shown qualitatively in Fig. 1.2 [76].

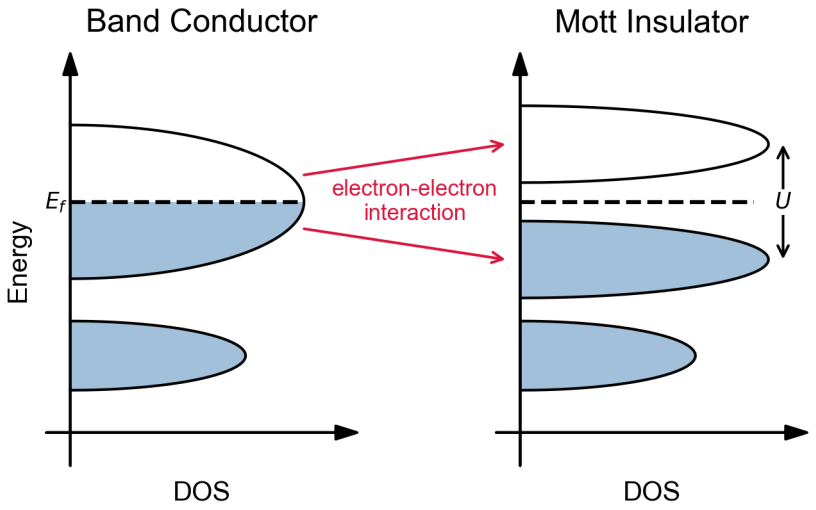

Figure 1.2: In the Hubbard model, half-filled conduction bands are split into upper and lower Hubbard bands due to electron-electron interactions. Depending on the balance between the bandwidth and the Hubbard energy, U, the material may be in a Mott insulating state.

Mott materials are good insulators, yet they are precariously close to being conductors. Insulator to metal transitions (IMTs) can be induced by various physical stimuli such as temperature, pressure, or doping [77]. It was recently shown that Mott insulators subjected to electric pulses can undergo an IMT linked with the creation of a conducting filamentary path [78]. This raises fundamental research questions about whether these transitions can be harnessed to create new scalable device technologies: are they suitable materials for memory, selectors, or neuromorphic elements [79]?

The binary vanadium-oxide system has been intensely studied since the early beginnings of condensed matter physics. It known to exist in over 20 different stable phases, many with IMT behavior due to correlated electron effects [80, 81]. The sesquioxide system (V_2O_3) is a famous example of a material

with a canonical Mott IMT that can be driven by pressure or doping, as shown in the phase diagram of Fig. 1.3. Under ambient conditions, V_2O_3 is a correlated metal, but substitution of a few percent of the V sites with Cr biases the material in a Mott insulating state. IMTs have been observed in $(V_{1-x}Cr_x)_2O_3$ thin films, where they are known to depend very sensitively on the stoichiometry, phase, and strain conditions [82–86]. Application of short electric pulses to the material can also induce an isostructural IMT, which is a possible basis for volatile or non-volatile memory applications [34, 35, 87]. Due to the advanced lithography nodes, very thin functional layers below 10 nm are a key requirement for all of these envisioned applications. At such small thicknesses, stoichiometry and phase purity become increasingly difficult to characterize, and the interface of the oxide with the substrate plays an important role which can serve either to facilitate or hinder the Mott transition behavior.

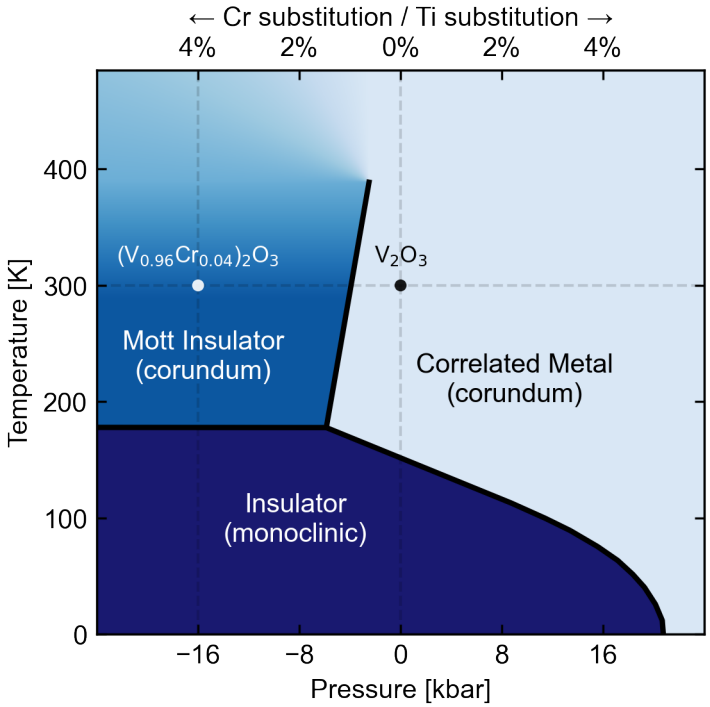

Figure 1.3: Phase diagram of $(V_{1-x}Cr_x)_2O_3$ vs. concentration of the Cr dopant (x), pressure and temperature. Doping brings the correlated metal into an insulating state, which can theoretically be switched back to the metallic state using an applied voltage. Pressure and doping are shown as having equivalent effects on the phase, though this has been refuted [88]. Data was extracted from [89].

In Chapter 4, the electronic structure of $(V_{1-x}Cr_x)_2O_3$ films as thin as 5 nm is probed by X-ray absorption spectroscopy and switching behavior in corresponding high-quality nanodevices are scrutinized through different kinds of electrical measurements.

1.3 Electrical measurements of resistive switching

An electrical measurement records the response (i.e. the current and/or voltage at specific nodes/branches) of a particular driving circuit when a test device is connected at its measurement terminals. In the case of RS, the stochastic two-terminal devices are elements of the circuit that simultaneously measures and controls them. Resistance transitions in RS materials can occur on timescales below 1 ns [66–68], but the nanoscale material volumes involved cannot normally survive prolonged exposure to the voltage required to initiate the transition, as the current density quickly reaches levels that can cause irreversible damage [90]. To prevent this and to keep the process in a reversible regime, a strategy is necessary to limit the duration and amplitude of voltage applied to the active part of the device during switching.

A challenging aspect of electrical characterization of RS devices is that measurements depend not only on the materials and structures used in device fabrication, but are also sensitively dependent, often in unexpected ways, on the measurement circuitry used. Generally, this is a consequence of coupling between the ultra-fast physical dynamics of the device under test (DUT) and the frequency-dependent feedback of the external driving circuit. Together, these coupled dynamics determine the current and voltage trajectories experienced by the DUT through time during its resistance transitions. Parasitics in the system have a large influence because they can store and release energy comparable to, or in excess of, what is needed to induce switching in the nano-scale material volumes. Thus, electrical measurements often do not probe device characteristics directly, and easily overlooked details of the test system, such as the model number of an instrument or the stray capacitance introduced by a short length of coaxial cable, can have an important impact on switching data [91–94].

Another fundamental challenge in using electrical measurement as an experimental technique relates to a concept known in control theory as observability [95]. Device stochasticity and path-dependence originate from nanoionic processes which occupy an extremely high-dimensional physical state space, but these internal states are difficult to infer using the few output signals available for electrical measurement. In other words, currents and voltages are a very indirect probe of what is really happening inside of the device, and the resulting data is therefore subject to many conflicting interpretations. On top of this, electrical instrumentation is technically complicated, and there are many sub-optimal ways to measure that still produce (I, V) data in a similar format. This data tends to be detached from critically relevant measurement details, which further impairs the chance for a productive interpretation.

Notwithstanding a diversity of experimental possibilities to address this seemingly bleak situation, reported RS electrical measurements overwhelmingly fall into the two categories represented in Fig. 1.4. To a large extent, these techniques were inherited from existing standards, methods, and equipment used in optimization of semiconductor devices, without specific adaptation to the unique challenges faced by RS devices [96–99].

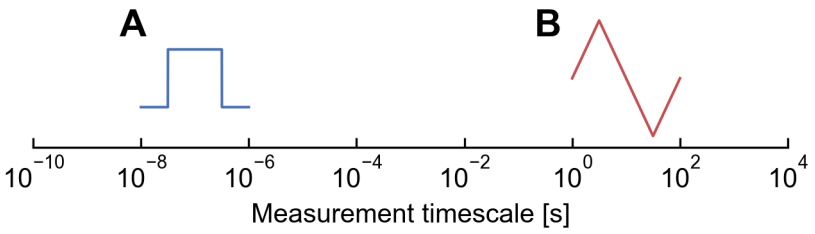

Figure 1.4: Conventional electrical measurements of RS devices largely fall into two rigid categories. Voltage driving signals are applied directly to the DUT and are either **(A)** square pulse waveforms on short timescales or **(B)** (quasistatic) sweeps at very low frequency.

Quasistatic (I, V) loops are a measurement where resistance switching is induced by an applied voltage that is continuously ramped at low speed ($\sim$1 V/s) between positive and negative values. The device current at each point of the sweep is sampled, and can be plotted against the applied voltage as shown schematically in Fig. 1.5. Such (I, V) loops are rich in information, and important features such as the (I, V) non-linearity, voltage and current switching thresholds, and details of the transition behavior can be extracted from each cycle. With very few exceptions [100, 101], sweeping (I, V) measurements are performed using commercial source measure units (SMUs), which can be contained inside instruments called semiconductor parameter analyzers (SPAs). These instruments, while offering extremely high resolution and dynamic range, rely on signal averaging on the scale of the power line frequency (50 Hz or 60 Hz) and are therefore limited to very slow measurement sweep and sample rates. In addition to making experiments involving more than a few hundred switching cycles impractical, this long timescale is 6 to 9 orders of magnitude larger than those relevant for applications, which puts excessive electrical stress on these highly time-dependent devices [102].

Sweeping (I, V) loop measurements are only generally possible with the use of a feedback mechanism to prevent runaway destruction of the RS device. However, externally implemented current limiting such as the current compliance (CC) function of commercial SMUs and SPAs are known to permit large current overshoots during resistance transitions in the time before CC is eventually enforced (shown in the test measurements of Fig. 1.6). These overshoots can lead to catastrophic damage to cells [103, 104] (see Fig. 1.7) and can otherwise strongly influence the measurements [105, 106]. At the same time, overshooting transients can not be recorded by the instrument itself because of its low bandwidth and sampling rate, and resistance cycling may procede while leaving no direct evidence of poor control in the experimental data.

Overshoots are only one manifestation of the deficiency of SMUs when applied to switching devices, as the instrument designs do not anticipate fast changes in a two-terminal conductance which is driven by the measurement voltage itself. Even using the appropriate (voltage/current) sourcing mode, the instrument can return spurious data as it induces oscillations in NDR

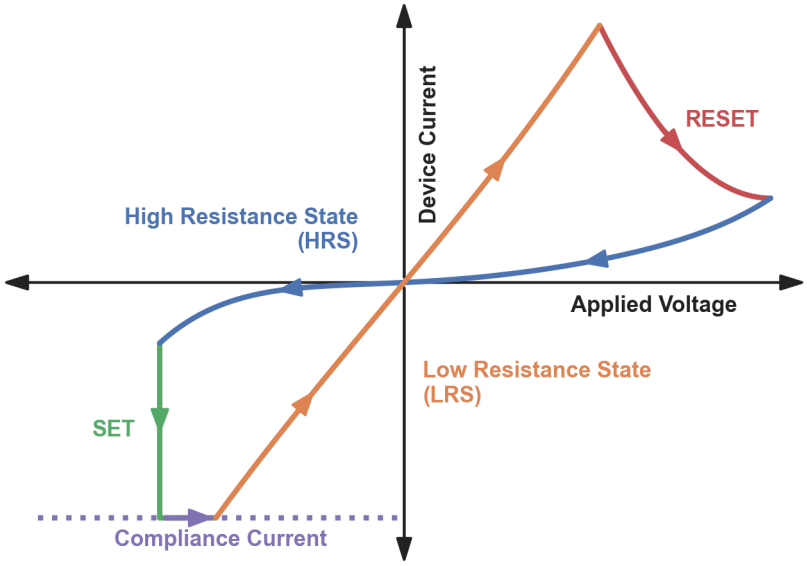

Figure 1.5: Schematic diagram of a single bipolar ReRAM (I, V) loop cycle. A voltage applied to the active electrode is swept in order to induce resistance switching while simultaneously measuring the device current. A feedback system interrupts runaway breakdown during the SET transition by attempting to enforce a current limitation (compliance). Important switching parameters such as threshold voltages and nonlinear resistance states can be extracted from (I, V) loop measurements, which show significant statistical variation both between devices and between cycles.

devices that it does not have the measurement bandwidth to resolve, as demonstrated in Fig. 1.8. This stability issue can arise when cells are driven by an ideal source with too high output capacitance, following the general mechanism to be described in Section 1.4.3, and also pointed out recently by Brown et al. [99]. However, the actual situation is far more fraught. To measure RS with an SMU is to investigate incompletely understood switching mechanisms by driving the cells with a slow, proprietary feedback circuit with an unpredictable dynamical response; conditions quite unlike any that would be produced in any integrated system of interest. Even in the subset of cases where instability is thought not to play an important role, the sheer uncertainty introduced by the misapplication of these machines makes them counter-productive to use for switching experiments, and doing so can cause long lasting damage to the understanding of RS effects.

Patterning RS devices directly on MOS transistors can provide superior current limiting and stability, but the required integration limits the materials available and necessitates long fabrication cycle times [105, 107, 108]. A simpler approach from the point of view of fabrication is to integrate fixed resistors on the chip in series with the devices [90, 100, 109, 110, 260]. However, the large linear feedback introduced by this relatively inflexible method significantly affects the switching behavior [111, 112], and can push the operating

voltage outside of a practical range. On the other hand, attempts to control RS using external current-limiting components such as resistors [90, 103–105, 110, 113–121] or transistors [91, 92, 122–125] struggle with issues related to the parasitic capacitance, among other shortcomings. In Chapter 2, we address these measurement challenges with two new experimental setups, which use external circuitry to measure (I, V) traces at high frequency using arbitrary waveforms while minimizing unwanted overshoots.

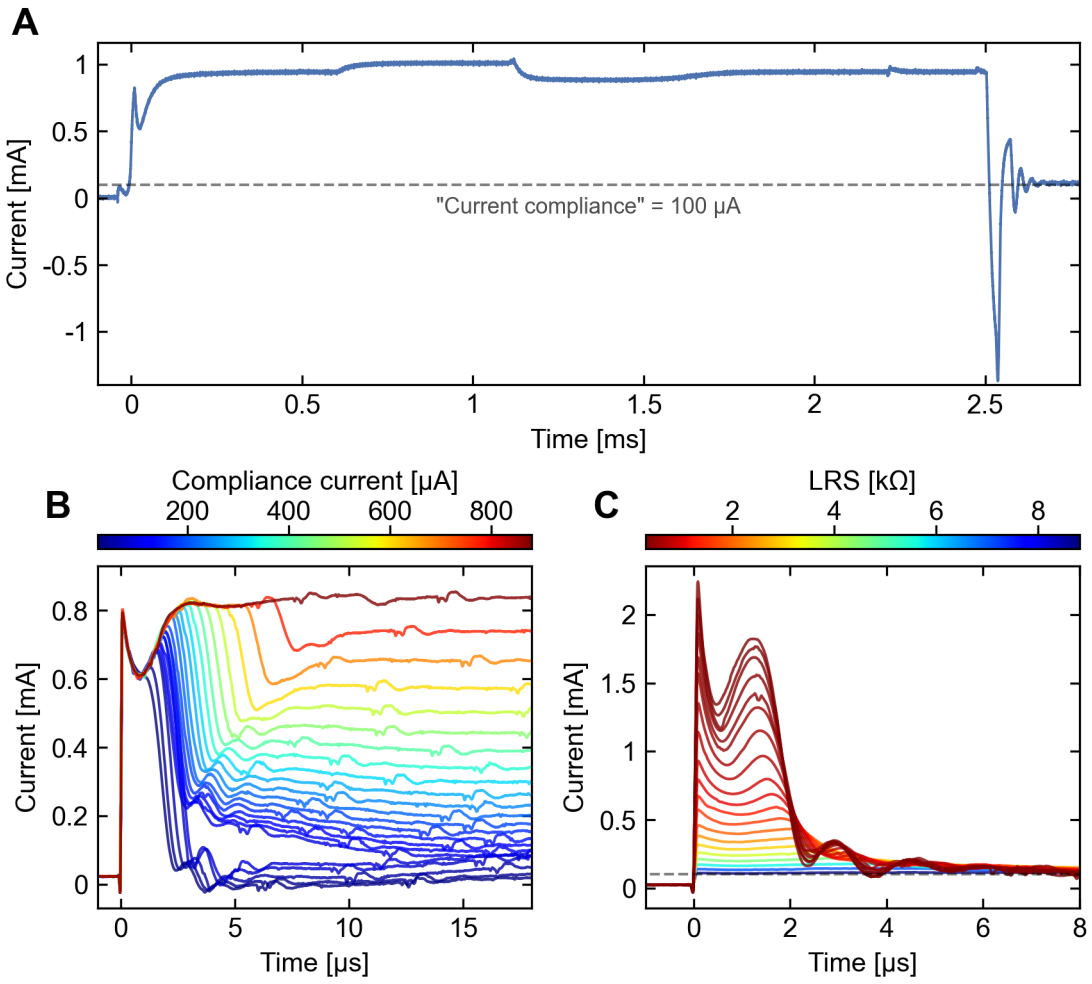

Figure 1.6: Current overshoot transients delivered to surface mount resistors by Keithley SMUs, set to source 1 V, as measured by an oscilloscope. At time $t = 0$ the DUT resistance was suddenly switched from an initially high resistance (HRS) to a lower resistance (LRS) via a low capacitance relay. **(A)** Model 2410 with CC level set to $I_{cc} = 100$ µA while switching from 1 MΩ to 1 kΩ. The current limiting feedback is unresponsive for 2.5 ms after switching; a very long duration relative to ReRAM switching times. The test resistor is also subjected to a large oscillatory current that includes reverse polarities. Model 2636B SMU shows an improved but still inadequate response when switching from an HRS of 43 kΩ, **(B)** with the LRS fixed to 1 kΩ and varying I_{cc}, and **(C)** with I_{cc} fixed to 100 µA while varying the LRS. These overshoots are invisible in the data returned by the SMUs, because their duration is far lower than the averaging period for a single measurement sample. No matter the CC setting, the transients are identical for at least 1 µs because the initial spikes are delivered directly by the coaxial cable.

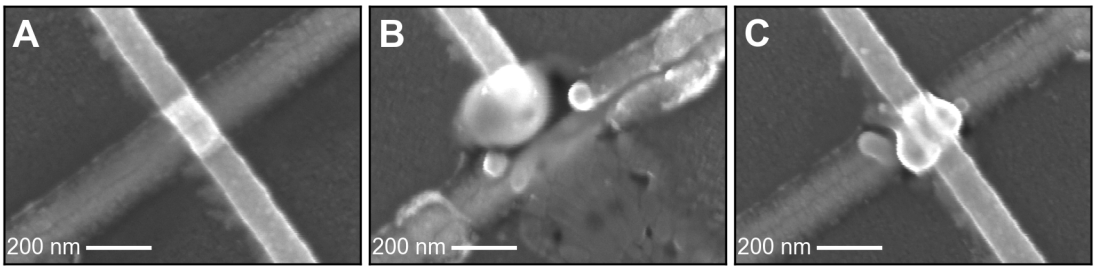

Figure 1.7: Damage induced by current overshoots in 100 nm crossbar structures with 30 nm Pt top and bottom electrodes and a 100 nm VOx switching layer. The initial structure before measurement is shown in **(A)**. Device **(B)** was subjected to a voltage sweep by an SMU with a 15 μA current compliance setting. Device **(C)**, protected by a 10 kΩ external series resistor, was swept similarly, but was still damaged by the capacitive discharge of an interconnecting 20 cm coaxial cable. Although the cells are visibly destroyed, both nevertheless continued to show measurable RS behavior as a conducting path still existed through what remained of the oxide material.

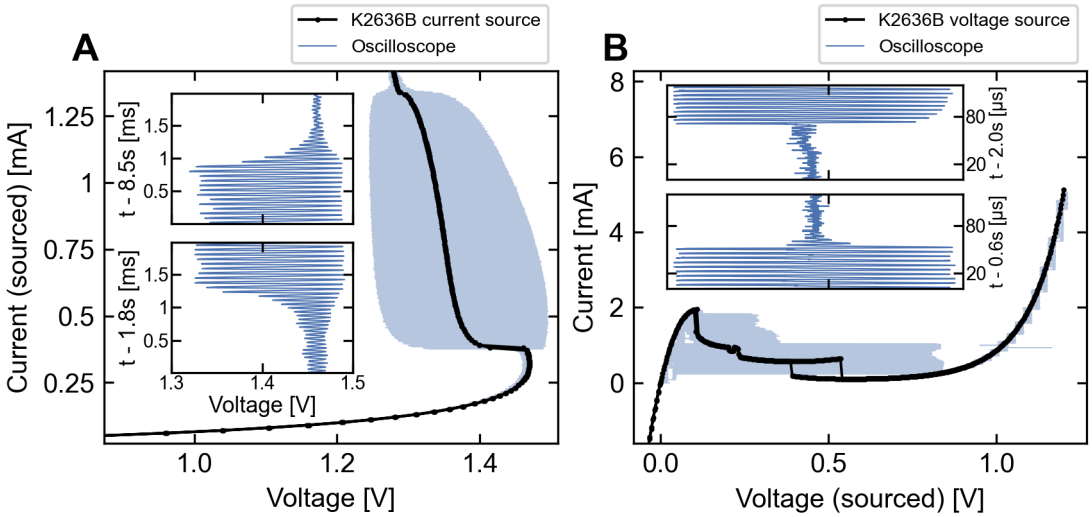

Figure 1.8: Source measure units (SMUs) cannot reliably measure NDR of either S- or N-type, as demonstrated in **(A)** an S-type $(V_{0.85}Cr_{0.15})_2O_3$ nanodevice ($500\times500\times30$ nm^3) and **(B)** an N-type 3I306E Ga-As tunnel diode. In both cases, oscillations are induced and the averaged curves returned by the SMU are inaccurate and potentially misleading.

1.4 The electro-thermal switching mechanism

When electrical current flows through a resistive medium, Joule heating always results; this is a basic consequence of work done on the charge distribution by the Lorentz force. At the same time, it is also common for a material's conductivity to be strongly affected by heating. For semiconductors, the conductivity usually increases through thermal activation of charge carriers over a band gap or over trap barriers [126, 127]. Surprisingly, these two properties alone can lead to highly unintuitive behavior in circuits built with electro-thermal components [128, 129]. Even in relatively simple circuits, the dynamics are coupled such that under different conditions they can exhibit bifurcations, oscillation, and chaos [99, 130, 131].

Due to its general ubiquity in nanodevices, it is important to first consider the effect of thermal feedback in isolation. Therefore, this section gives a conceptual framework for understanding electrical measurements of two-terminal devices that have a large temperature dependence of their conductance. Using suitable simplifications, a relatively tractable analysis serves as a good prototype for surveying the variety of dynamics that can occur. This straightforward example also provides an essential base layer for understanding the behavior of more complex devices, for which such thermal effects are often strongly involved [132, 133]. In particular, this section provides the context in which the various results of Chapter 4 can be best interpreted.

1.4.1 Thermistor dynamics

To introduce a deliberately minimal example, suppose two metallic electrodes make contact with a uniform cuboid of semiconducting material on opposing faces. Assume uniform electric field lines perpendicular to the electrode planes and that the temperature T inside the cell volume is always uniformly distributed. Suppose the medium has thermally activated charge carriers such that the total device conductance follows a voltage-independent Arrhenius equation,

$$G = \frac{I}{V} = A \exp\left(\frac{-E_\mathrm{b}}{k_\mathrm{B} T}\right), \tag{1.1}$$

where I and V are the device current and voltage drop, A is a pre-exponential current constant, E_b is the activation energy barrier, k_B is the Boltzmann constant.

The device heating rate can be calculated as a balance between the power generated by Joule heating and that lost to the cell surroundings by conductive heat transfer. Assuming a constant lumped heat capacity C_th of the cell, we have

$$C_\mathrm{th} \frac{dT}{dt} = P_\mathrm{in} - P_\mathrm{out}. \tag{1.2}$$

The Joule heating term is a product of the instantaneous device current and voltage difference across its electrodes,

$$P_{\text{in}} = IV, \tag{1.3}$$

and Newton's law of cooling can be applied to approximate the rate of heat conduction out of the cell boundaries as

$$P_{\text{out}} = \frac{T - T_0}{R_{\text{th}}}. \tag{1.4}$$

Here, the cooling rate is simply proportional to the difference between the cell temperature and the ambient level T_0. The thermal resistance R_{th} quantifies the thermal contact with the cell environment and depends on the cell size as well as the surrounding materials.

A convenient property of the conduction Eq. 1.1 is that any possible non-zero (I, V) condition unambiguously identifies the temperature state of the cell. This allows equations 1.1–1.4 to be combined in order to express the temperature dynamics as a single ordinary differential equation (ODE) in terms of only the electrically observable variables I and V,

$$\frac{dT}{dt} = \frac{1}{C_{\text{th}}}\left[IV + \frac{1}{R_{\text{th}}}\left(\frac{E_{\text{b}}}{k_{\text{B}} \log\left(\frac{I}{AV}\right)} + T_0\right)\right]. \tag{1.5}$$

A visualization of this unwieldy equation (Fig. 1.9) is a useful step toward a qualitative understanding of its dynamical behavior. For the sake of concreteness, values for a hypothetical nano-scaled cell are assigned to the model parameters, following the reasoning in Appendix A.

The first task in analyzing any dynamical system is to identify the conditions under which the state variables do not change in time. In the case of this first-order system, parameter and variable values consistent with the condition $dT/dt = 0\,\text{K/s}$ are called the fixed points of the model. Referring to Fig. 1.9, we see that for a given set of parameters, a projection of dT/dt on the (I, V)-plane shows a smooth continuous curve of fixed points that separate two regions of net heating ($dT/dt > 0$ K/s) to the right and cooling ($dT/dt < 0$ K/s) to the left. This curve can be understood as the set of (I, V) points for which the input power equals the power lost to the cell surroundings, such that its temperature remains stationary. As the distance of an (I, V) point from the curve increases, the disparity between the input and output power grows, which leads to larger rates of heating or cooling. Along the curve, the steady-state temperature is continuously elevated above the ambient level in proportion to the total input power, with

$$T_{\text{fixed}} = T_0 + R_{\text{th}} IV. \tag{1.6}$$

The fixed point curve is an important characteristic of the thermistor whose shape depends on a balance of electrical and thermal factors. A particularly

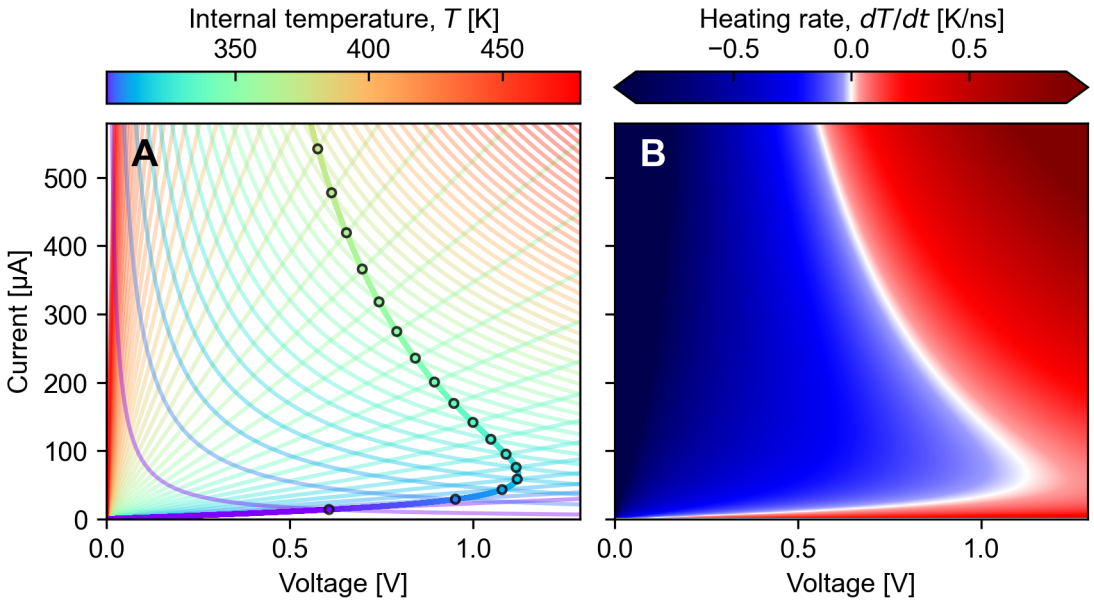

Figure 1.9: A projection of thermistor fixed points ($dT/dt = 0$ K/s) on an (I, V)-plane show a continuous curve with S-type NDR. **(A)** Solutions are visualized as the intersections of like-colored lines. Lines radiating from the origin represent the conduction Eq. 1.1 and lines of constant power show the corresponding conditions for thermal equilibrium $P_{\text{in}} = P_{\text{out}}(T)$. **(B)** The rate of change of the thermistor temperature depends on the instantaneous voltage and current through the device with respect to the NDR curve. On the left of the curve there is net cooling, on the right there is net heating, and the cooling/heating rate increases rapidly with the distance to the curve.

consequential feature that can arise here is NDR, or a negative value of dV/dI somewhere along the curve. In this case, the set of (I, V) fixed points can also be called the NDR curve of the cell. Unfortunately, substitution of T_{fixed} into a conduction equation with an Arrhenius-like temperature dependence (e.g. Eq. 1.1), tends not to have an analytical solution for $V(I)$, but NDR curves can still be treated numerically and graphically. Regions of NDR can appear and disappear as the model parameters vary. The general topology of this can be seen by extending fixed point projection into a third dimension in the vicinity of an (I, V)-plane with NDR. For example, as the ambient temperature T_0 varies, the fixed points lie on a continuous, multivalued surface reminiscent of the cusp geometry studied in catastrophe theory (Fig. 1.10) [134].

Although its temperature and therefore its conductivity can change significantly under external bias, the thermistor returns to its original temperature state ($T = T_0$) after a period of time without bias. Therefore, in a memory context it can be referred to as a *volatile* switching device; resistance state changes are temporary and the device does not retain any memory of past states after power is removed for a sufficient amount of time.

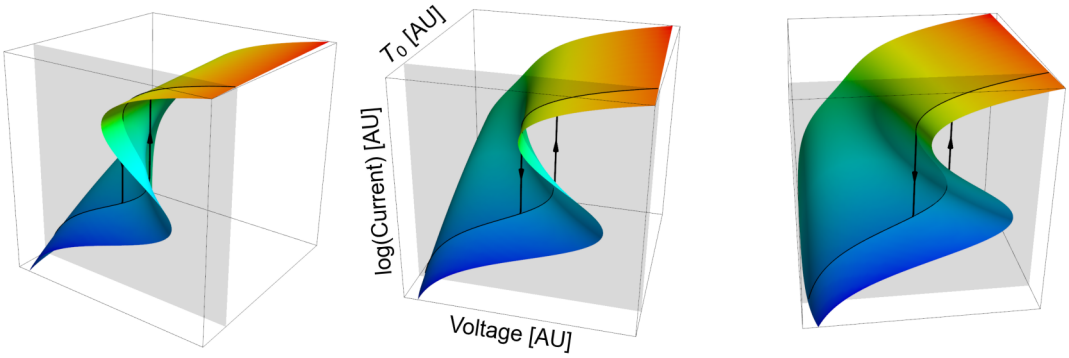

Figure 1.10: Projection of the thermistor fixed point manifold ($dT/dt = 0$ K/s) in the 3D space of voltage, current, and ambient temperature T_0, shown from three separate points of view. The surface color corresponds to the internal cell temperature T. The black path shows transitions in the stable points of the system when the voltage is viewed as a swept control variable at room temperature. For high enough values of T_0, the NDR disappears.

1.4.2 The geometry of volatile switching

The discussion leading up to the differential equation Eq. 1.5 considers only the internal properties of a thermistor device. However, the NDR curve alone does not tell us what happens in a measurement of the device, nor whether any of the NDR points will necessarily be observed. In order to talk about a measurement, we first need to specify the external circuit that drives it. Mathematically, this may require adding extra equations to the system that act as constraints or add additional dynamical state variables. An analysis of measurable behavior and types of control then centers around an understanding of stability of the overall circuit.

For the purposes of this discussion, we confine our attention to an external measurement circuit containing an external series resistance and a (parasitic) parallel capacitance. We will also consider the effect of internal resistances, which are often an unintentional side effect of fabrication, coming from the electrode leads for instance, but can also be a deliberate part of the cell design. The connection of the considered circuit components is shown in the schematic of Fig. 1.11. Variations of the component values in this circuit encompass many common measurement configurations, both ideal and practical. For example, when $C_p = 0$ pF, limiting cases include ideal voltage sourcing ($R_{s,ext} \to 0\ \Omega$) and current sourcing ($R_{s,ext} \to \infty\ \Omega$). In practice, a non-negligible value of C_p is present, and its consequences will also be discussed. Although capacitances between the internal nodes of the circuit may also be present, they are typically much smaller in suitably designed nano-scale test structures than the external value due to the probing interconnect (femtofarads vs. picofarads), and will be neglected here for simplicity.

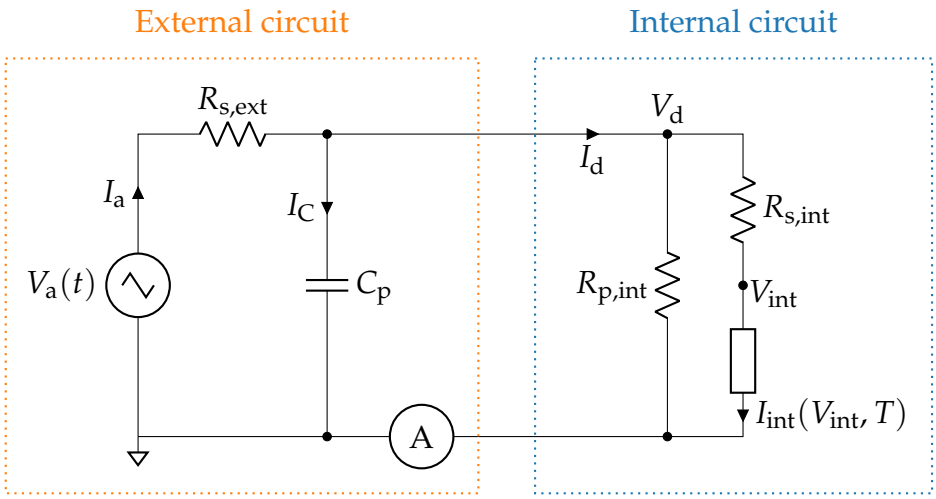

Figure 1.11: Generic form of the measurement circuit under consideration, split into sections that are external and internal to the test chip. Internal resistances may be present in series and in parallel with an electro-thermally active volume. A time varying voltage source drives the measurement and an ideal ammeter samples the current on the return path.

Two-dimensional plots of data are arguably the most important tool in the scientific arsenal for augmenting human intelligence. As such, electrical measurement data are commonly represented on plots of important variable pairs such as current vs. voltage. However, there are several different voltages and currents that appear in a measurement circuit which could be potentially measured and plotted, and the form of the (I, V) curves are meaningfully different depending on which I and V are selected, as well as on the methods used for their measurement. Furthermore, it is not always practical to make a direct measurement of the most relevant voltage/current for a given analysis, which affects the way the data should be interpreted. Therefore, to avoid ambiguity, (I, V) plots always need to be given in context of a circuit diagram that clearly identifies the locations of the measured signals. In the present case, for non-negligible internal resistance values, the measurable device voltage V_d and current I_d are different from the variables V and I in the equations of the preceeding discussion, which are replaced in the circuit schematic by $V \rightarrow V_{int}$ and $I \rightarrow I_{int}$. These internal values are those across the active volume of the cell which drive possible resistance transitions, but are not measured directly.

In this simplified case where there are no internal reactive components, the voltages and currents in question are related by the linear transformation

$$\begin{bmatrix} V_d \\ I_d \end{bmatrix} = \begin{bmatrix} 1 & R_{s,int} \\ \frac{1}{R_{p,int}} & \left(1 + \frac{R_{s,int}}{R_{p,int}}\right) \end{bmatrix} \begin{bmatrix} V_{int} \\ I_{int} \end{bmatrix}. \tag{1.7}$$

Importantly, this transformation affects the shape of NDR curves from the

point of view of the external measurement circuit, as shown in Fig. 1.12. Therefore, the values of the internal resistances can determine whether and in which range NDR appears, and is also one of several ways for the curve to return to positive differential resistance (PDR) after an intermediate region of NDR.

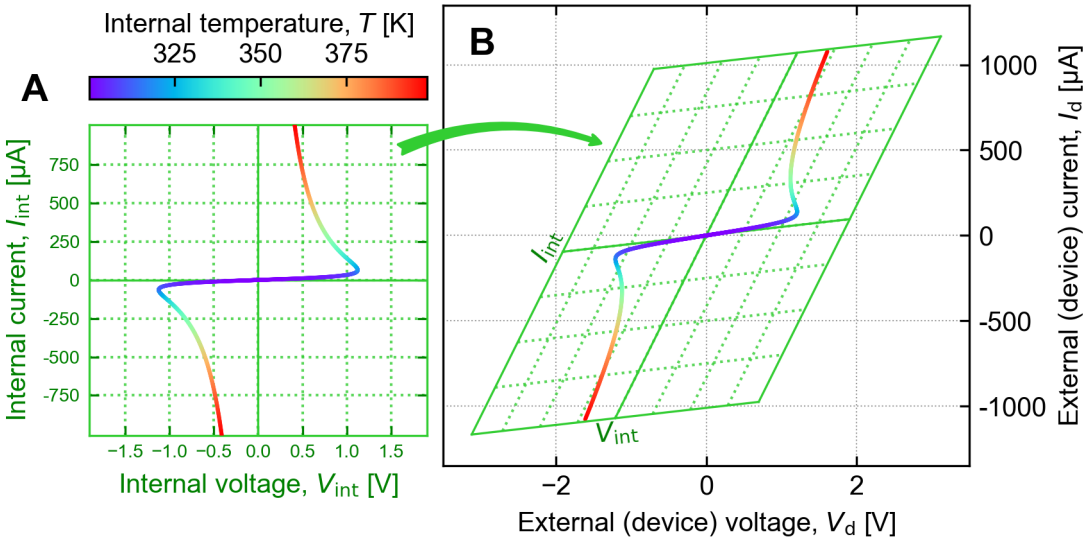

Figure 1.12: **(A)** Internal (I_{int}, V_{int}) characteristics are linearly transformed into **(B)** external (I_d, V_d) characteristics by measurement through internal resistances. This has important implications for the stability of the overall circuit because it distorts the external view of NDR curves. The plotted NDR data is the simulated steady-state of a thermistor volume with $R_{s,int} = 1.2\ k\Omega$, $R_{p,int} = 20\ k\Omega$.

As for the influence of $R_{s,ext}$, consider first the case where C_p=0 pF. Here, any measurable switching trajectory (I_d, V_d vs. t) is constrained by a one dimensional "load line", fixed by the values of $R_{s,ext}$ and the applied voltage V_a:

$$V_d = V_a - I_d R_{s,ext}. \tag{1.8}$$

The space representing each possible state of a dynamical system is called its phase space. Although the cell temperature T is the sole physical state variable of this system, a segment of the load line can also be viewed as a proxy phase space, due to a one-to-one correspondence with T. The load line can have up to three intersections with the (I_d, V_d) NDR curve, which define the remaining, discrete fixed points of the overall circuit ($dI_d/dt = 0$ A/s and $dV_d/dt = 0$ V/s). In the autonomous case (V_a constant in time), (I_d, V_d) will move along the load line from the initial condition toward one of the stable fixed points, corresponding to a final steady-state cell temperature. However, in loose analogy with the Newtonion situation of a ball on a hilly landscape, not all of the NDR intersections are *stable* fixed points (see Fig. 1.13). Near any fixed point, the heating and cooling rates are relatively slow, and stability is determined by whether or not the circuit restores the temperature to the fixed point level after an inevitable perturbation.

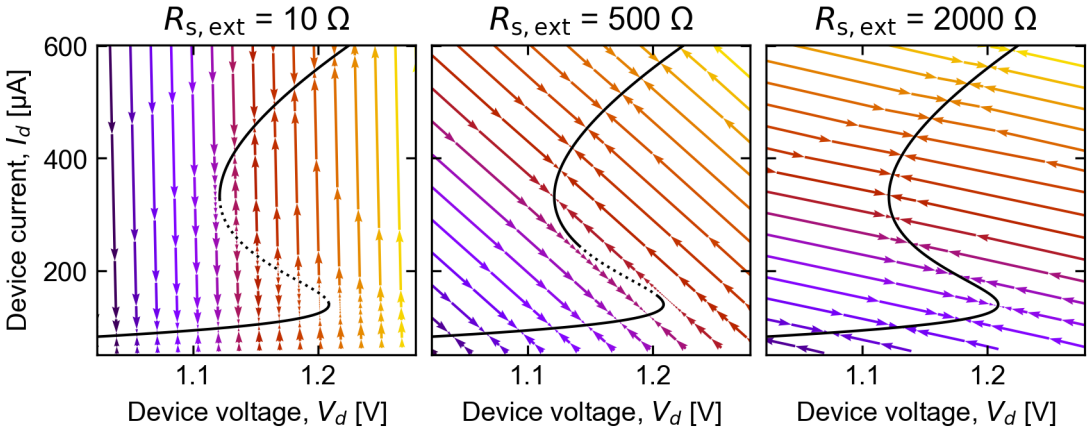

Figure 1.13: A visualization of the load-line dynamics in a thermistor measurement with negligible capacitance, C_p. Each arrow represents a duration of 0.8 ns, following the load lines colored according to different applied voltages V_a, for three different values of external series resistance $R_{s,ext}$. Depending on $R_{s,ext}$, the NDR curve is divided into monostable and bistable ranges with respect to V_a. For low $R_{s,ext}$, the dotted portion of the NDR curve is unstable, and for high $R_{s,ext}$, the entire NDR curve is stable.

Translating this to a specific measurement context, suppose a room temperature cell is characterized by applying a square voltage waveform (Heaviside function) to its electrodes. The corresponding measurement trajectory begins at the intersection of the load line with the isothermal device (I_d, V_d) curve evaluated at $T_0 = 300$ K, and moves along the load line toward the first intersection with the NDR curve, where it eventually settles (see Fig. 1.14). Different speeds along this path are possible, which can be understood with reference to the NDR curve and the external load characteristic. For (I_d, V_d) values lying nearby the stationary points, particularly near the "knee" of NDR onset, the rate of change of the temperature is low because a large fraction of the input power is dissipated to the external environment. Therefore, if the load line travels nearby the thermal equilibrium points of the NDR curve, the result will be a long delay as the applied power does not lead to rapid temperature increase of the cell. On the other hand, a larger series resistance combined with a larger applied voltage allows the switching trajectory to avoid these near-equilibrium values, and reach the same final state in a shorter time and potentially with lower total input energy [45].

There is a temptation to associate a time constant with heating processes, as the familiar scenario of heating a lumped-capacitance object with a constant power source admits exponential solutions for the temperature vs. time. However, this is very untrue in the case where the rate of heating is coupled to the temperature through the thermally activated conduction mechanism. According to the model, all heating rates are proportional to the cell heat capacity, but orders of magnitude different switching timescales can be induced in the same

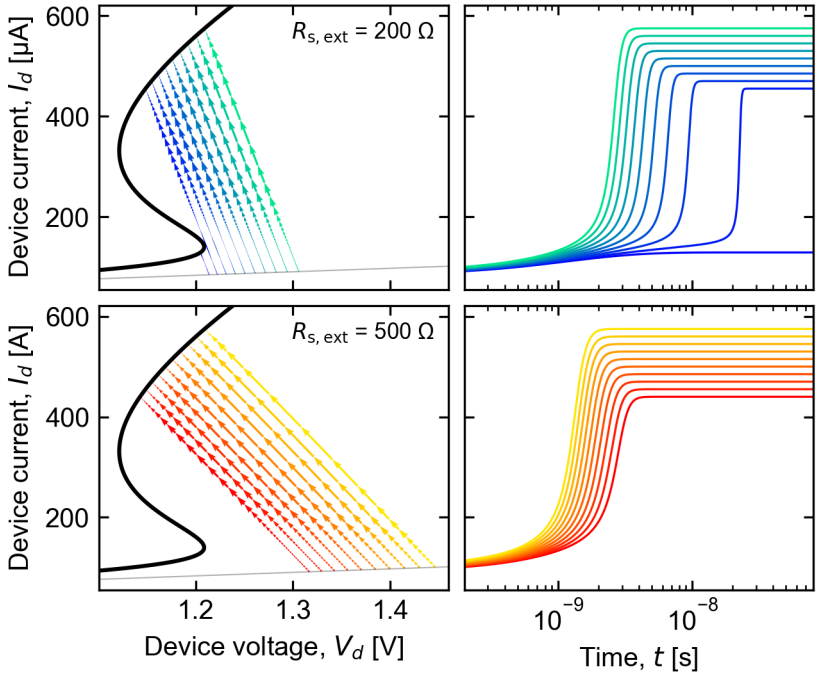

Figure 1.14: Time-dependent $(I_\mathrm{d}, V_\mathrm{d})$ trajectories following ideal voltage pulses with different amplitudes (color), with the cell initially at ambient temperature $(T = T_0)$. Using two slightly different series resistances ($R_{\mathrm{s,ext}} = 200\ \Omega$ above, and $R_{\mathrm{s,ext}} = 500\ \Omega$ below), the trajectories eventually reach the same final points of the NDR curve, but at different speeds. The farther the load line lies from the NDR curve, the faster the heating process. Delay-before-transition behavior is caused by load lines that first pass nearby, and then away from the NDR curve.

cell even in a narrow range of applied voltage. The thermistor temperature vs. time is not exponential in general and can show a long initial delay before eventual transition to the final high temperature state. These delays are caused by inefficient heating due to a switching trajectory passing nearby a condition of thermal equilibrium, where the cell is forced to spend long periods of time at elevated temperature while losing energy to its surroundings. Therefore, the delay is not an inescapable fact of the device heating process, but is due to a constraint imposed by the external driving circuit.

Measurements using a continuously swept voltage source are also common and important to consider. Again assuming negligible C_p, using high enough $R_{\mathrm{s,ext}}$ means that every point of intersection of the (time dependent) load line and NDR curve is a stable fixed point and therefore the NDR curve can be entirely measured using a slow enough sweep. The condition for a complete, stable characterization of the NDR curve is that the external resistance is equal to or larger than the maximum amount of NDR along the curve,

$$R_{\mathrm{s,ext}} \geq \max\left(-\frac{dV_\mathrm{d}}{dI_\mathrm{d}}\right), \tag{1.9}$$

or equivalently that NDR does not appear when plotted on the (I_d, V_a)-plane.

Using lower series resistance than the condition of Eq. 1.9 leads to a different type of behavior in a sweeping measurement (see Fig. 1.15). At two separate points during the sweep, where the NDR slope crosses the value of $-R_{s,ext}$, there are abrupt changes in the stability of the system (bistable $\rightarrow$ monostable). As V_a is swept, a previously stable fixed point is annihilated (along with an unstable one) as the load line suddenly ceases to intercept the NDR curve at those points. If the annihilated fixed point was the previous operating point, the system transitions along the load line, forced into regions of relatively high heating/cooling rates. In dynamical systems theory, this well-known behavior is called a saddle-node bifurcation [135]. Following the bifurcation, the system may settle at a new and potentially distant operating point; but if there are no remaining fixed points, runaway heating can lead to a cascade of increasingly destructive effects, eventually causing an irreversible breakdown of the cell [90, 120].

In a measurement, a saddle-node bifurcation effectively means that at a certain threshold level of V_a, a rapid and sometimes dramatic jump in I_d and/or V_d (with $dV_d/dI_d = -R_{s,ext}$) will be recorded as the stability landscape of the system suddenly changes with the biasing level. The common name given to measurements of these reversible jumps during slow voltage sweeps is threshold switching (TS). This term is usually used in context of memory selector devices, which are needed to avoid sneak paths in resistive crossbar arrays [136–139].

Although thermistors are not internally bistable devices, bistability appears as a property of certain circuit configurations. Nevertheless, the terminology "threshold switch" is commonly applied to such devices, even though the device alone is not truly a switch [40, 140–142]. A typical mental abstraction is that the TS device has two distinct resistance states; starting in its high resistance state (HRS), it transitions at a certain speed into its low resistance state (LRS) after it sees its "threshold" voltage, and when the voltage is reduced below its "hold" voltage it transitions back to its original HRS. However, in light of the present discussion, we recognize that this is not a complete picture for a thermistor-type switch. In fact, the device temperature continuously increases along its NDR curve, which can potentially be divided into stable and unstable regions depending on the load presented by the external measurement/driving circuit. The "HRS" and "LRS" here are not static device states but are two stable segments of the continuous NDR curve, and the so-called threshold and hold (I_d, V_d) points occur due to saddle-node bifurcations during sweeping. In each case, the locations of these thresholds can be predicted from the shape of the device NDR curve and the load characteristic, but neither switching speeds nor the threshold current densities and electric fields can be directly interpreted as a property of the device materials alone [143].

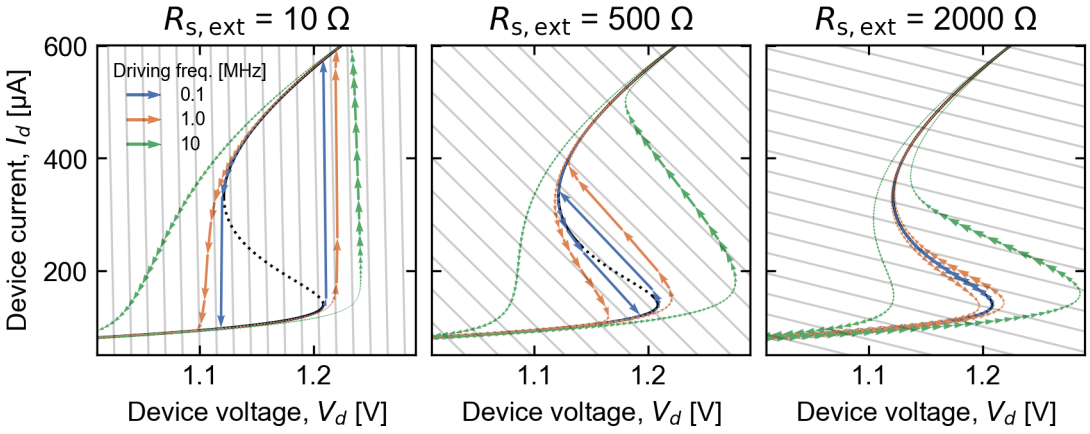

Figure 1.15: Simulated (I_d, V_d) trajectories for a thermistor driven by a sine waveform $V_a(t)$ with different frequencies and using three different external series resistances $R_{s,ext}$. For low $R_{s,ext}$ and low driving frequency, the system suddenly transitions along the respective load lines at the saddle-node bifurcation points. For high $R_{s,ext}$, the entire NDR curve can be traced out at low frequency. Higher frequencies push the cell further out of equilibrium due to thermal inertia, which widens a rate-dependent hysteresis in the trajectories. Load line slopes (gray) are plotted in the background for reference.

1.4.3 Oscillations and excitability

Finally, consider the case where $C_p > 0$ pF and is large enough to significantly affect the dynamics of the system. The load line constraint of Eq. 1.8 in this case is replaced by a differential equation that describes the charging of the capacitive V_d node,

$$C_p \frac{dV_d}{dt} = \frac{V_a - V_d}{R_{s,ext}} - I_d. \tag{1.10}$$

With the introduction of this equation, we now have a second-order (or planar) dynamical system, and we can consider the (I_d, V_d)-plane as a proxy phase space, because (I_d, V_d) identifies both the charge on the capacitor and the cell temperature state. The load line is still relevant as it becomes the condition where $dV_d/dt = 0$ V/s, which can also be called the V_d-nullcline. Therefore, the fixed points of the system ($dI_d/dt = 0$ A/s and $dV_d/dt = 0$ V/s) are not modified by the presence of C_p. However, the increased dimensionality introduces further possibilities for stability/instability of these fixed points [135, 144]. In particular, excitable and oscillatory modes exist depending on the particular biasing conditions and the value of C_p, as shown in Fig. 1.16. Operating in the oscillatory mode, the circuit is classified as a Pearson-Anson relaxation oscillator [145].

The oscillation behavior in this system can be understood as the alternating movement of energy between the reservoirs of cell temperature and the capacitor charge. After the bias V_a is initially turned on, the parallel capacitance

starts charging through the external series resistance. Eventually the device current and voltage are large enough to cross a spiking threshold and initiate thermal runaway in the thermistor material. As the device resistance rapidly drops, the capacitor discharges and delivers additional energy to the cell, raising its temperature well above the fixed-point level. Following the discharge, the elevated temperature cannot be maintained because of the current limiting feedback of the series resistance. By the time the temperature reduces back to the fixed-point level, the capacitor has lost too much charge to sustain it, so the device continues to cool down before the cycle repeats and approaches a limit cycle.

Sustained oscillation can only occur when the system is biased inside of (I_d, V_d) range with NDR [146, 147]. Another necessary condition for oscillation is that the electrical capacitance must be large enough relative to the thermal capacitance, in order for its discharge to elevate the device temperature to a level high enough above the steady-state temperature set by the series resistance and applied voltage. The exact biasing range where oscillations arise depends on the value of C_p, as shown in Fig. 1.17. There are at least two important timescales involved in the oscillations that affect the frequency. The relaxation time or refractory period is determined by electrical RC charging ($R = R_d||R_s$, $C = C_p$), and the duration of the falling edge of the spike is limited by the thermal discharge time $R_{th}C_{th}$ of the device.

In computational neuroscience, simple dynamical models are used to mimic the evolution of the neuron membrane potential, with the Hodgkin-Huxley model being the most famous example [148, 149]. Excitability in these systems is associated with quiescent points lying near bifurcations, with different types of bifurcations giving rise to various types of spiking and bursting patterns thought to be connected to the computational properties of neurons [150–152]. Simple circuits including a thermistor-like component, such as the one considered here, can be shown to be topologically equivalent to simplified planar neuron models such as the FitzHugh-Nagumo model [153, 154], and are thus capable of generating same set of spiking bursting patterns when biased in specific ways. This correspondence is responsible for the interest in using thermal feedback circuits as artifical neurons in NC.

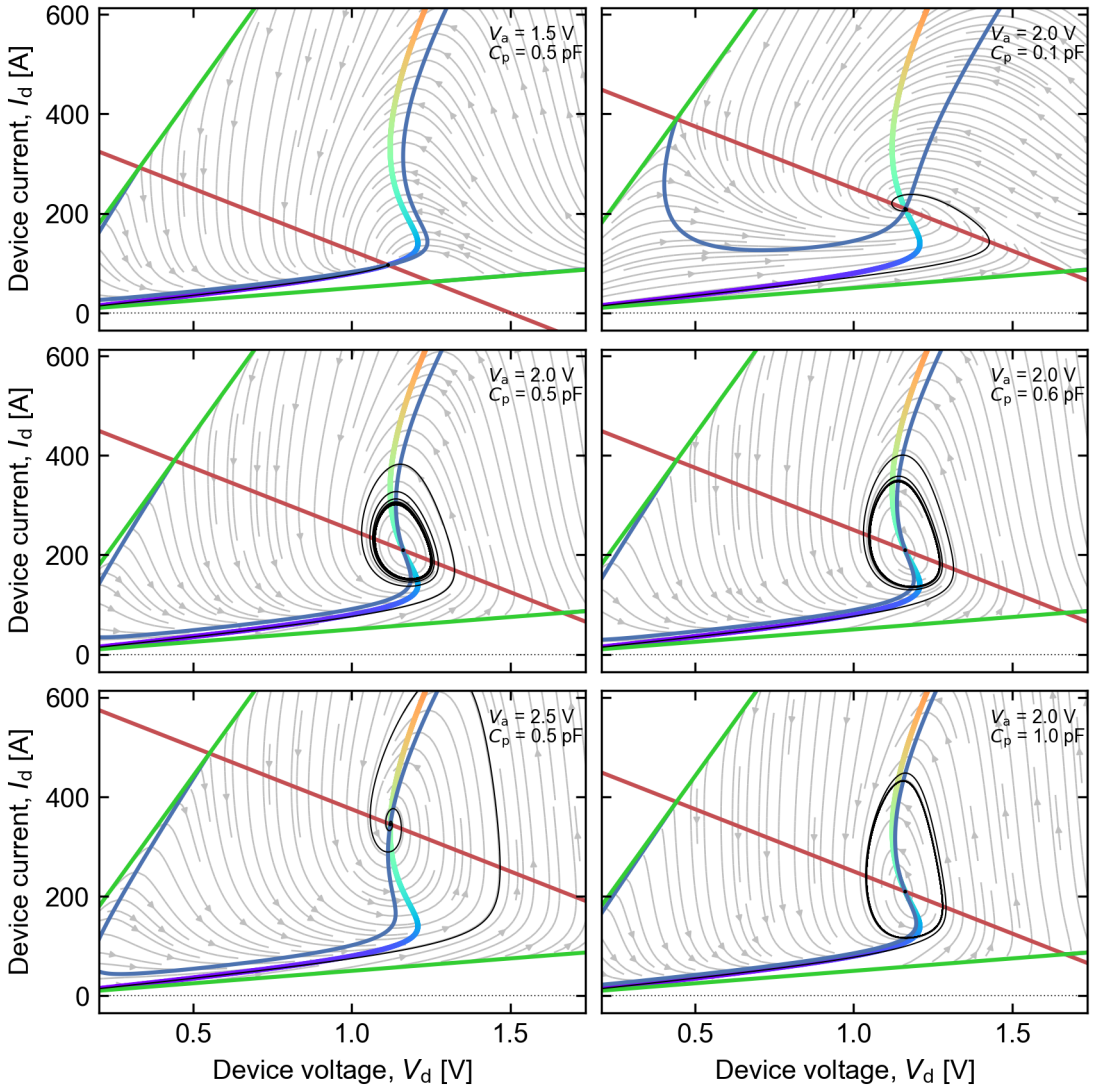

Figure 1.16: Phase portraits for a thermistor in a circuit with parallel capacitanc and external series resistance $R_{s,ext} = 4$ kΩ, under six different biasing conditions. Left and right columns show the effect of increasing V_a and C_p, respectively. The red and blue lines are the V_d- and I_d-nullclines, respectively, and the multicolored (according to T) line is the NDR curve, or T-nullcline. Lime-green lines bound the phase space (due to the internal resistances), and the black curve shows the trajectory with initial condition $T = 300$ K and $V_d = 0$ V. For a range of conditions, the trajectories are oscillatory.

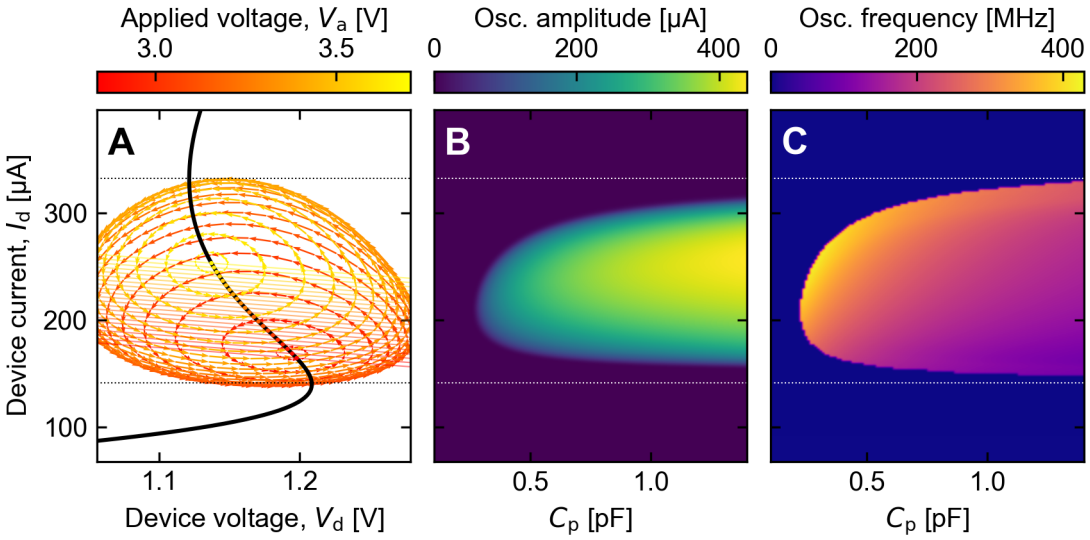

Figure 1.17: Simulated oscillatory limit cycles in dependence of the biasing level V_a and the parallel capacitance C_p, with $R_{s,ext} = 10$ kΩ. **(A)** NDR curve with limit cycles occuring in a subrange of operating points when $C_p = 0.35$ pF. Peak-to-peak current amplitude **(B)** and frequency **(C)** vary with C_p and the fixed-point current. Crossing the discernable boundary between non-oscillation and oscillation is called a Hopf-bifurcation.

1.4.4 Field effects

This section has explained a range of volatile switching phenomena in devices with temperature dependent conductivity while assuming linear voltage dependence of the current. The exponential temperature term in the conduction mechanism is all that is necessary to see these effects, which can also be observed in commercial temperature sensing devices (see Appendix B). In this case, the device is called a thermistor and any observed $(I_\mathrm{d}, V_\mathrm{d})$ non-linearity must be caused by device heating. However, most if not all scaled cells show a non-linear voltage dependence due to various other impacts of the high electric fields produced across the short insulating gaps.

Along these lines, a distinction was already pointed out in the 1960s and 1970s between thermistor and electro-thermal action [155–158]. More modern works considering nanodevices for technological applications usually employ electro-thermal models with non-linear voltage dependence, where Poole-Frenkel-type conduction is an especially popular assumption [37, 159, 160]. If the voltage dependence is monotonic and does not introduce further relevant state variables to the system, it can be considered a point of detail useful for matching measurement data, but it does not fundamentally change the thermal character of the system.

The conditions for NDR in a simplified electro-thermal case were discussed over 60 years ago by Burgess [155], and were more recently elaborated by Gibson [161]. Among the results is that, mathematically, the temperature dependence of conductivity need only increase more than linearly for NDR to appear at some current level. However, additional physical effects such as ionic migration, phase transitions, crystallization, and melting are simultaneously possible, and are similarly capable of inducing resistance transitions in the material. These different possible effects act in parallel and can come into play before or after the thermal NDR onset point, depending on the material, the timescale, and other factors. Whether the presence of electro-thermal effects are decisive in the overall observed switching process, merely incidental, or something in between, can only be evaluated with an understanding of the various consequences of thermal feedback covered here.

Chapter 2

Electrical measurement setup

This chapter is partially based on Ref. [162] and Ref. [163], and was reproduced with the permission of AIP Publishing.

2.1 Digital potentiometer stabilization circuit

Electrical measurements of patterned devices are inevitably carried out in the presence of resistance in series with the active material volume of the cell. This series resistance, commonly of unknown value [164, 165], may originate from a combination of the electrode leads, inactive layers of the material stack, or the triode region of a series FET current limiter. Internal and external series resistance adds current-voltage feedback to the system that affects stability and influences the operational behavior in important ways. Modification of switching speeds, threshold voltage/currents, and the range of achievable resistance states have all been observed and discussed theoretically [45, 100, 109, 111, 112, 166].

A series resistance is often deliberately placed to play the necessary role of an energy limiting mechanism , where its value can mean the difference between a functioning and non-functioning device. As an experimental technique, it is useful to be able to place different resistance values in series with the device under test (DUT) to examine the effect on switching processes. An important advantage of this simple two element series configuration (2R) is that the circuit response is easily predictable through load line analysis in the ideal case, and is also straightforward to model analytically in the presence of commonly encountered parasitics.

Another advantage of the 2R configuration is ease of implementation relative to integration of active FET devices on a test chip, with the latter requiring substantial fabrication cycle time. However, integrating calibrated series resistances is inflexible because each cell is attached to a single static resistance value that cannot be changed or removed. Scenarios often arise that give good reason to alter or remove the series resistance *in situ*. Notably, devices possessing steady states with S-type or N-type NDR each have different criteria for stable characterization, and both types are commonly present in the SET and RESET processes of ReRAM, respectively. This imposes different requirements for the series resistance value even within a single switching cycle.

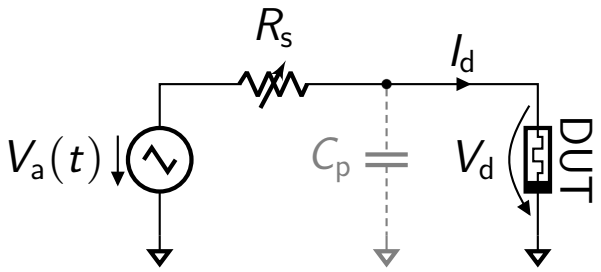

Figure 2.1: A simple circuit configuration for device characterization uses a waveform generator and an external resistance in series with the DUT. In practice, the effect of the parasitic capacitance in parallel with the device requires careful attention.

Where an adjustable series resistance is required, it must be implemented externally to the wafer. The main practical challenge associated with this is that parasitic capacitance C_p at the node shared with the DUT is highly detrimental and difficult to avoid (Fig. 2.1). This stray capacitance slows down the dynamic response of the circuit, degrading the ability to control and to measure the voltage and current experienced by the active cell volume versus time. Coupled with rapid conductance transitions of the DUT, harmful overshoot transients are generated that strongly impact the observed switching behavior and can cause irreversible damage [90, 103, 105, 120].

While singular through-hole resistors are a common external solution, their use entails manually switching between resistance values where required. However, the stochastic nature of resistive switching cells is such that they benefit greatly from a statistical treatment using automated measurements with programmable parameters. In this section we present an external circuit design providing an adjustable linear series resistance for flexible wafer-level device characterization. The circuit, based on a digital potentiometer (digipot) chip, is remotely programmable over USB between 528 resistance levels. Importantly, the voltage signal at the low-capacitance DUT node is directly amplified for synchronous measurement with the DUT current with a bandwidth over 200 MHz. We demonstrate the circuit operation for automated characterization of NDR devices and for resistance cycling of bipolar ReRAM cells with high speed voltage sweeps.

2.1.1 Design principles

Applying Kirchhoff's current law, the dynamical equation governing the time evolution of the device voltage in the circuit of Figure 2.1 is

$$C_p \frac{dV_d(t)}{dt} = \frac{V_a(t) - V_d(t)}{R_s} - I_d(t, \ldots), \tag{2.1}$$

where t is time and I_d in general depends on V_d and other internal state variables of the DUT. Possible steady state solutions lie on the V_d-nullcline,

$$V_d = V_a - I_d R_s, \tag{2.2}$$

also known in static analysis as the load line. For fast conductance switching events that are common in the targeted material systems, transient deviations from the load line occur as seen in the simplified situation of Fig. 2.2. During such transients, the excess energy delivered to the DUT due to capacitive discharge is significant and can strongly influence the end result of the switching process.

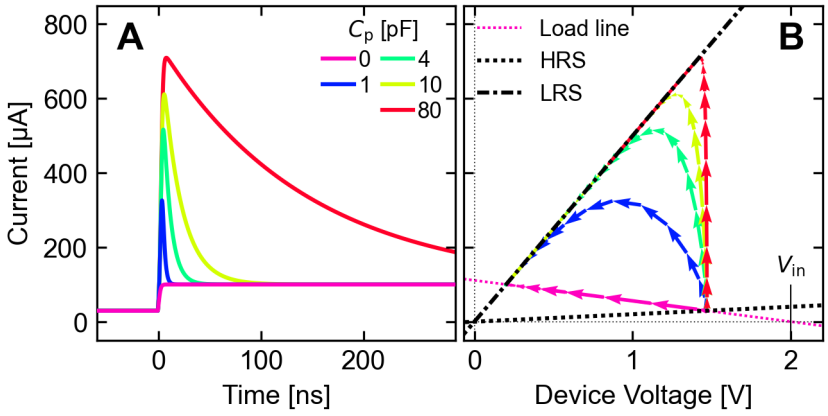

Figure 2.2: Simulations (using Eq. 2.1) of I_d, V_d transients following a rapid resistance transition of the DUT with $V_a = 2$ V and different values of C_p. Subplot **(A)** shows I_d vs. t while **(B)** shows I_d vs. V_d of the same simulations. The DUT resistance value is assumed to change exponentially in time from a high resistance state (HRS) of 50 kΩ to a low resistance state (LRS) of 2 kΩ with time constant 1 ns. During and following the transition, the device is subjected to excess currents relative to the load line, an effect which is reduced by using lower C_p values.

While the potential for overshooting transients is unavoidable in the context of a passive feedback arrangement, it is important to control them to the extent possible and to accurately measure them when they occur. The only way that overshoots can be reduced in the discussed configuration is by minimizing the value of C_p. Practically this means that a coaxial cable, acting approximately as a parasitic capacitance of 100 pF/m, cannot be used to connect R_s to the DUT. The series resistance should rather be placed as close as possible to the DUT, with the components carefully selected and the printed circuit board (PCB) layout designed for low contribution to the total C_p.

High fidelity current measurements can be achieved by amplification of the voltage across a ground referenced shunt termination following transmission over a coaxial line. Using this type of current measurement, positioning the DUT (rather than R_s) adjacent to the shunt is generally preferable because it avoids low pass filtering of the I_d signal. This allows measurement of I_d at

a high bandwidth that is independent of both R_s and the resistance state of the device. Experiments that prioritize application of high frequency voltage signals to the DUT may benefit from connecting R_s in the opposite orientation, but the voltage transfer function in this alternate configuration can distort wideband signals.

With prior knowledge of R_s, Eq. 2.2 is often used to calculate the V_d from a measurement of I_d and V_a, but there are several drawbacks associated with this method. One is the inaccuracy that comes from neglecting the capacitive currents of the left-hand side of Eq. 2.1. Another problem is measurement noise introduced by the $I_d R_s$ term, as the small I_d signal with high relative error is multiplied by a potentially large R_s value. It is therefore strongly advantageous to directly amplify the voltage at the DUT electrode rather than attempt to calculate it from other measured signals.

Following from these considerations, the basic intended configuration of external instruments and the designed circuit is shown in Fig 2.3. If sampling the current with a bare (50 Ω) oscilloscope input does not provide sufficient resolution, additional voltage amplification should be placed at the termination, and several output stages can be beneficial for dynamic range. Note that the length of the coaxial lines for DUT voltage and current sampling should be matched so that post-processing is not needed for high speed signal synchronization.

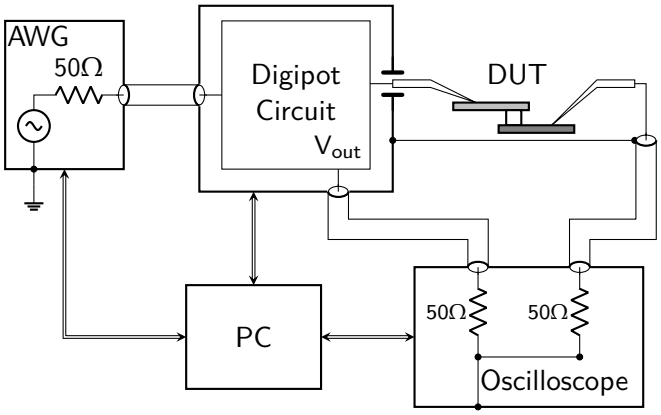

Figure 2.3: Schematic depiction of the overall measurement setup. An arbitrary waveform generator (AWG) produces the driving signal $V_a(t)$, and the resulting current is sampled after the right electrode via the 50 Ω shunt of the oscilloscope input. A second oscilloscope channel simultaneously captures the amplified voltage at the left electrode. A ground jumper provides a low inductance return path and reduces RF interference. All instruments are under computer control.

2.1.2 Implementation

A commercial integrated circuit (IC), the DS1808 digipot from Maxim Integrated, was chosen as the central component to control the series resistance R_s.

It contains two separate potentiometers internally, each consisting of a chain of 32 resistors whose junctions can be connected to a "wiper" output via a set of CMOS transmission gates (analog switches). For each potentiometer, there are 32 available resistance settings spaced logarithmically (piecewise) from approximately 300 Ω to 45 kΩ. According to the published specifications, the DS1808 has a maximum parasitic capacitance of 10 pF and a maximum voltage range of ±12 V [167].

To increase the coverage of R_s values, the PCB is routed in a way that allows connection of both potentiometers either in series or in parallel by connecting or opening solder jumper pads. While a connection to a single potentiometer remains possible, the number of unique settings is increased to 528 between 600 Ω − 90 kΩ for the series combination and between 150 Ω − 22.5 kΩ for the parallel combination. Because the individual digipots do not provide a resistance setting below 300 Ω, a reed switch was also included on the PCB to add an option to short the input directly to the output.

For amplification of the output voltage, the THS3091 current-feedback operational amplifier from Texas Instruments was used in a non-inverting configuration. This device features low distortion, low noise, a bandwidth of 210 MHz, and a slew rate of 7300 V/µs while adding only 0.1 pF parasitic capacitance [168].

All on-board settings are controlled via an Atmega32u4 microcontroller programmed as a USB serial interface to the PC. Control of the R_s value is accessible using any programming language able to open a serial COM connection to the microcontroller. This is done by sending a simple command composed of three integer values corresponding to the wiper positions and the state of the bypass relay. The total time from issuing a serial command to R_s reaching a new value is limited by USB / I^2C communication, and is typically less than 300 µs. The overall circuit design is visualized in the a block diagram of Fig 2.4, and a corresponding fabricated PCB is pictured in Fig. 2.5.

Further details regarding the digipot circuit implementation were recorded in the thesis of Wichmann [169].

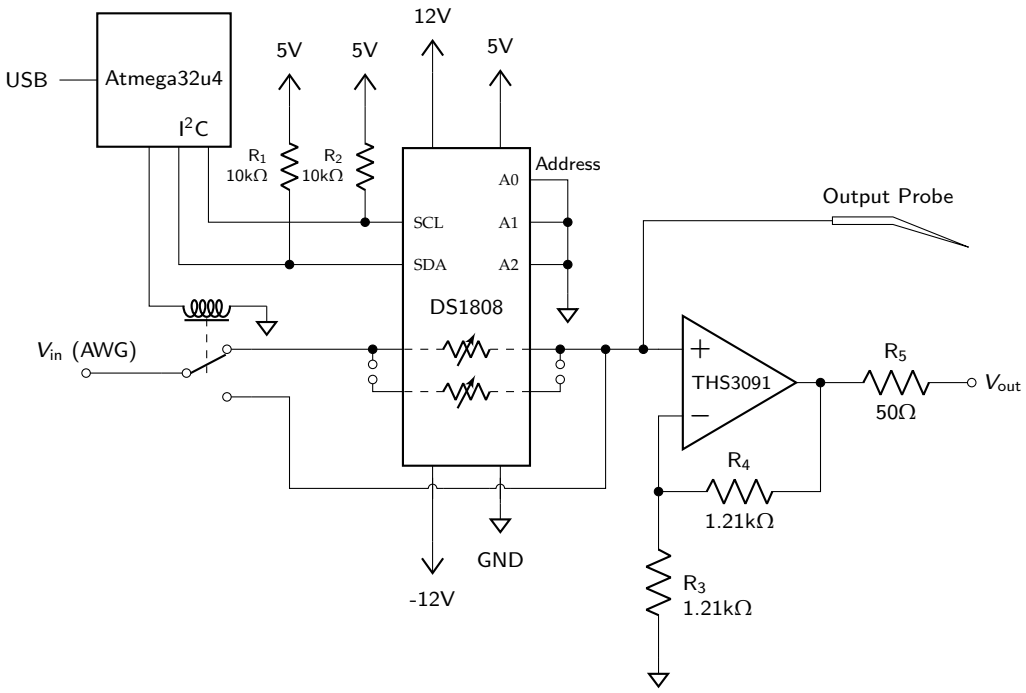

Figure 2.4: Simplified schematic of the digipot measurement circuit. An Atmega32u4 microcontroller USB-serial interface communicates to the DS1808 digipot via an I^2C bus. A SPDT reed relay can be actuated in order to bypass the digipot and make a direct connection between input and output. The voltage at the output is amplified by a THS3091 non-inverting follower.

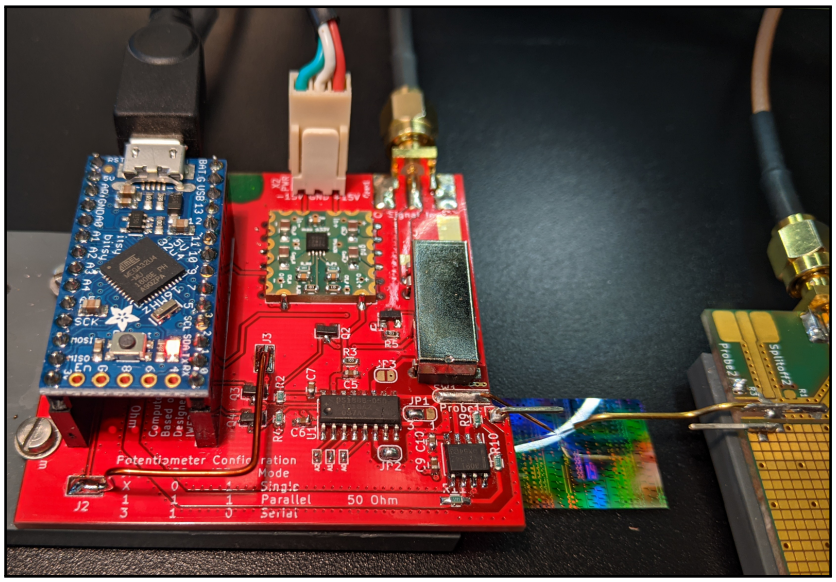

Figure 2.5: A photograph of the probing PCB contacting a test chip. A non-coaxial BeCu probe tip is soldered directly to the output of the main PCB (red), which uses SMA connectors for additional input and output signals. An elevated PCB (blue) contains the microcontroller USB interface (Adafruit ItsyBitsy 32u4). A square PCB module (green) functions as a low noise dual voltage regulator providing $\pm$ 12 V to the system. The right probe is directly connected to a 50 Ω oscilloscope input.

2.1.3 Measurement demonstrations

In quasistatic measurements of classical NDR materials using a series resistance, saddle-node bifurcations can occur that separate the NDR characteristic into stable and unstable regions. The range of the unstable region is determined by the value of the series resistor, with the bifurcations occurring where the derivative of the NDR curve voltage with respect to current crosses $-R_s$. While sweeping voltage, sudden current jumps are observed for sufficiently low values of R_s in S-type NDR (Fig. 2.6[A]) and for sufficiently high values of R_s in N-type NDR (Fig. 2.6[B]). Thus, an adaptable R_s value allows control of the conditions under which each of these characteristic curves, which contain important information, can be measured.

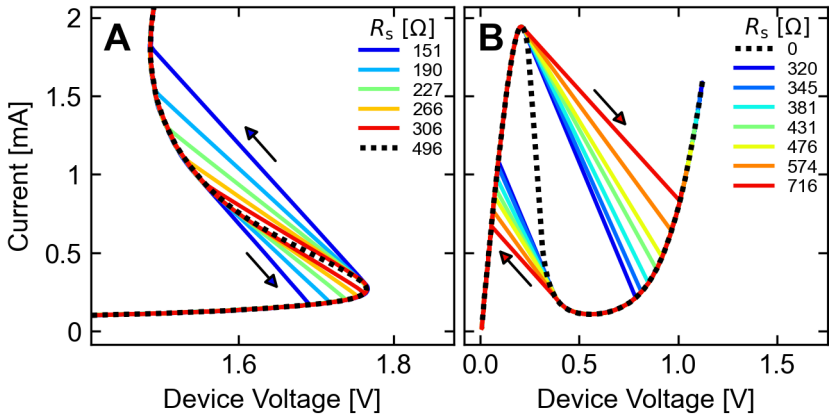

Figure 2.6: Voltage sweeping measurements of NDR devices using different resistance settings. **(A)** $90\times500\times500$ nm^3 S-type VCrOx device [44], stabilized for $R_s > 400\ \Omega$. **(B)** N-type Ga-As tunnel diode 3I306E, stabilized for $R_s = 0\ \Omega$.

Where the material mechanism of NDR is dynamic and reversible, the presence of C_p makes the measurement circuit prone to transient oscillations, and stable oscillatory limit cycles can also occur. In these cases, the presented circuit is able to capture high speed transients for accurate projection onto the (I_d, V_d)-plane (Fig. 2.7). This data is useful for device modeling and circuit simulations, each relevant for example in present investigations of coupled oscillatory devices in neuromorphic systems.

In ReRAM devices, NDR behavior is mediated by a combination of Joule heating and migration of point defects in the oxide material that locally increase its conductivity [20]. Altering the R_s value allows these transitions to be probed in different ways, as seen in the example measurements of Fig. 2.8. In analogy to the NDR measurements of Fig. 2.6, a fixed value of R_s can result in sudden and unstable transitions for one or both of the SET or RESET processes. By switching the value of R_s during the measurement (Fig. 2.8[C]) it is shown that runaway load line transitions can be suppressed by appropriate selection of the external feedback.

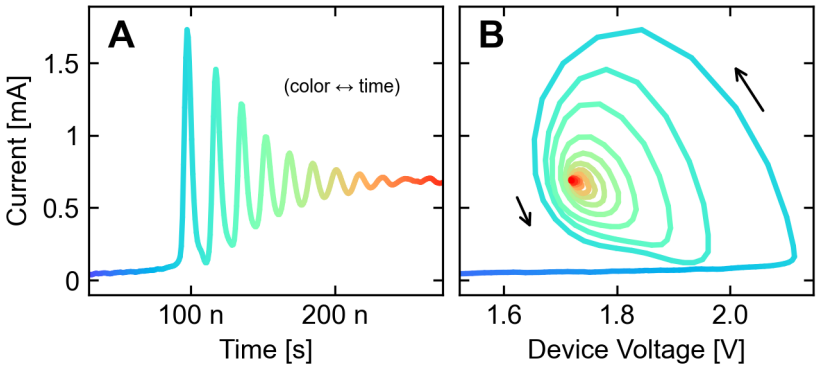

Figure 2.7: Oscillations (57 MHz) occurring in a 30×250×250 nm S-type VCrOx NDR device [44] following a square voltage pulse $V_a = 0\ V \rightarrow 2.5\ V$ using $R_s = 1083\ \Omega$. With the line color mapped to time of measurement, **(A)** shows I_d vs t of the transient, and **(B)** shows the trajectory of the same data on the (I_d, V_d)-plane.

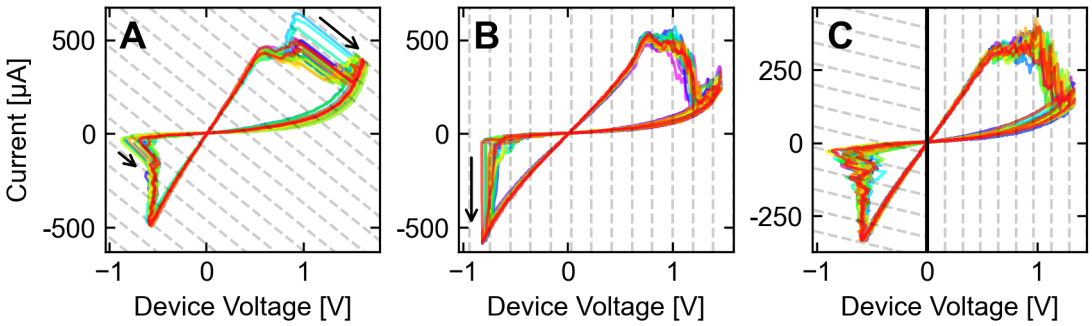

Figure 2.8: Cycling measurements of a 100 nm ReRAM device [163] using bipolar triangular voltage sweeps, each lasting 1 ms. Each subplot contains 20 consecutive switching cycles differentiated by color. The value of added series resistance is indicated by the dashed lines with gradient R_s^{-1}, and rapid load line jumps are indicated with arrows. Transition behavior differs considerably when using **(A)** 2.4 kΩ, **(B)** 0 Ω, and C 11 kΩ for negative polarity (SET) and 0 Ω for positive polarity (RESET).

2.2 Current limiting amplifier

Resistive switching devices face significant challenges related to control of delicate filamentary states in the oxide material. As a device switches, its rapid conductivity change is involved in a positive feedback process that would lead to runaway destruction of the cell without current, voltage, or energy limitation. Typically, cells are directly patterned on MOS transistors to limit the current, but this approach is very restrictive as the necessary integration limits the materials available as well as the fabrication cycle time.

In this section we propose an external circuit to cycle resistive memory cells, capturing the full transfer curves while driving the cells in such a way to suppress runaway transitions. Using this circuit, we demonstrate the acquisition of 10^7 (I, V) loops per second without using on-wafer current limiting transistors. This setup brings voltage sweeping measurements to a relevant timescale for applications, and enables many new experimental possibilities for device evaluation in a statistical context.

2.2.1 Design principles

For the purpose of rapidly testing devices with minimal nano-fabrication overhead, compatibility with isolated two-terminal structures is necessary and should be provided by an external current limiting amplifier (CLA) circuit placed in series with the device under test (DUT) in a setup similar to that shown in Fig. 2.9. When the series combination is driven by a voltage waveform, the circuit should provide a variable current limit in the approximate range 10 μA to 1 mA in the forward polarity (SET direction). Because of the self-limiting nature of the RESET process under voltage control, current should flow through the circuit unimpeded in the reverse polarity (RESET direction).

To avoid any influence of the circuit on the switching process before the current limit is reached, the circuit should present a negligible impedance for all currents below the limit. Only once the DUT current reaches the limit, the circuit should rapidly transition into a current source behavior to terminate the runaway switching process. In other words, the circuit should ideally present a frequency independent (I, V) characteristic as shown in Fig. 2.10(A) in series with the device. The circuit should be highly stable for a variety of loads, and its design should be as simple as possible in order to easily distinguish the role of the DUT in measurements of the overall electrical response.

Crucially, any overshoot above the current limit following a SET transition should be suppressed as much as possible. Because such overshoots are caused by the stray capacitance at the terminal of the current limiting circuit, this capacitance is considered a critical design parameter to be minimized. It is therefore not an option to connect the CLA to the DUT over a length of coaxial cable, as this would present an effective capacitance of 100 pF/m. To reduce this capacitance, the probing circuit needs to be mounted as close as possible

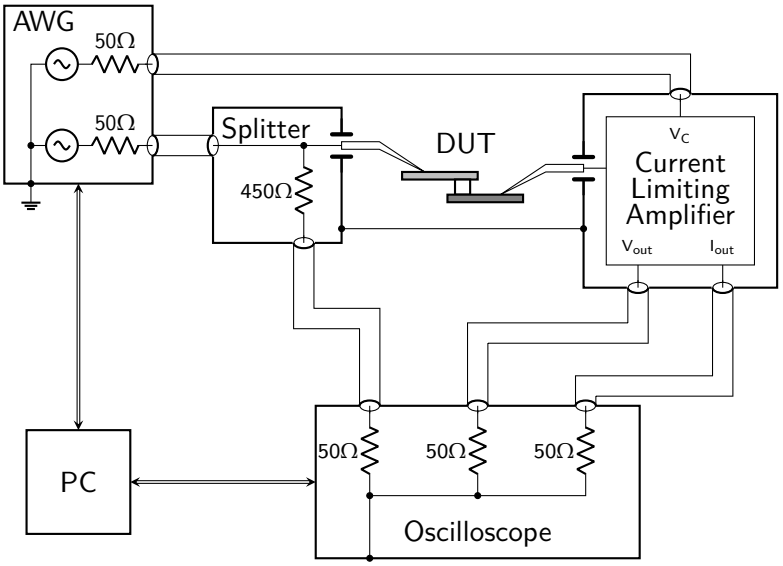

Figure 2.9: Schematic of a measurement setup using the current limiting amplifier circuit. A two channel arbitrary waveform generator (AWG) applies a driving signal to the DUT as well as a signal (V_C) to control the value of the forward current limit. An oscilloscope measures simultaneous samples of the voltage at both electrodes, as well as the device current. A jumper connects the ground planes of the left and right probes to reduce interference and inductance in the signal path.

to the DUT, and a short unshielded probe needle should be mounted directly to its circuit board.

In this type of measurement setup, two important bandwidths can be distinguished. The first is the bandwidth of the application of voltage signals to the DUT, and the second is the bandwidth of the measurement of current through the DUT. Simplistic external current limiting approaches using a series resistor or a common-source FET have the side effect of forming a low pass filter that limits one of these bandwidths depending on which side of the DUT the current limiter is positioned. In such setups, the limited bandwidth also depends on the resistance state of the DUT and on the current limit used. These bandwidth limiting effects should be circumvented in the CLA design. For all current limit settings and DUT states, the bandwidth of voltage application should be limited only by the AWG (100 MHz) and the DUT parasitics. The bandwidth of the low-noise current measurement should be large enough to accomodate the detection of rapid switching events with a rise time below 100 ns.

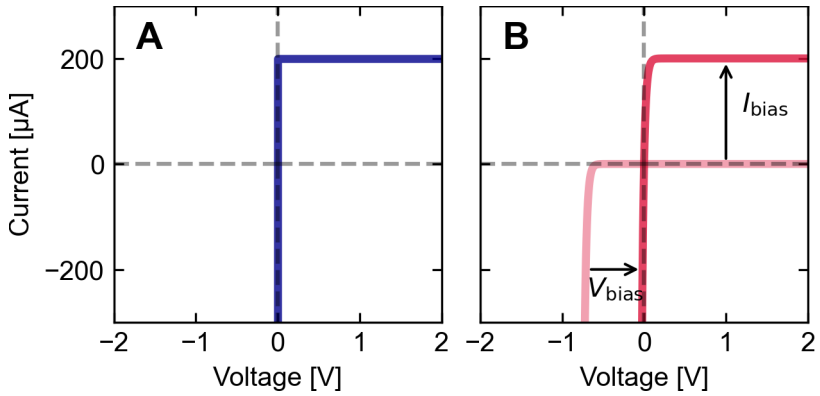

Figure 2.10: The current limiting (I, V) characteristic intended to be placed in series with the DUT. In the ideal case **(A)**, the differential resistance is zero below the adjustable current limit (here 200 µA), and infinite above. An approximation of the ideal characteristic **(B)** can be realized using a common-base amplifier with voltage and current bias.

2.2.2 Implementation

The fundamental idea behind the presented circuit design is to use a single bipolar junction transistor (BJT) (I, V) characteristic to implement the desired current limiting response while also providing transimpedance amplification of the DUT current. Packaged discrete BJTs for radio frequency applications are available with very low parasitic capacitance, making them highly suitable here for use in the input stage. The common-base (CB) amplifier configuration is of particular interest as a high-bandwidth current buffer, featuring a low input impedance and small feedback capacitance that does not suffer from the Miller effect. With voltage and current biasing, a CB amplifier can closely approximate the targetted current limiting (I, V) characteristic shown in Fig. 2.10. A simplified schematic of the input stage used to accomplish this is shown in Fig. 2.11.

The basic operation of this input stage is straightforward to analyze. Applying Kirchhoff's current law at the input node, it can be seen that whenever the DUT current I_d is less than the bias current I_bias, the BJT emitter current I_E is positive and transistor will be in forward-active mode. In this mode, with an appropriate setting of $V_\mathrm{bias} \approx 0.7$ V, the input voltage will be held close to 0 V due to the high forward transconductance of the BJT. Thus, for either positive or negative voltages applied to the DUT, the input stage effectively presents a low impedance to ground as long as $I_\mathrm{d} < I_\mathrm{bias}$. As I_d approaches I_bias, the BJT enters cut-off mode, where its effect in the circuit can be ignored and the input behaves as a current source with $I_\mathrm{d} = I_\mathrm{bias}$.

Ideally, the voltage bias V_bias should be chosen so that the input current is zero for an input voltage of zero (such that the curve of Fig. 2.10 intersects the origin). Considering an approximated Ebers-Moll model of the BJT, it follows

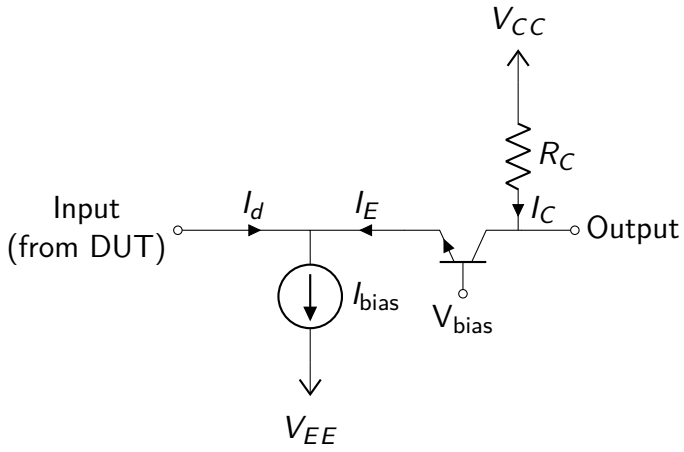

Figure 2.11: A simplified diagram of a circuit implementing unipolar current limiting and transimpedance amplification. The value of the forward current limit is set by I_{bias}, and the input voltage is approximately 0 V for input currents below this limit.

that

$$V_{bias} = -nV_T \log\left(\frac{I_{bias}}{I_s} + 1\right), \tag{2.3}$$

where I_s is the saturation current of the base-emitter junction, $V_T \approx 26$ mV is the thermal voltage, and n is the diode ideality factor. The output of this stage then gives an amplified voltage signal V_{out} that is linearly related to the input current

$$I_d = I_{bias} - \left(\frac{1+\beta}{\beta}\right)\left(\frac{V_{CC} - V_{out}}{R_C}\right), \tag{2.4}$$

where β is the forward common-emitter current gain of the NPN transistor.

A full circuit diagram expanding on this concept is given in Fig. 2.13, with a prototype PCB layout also pictured in Fig. 2.12. Here, Q_1 is the CB amplifier corresponding to that depicted in Fig. 2.11, and a nearly ideal voltage controlled current source is realized by the emitter degenerated cascode amplifier formed by Q_2, Q_3, and R_2. The dependence of the current limit I_{bias} on the control voltage V_c, which is approximately linear for $I_{bias} > 100$ µA, is calibrated for V_c values between -10 V and -1 V by an SMU measurement. The V_c signal is then generated according to interpolation of the calibration table at the desired I_{bias} values.

Further circuitry in Fig. 2.13 is included to null voltage offsets and condition the output signals for transmission to 50 Ω oscilloscope inputs. According to Eq. 2.3, the ideal value of V_{bias} depends slightly on the value of I_{bias}. Therefore, simply using a constant value of V_{bias} would create offset voltages at the input terminal on the order of $10 - 100$ mV as I_{bias} is varied. To automatically compensate this effect, a reference path R_3, Q_4, Q_5, Q_6, R_4 mirrors the components R_1, Q_1, Q_2, Q_3, R_2, and is used to actively zero the input offset for all values of I_{bias} via OPA277. This same structure also generates a reference voltage for a differential measurement performed by AD8130, producing

a low-offset output signal I_{out} proportional to the input current. A voltage follower (THS3091) with very low input capacitance (0.1 pF) is also placed directly at the input node, providing a simultaneous measurement of the DUT voltage drop.

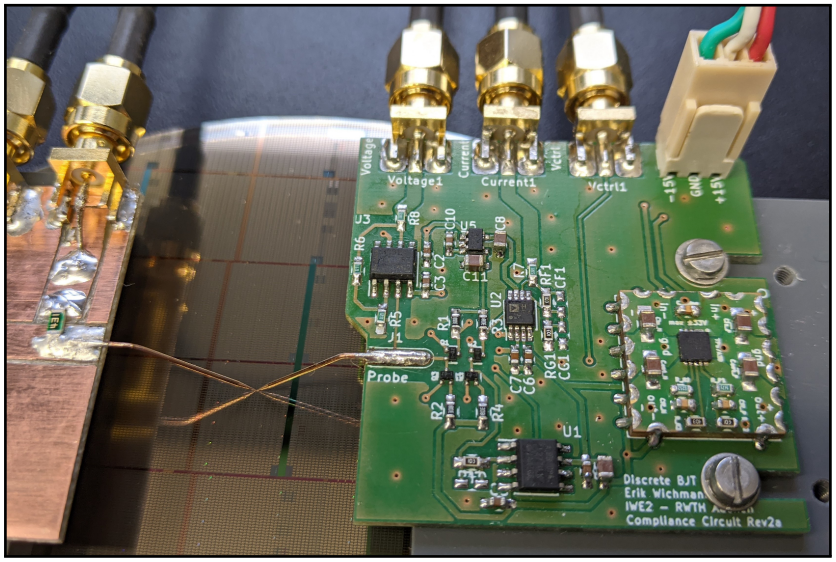

Figure 2.12: A photograph of the probing circuit board contacting a prototype ReRAM device. Left and right probes are mounted on independent micropositioners (ground jumper not pictured).

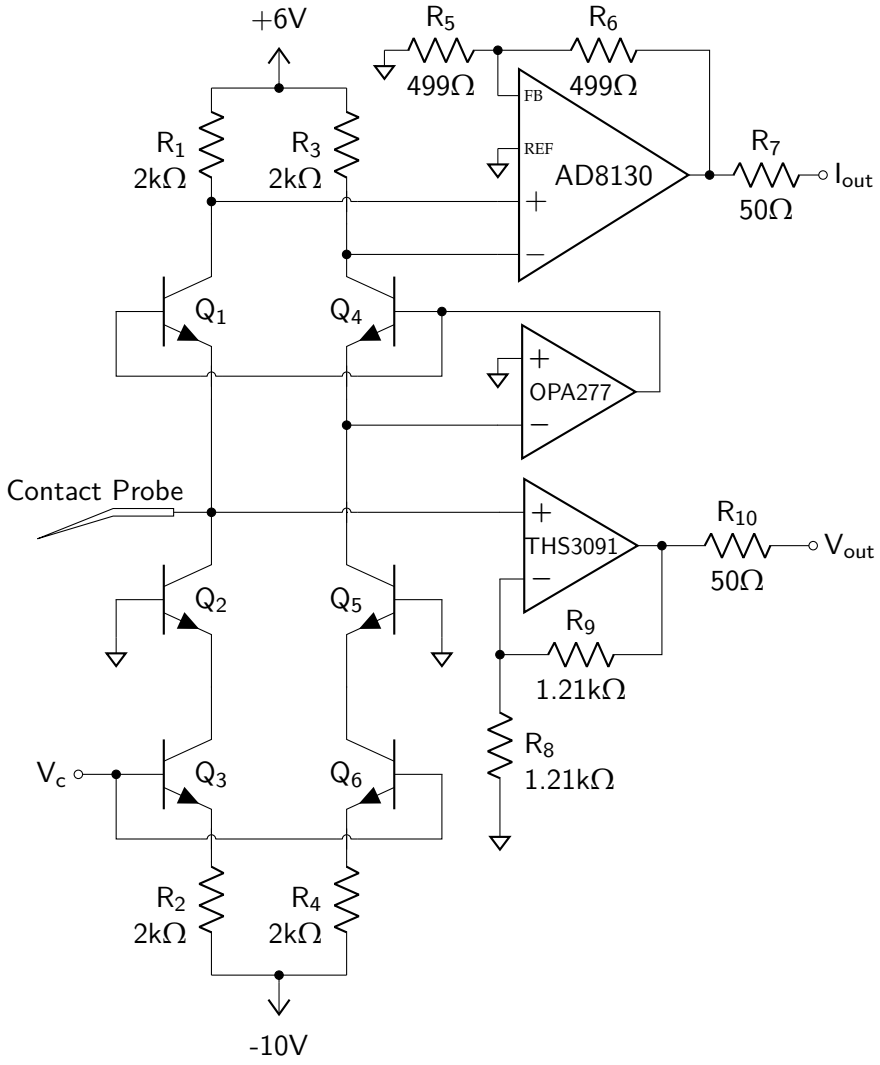

Figure 2.13: Full schematic for the current limiting probing circuit. All BJT devices are ON Semiconductor part no. NSVF5501SKT3G. Transistors Q1-Q3 perform the current limiting function, with the current limit controlled by the input signal V_c. Regulated power supplies providing ± 10 V and $+6$ V are not shown.

2.2.3 Measurement demonstrations

Overshoot characterization

Current overshoots accompanying sudden negative resistance transitions are suppressed in this measurement scheme by minimizing the capacitance at the input node of the CLA. This is done by careful selection of the input transistors and by avoiding proximity of input traces to the ground plane. However, the parasitic capacitance cannot be fully eliminated and the potential to create overshoots inevitably remains. Since overshoot transients tend to play a critical role in switching behavior in practice, it is important for them to be characterized and modeled.

In general, the time-dependent (I, V) trajectory of a current overshoot is not solely a characteristic of the measurement setup, but is determined by the coupled dynamics of the DUT conductance and the driving circuitry. The duration and amplitude of the overshoot therefore depends on the type and history of the RS cell being measured, and is not easily reproducible. To measure the overshoot characteristic in a standardized way, a test sample designed to imitate the resistive switching action was constructed using surface mount components. A mechanical reed relay in series with a 1.2 kΩ and in parallel with 100 kΩ was found to be well suited for this purpose, providing a controllable sub-nanosecond transition between two discrete resistance levels with negligible parasitic effects.

With the reed switch connected in the position of the DUT and biased by 1 V, the current transient following a resistance transition was measured across a 50 Ω scope shunt with 350 MHz bandwidth (Fig. 2.14). Close agreement of the transient was found with the solution of a differential equation describing the charging of the CLA input node,

$$C_\mathrm{p}\frac{dV_\mathrm{d}}{dt} = I_\mathrm{bias}\left[1 - \exp\left(\frac{V_\mathrm{d} - V_\mathrm{a}}{V_\mathrm{T}}\right)\right] - \frac{V_\mathrm{d}}{R_\mathrm{d}}, \tag{2.5}$$

where $C_\mathrm{p} = 5.7$ pF is the parasitic capacitance at the input, V_d is the DUT voltage drop, V_a is the applied voltage, and R_d is the DUT resistance (here assumed a step function in time). Note that C_p includes the self-capacitance of the measured cell, which is approximately 0.5 pF for the reed relay circuit. This should be taken into consideration in the memory cell design itself, where thin dielectric layers and large contact pads or device areas can contribute significantly to the total C_p, which intrinsically degrades the overshoot performance. Given the single parameter C_p, the simple model of Eq. 2.5 is expected to accurately characterize the transient response of the CLA circuit, and should be incorporated with a physical device model to properly model the complete coupled system during a measurement.

For comparison, the current overshoot transient induced using a modern SMU was measured under identical conditions. For the first 1 µs after the resistance transition, the transient begins with the discharge of a 1 m coaxial cable that was used to connect the instrument. Between 1 to 10 µs, a proprietary

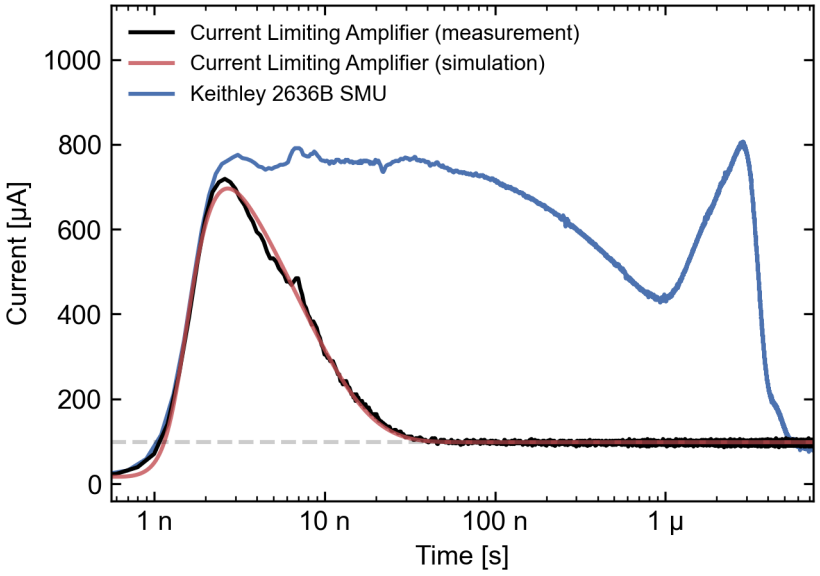

Figure 2.14: Current overshoot characterization using a reed relay to abruptly switch from 100 kΩ to 1.2 kΩ at time 1 ns with 1 V applied and with a current limit of 100 μA. Under these conditions, the CLA returned to the current limit in ∼20 ns, whereas a commercial SMU produced a more complex overshoot response lasting several microseconds.

feedback circuit is engaged and produces a long unpredictable current excursion before undershooting and eventually settling to the programmed current compliance level. Relative to this, the overshoot duration is reduced in the CLA measurement by over two orders of magnitude.

As a separate demonstration, we compare the CLA and SMU overshoots during the electroforming operation using a ReRAM device with material stack (bottom to top) Pt(20 nm)/TaO(20 nm)/TaOx(5 nm)/Pt(20 nm) and lateral dimensions 100 nm × 100 nm. The forming operation was used because of the similar starting device states and because the device is particularly vulnerable to overshoots during forming, due to the high voltage initial condition and the rapid runaway mechanism. The final state that a formed device arrives in is correlated with the overshoot transient, and in particular the resistance level and the first RESET characteristic are affected [105].

The devices were formed with the CLA and SMU with current limit set to 300 μA, using voltage sweeps on a similar timescale, 0.1 s and 1 s respectively. The current transients during forming event were each sampled using an oscilloscope with sample rate 625 MHz. The results, plotted in Fig. 2.15, show overshoot transients similar in character to those recorded for the reed relay switching. The SMU overshoot had over 100× longer duration, reached over 2× higher current levels, and resulted in a dramatically different formed state (∼ 1.5 kΩ) compared to the CLA-formed device (∼ 3.2 kΩ). The SMU-formed cell, which was subjected to a current of 5 mA during the forming overshoot, required a large RESET current above 2 mA. The CLA-formed cell required a

much lower RESET current, approximately equal to the current limit setting. In the CLA measurement, the overshoot lasted under 10 ns, and yet was still responsible for the entirety of the resistive switching during forming. That is, the entire switching trajectory from the initial (pristine) to the final (formed) state occured above the current limit, with the excess current coming directly from the ∼6 pF parasitic capacitance.

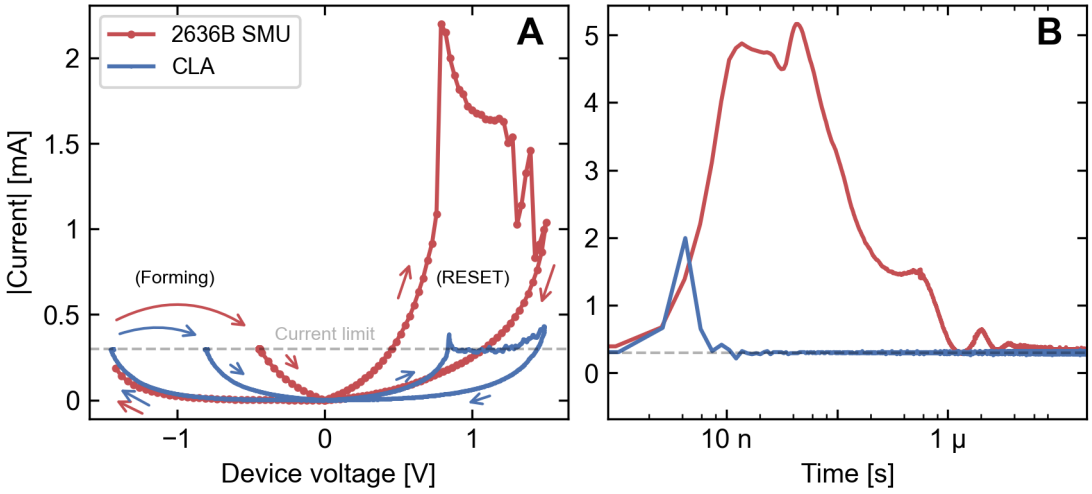

Figure 2.15: A comparison of a ReRAM forming operation using the CLA circuit and an SMU with current compliance. **(A)** shows the measured quasistatic (I, V) loops for the initial forming step (invalid transition datapoints removed) followed by a first RESET, and **(B)** shows the current vs. time during the forming events. The SMU controlled overshoot lasted over 100× longer and had over 2× higher amplitude than the CLA overshoot, with corresponding consequences for the formed states.

High frequency ReRAM cycling

To demonstrate the basic RS cycling operation using the external CLA circuit, we tested the ReRAM device with the CLA input connected to the top (active) electrode. The current limit was set to 300 µA and a triangular voltage signal with period 10 µs and amplitude 1.5 V was applied to the DUT bottom (ohmic) electrode using a Rigol DG5102 AWG. The applied voltage and the amplified current signal were sampled at 1.25 GS/s using a Picoscope 6403D deep-storage oscilloscope. In a single measurement lasting only one second, 10^5 full (I, V) loops were successfully collected, each containing 1,564 8-bit (I, V) samples (Fig. 2.16). It is furthermore possible to collect millions of such cycles in a practical amount of time by collating multiple measurement shots, creating powerful datasets for statistical evaluation of RS devices.

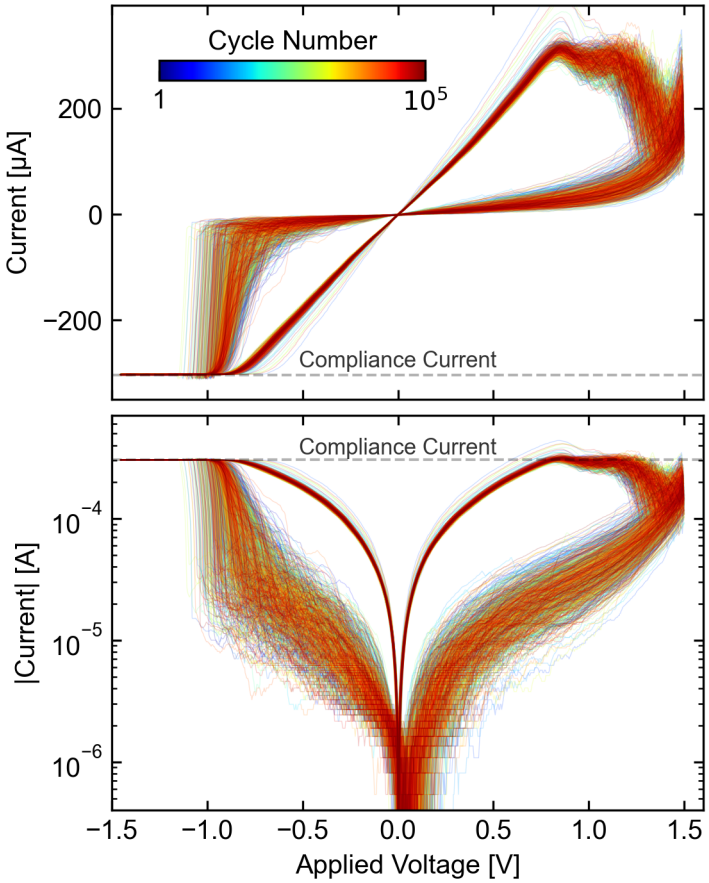

Figure 2.16: A measurement of 10^5 consecutive (I, V) loops collected in one second with the CLA circuit using a triangular voltage excitation and a 300 μA current limit. Data is smoothed by a 15 sample moving average, and every 100th cycle is plotted on a linear current scale (top) and log scale (bottom). To conform to plotting convention with respect to polarity, the applied voltage is defined as the negative of the AWG voltage.

Hybrid sourcing

Beyond the voltage-controlled cycling scheme just demonstrated, there are additional operating modes supported by the CLA circuit that are also useful for characterization of devices with NDR in their switching characteristics.

It is quite common for bipolar ReRAMs to exhibit S-type NDR for the SET polarity as well as N-type NDR for the RESET polarity. These different kinds of NDR are also referred to as current controlled and voltage controlled types, respectively, as these sourcing modes can support single-valued stable states in each case. In practice, a compromise can be made by driving both SET and RESET transitions through a series resistance. However, measurement circuits that put a large resistance in series with the DUT are widely acknowledged to cause an unnecessary acceleration of induced RESET processes due to positive feedback, which has been referred to as "the voltage divider effect" [111, 114]. Because of this, a hypothetical attempt to drive a RESET process using

a current source, which has an infinite Thévenin equivalent series resistance, would tend to cause a runaway RESET and result in a poorly controlled transition and/or destruction of the device [170]. Accordingly, current sourcing during a RESET operation is never done in practice. On the other hand, using a voltage source to drive a SET process is a directly analogous scenario but is a completely standard procedure.

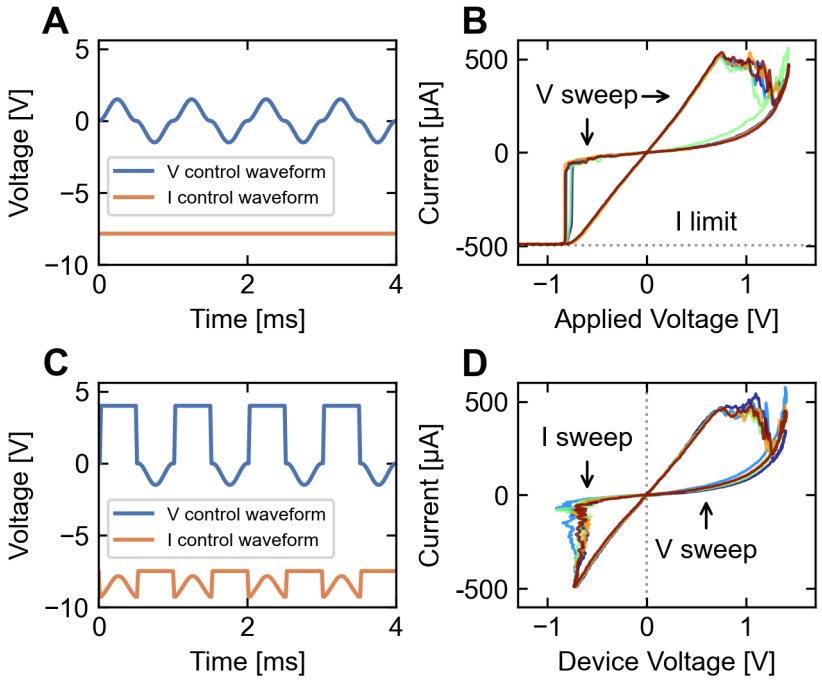

Figure 2.17: Control waveforms and the resulting (I, V) measurements of a 100 nm ReRAM cell in two different schemes. In a more conventional scheme, **(A)** shows a periodic signal controlling the device voltage while the current limit is set to a DC level (500 μA). This results in the device electrodes seeing the full control voltage waveform up until the current limit, producing rapid runaway SET transitions **(B)**. In a new scheme for collection hybrid (I, V) loops, synchronized I and V control waveforms **(C)** avoid runaway transitions for both SET and RESET, allowing gradual control of both resistance transitions **(D)**.

While practical details may justify the eventual use of direct voltage control in certain applications, much understanding can nevertheless be gained by using alternate characterizations. An obvious objective is to control each transition in the most gradual way possible while collecting detailed information each cycle. In light of the dual-NDR nature of ReRAM, an advantage of the CLA design is that by using suitably synchronized AWG control signals, a seamless transition between the two different sourcing modes can be controlled at high speed. This enables continuous collection of what we will call "hybrid" (I, V) loops, which are demonstrated in Fig. 2.17.

Experimental SET transitions measured under external voltage control are commonly observed to occur suddenly, taking place over a short duration that

can be well below 1 ns [68]. Unless overdriven, these rapid transitions occur only after a much longer stochastic delay [171]. This behavior, which makes the SET transition difficult to predict and to control, is particular to the kind of sourcing used. By instead sweeping through the SET transition with a low capacitance current source, we demonstrate that intermediate states can be stable over relatively long durations (Fig. 2.18), which allows a continuum of resistance states to be more reliably programmed in both polarities.

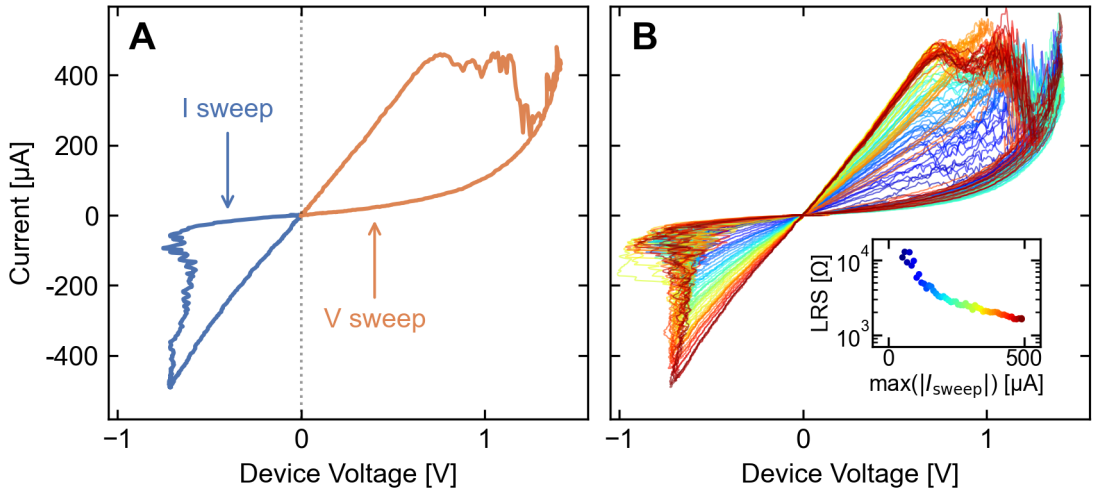

Figure 2.18: Demonstration of gradual SET switching induced by hybrid-sourced (I, V) loops, continuously collected on a timescale of 1 ms per cycle. A single loop is shown in **(A)** with an identification of the source types used for each polarity. Superposed sweeps using progressively higher maximum current levels is shown in **(B)**, with an inset displaying the resultant LRS values after each sweep.

Cycling with *in-situ* overshoot sampling

Even a carefully minimized parasitic capacitance still has the potential to play a dominant role in switching processes, or at least to have important effects on the tails of the switching parameter distributions. RF measurements have shown SET switching using energies of only ∼100 fJ [65], while a capacitance of just 1 pF discharged from 1.5 V to 1 V during a SET transition releases an additional $C_p(\Delta V_d)^2/2 = 625$ fJ into the cell. Despite this situation, measurements are almost never set up to record overshoot transients as part of a current-limited ReRAM cycling procedure. Several one-off measurements have shown current vs. time traces for individual switching events [91, 94, 103, 119], but the switching trajectories (actual device current *and* device voltage vs. time) during overshoot are not shown, and statistically significant measurements have not yet been reported.

Measuring the full device (I_d, V_d, t) trajectories with high bandwidth during switching is a direct way to quantify and to study the impact of overshoot transients, and more generally to inform the design of control systems

for ReRAM. The ability to record these trajectories while also using the CLA limiting function benefits from reconfiguring the CLA as a source-side current limiter. In the original configuration, the CLA and AWG are positioned on opposite sides of the DUT, and currents from the parasitic capacitance are not amplified because they flow outside of the measurement path. Positioning both the CLA and AWG on the same side of the DUT opens up the counter electrode for a separate current measurement over a terminated, ground referenced transmission line that can detect the DUT current at the full scope bandwidth (350 MHz). This small modification is realized by using the AWG voltage source channel to drive the non-inverting input of the OPA277 instead of the DUT counter electrode, and is also advantageous because V_{out} gives a more direct reading of the device voltage than in the original configuration.

Full switching trajectory measurements are demonstrated for a forming event in Fig. 2.19 and for an individual switching cycle in Fig. 2.20. Each of these reveal aspects of the switching process that are often misrepresented by conventional measurements. These examples also highlight that, even with a minimized parasitic capacitance, scenarios exist where the majority of SET switching occurs during a short nanosecond-scale transient that vastly exceeds the CLA current bias. These overshoots should not be interpreted as a failure of the control circuitry, but rather as an inevitability of external measurements of RS that should not escape detection. Relative to commercial solutions currently in widespread (mis)use, this configuration not only improves the control situation by orders of magnitude, but also makes the remaining non-idealities visible. This provides new opportunities to evaluate a larger class of device designs that would not necessarily survive switching under SMU control, while monitoring the actual voltages and currents seen by the device at high speed during the entire process. Therefore, the implications of control non-idealities can be studied directly, for essentially the first time.

With this setup, it is possible to measure current limited $(I_{\mathrm{d}}, V_{\mathrm{d}}, t)$ loops at extremely high speed. In Fig. 2.21 we demonstrate 1 MHz and 10 MHz cycling rates, where up to 10^6 full loops were captured in a single continuous 100 ms measurement shot. To our knowledge, this is a faster cycling rate than any sweeping measurement previously reported, even those using integrated FETs [100, 101]. Interestingly, highly uniform switching characteristics were measured at the short 100 ns per cycle timescale, which could be due to the overdrive used in the measurement but could also indicate that faster cycling is beneficial for switching uniformity.

The huge amount of cycling data produced in a short time frame enables many new experimental possibilities for device evaluation in a statistical context and at a timescale relevant for applications. Parameters such as switching thresholds and resistance levels can be extracted from each of the full $(I_{\mathrm{d}}, V_{\mathrm{d}}, t)$ loops, as shown in Fig. 2.22. Here, R_{HRS} is the resistance of the HRS, V_{set} is the SET threshold voltage, I_{max} is the maximum recorded current in the SET transient, R_{LRS} is the resistance of the LRS, and I_{reset} is the current at the beginning of the RESET. Due to large number of samples, the overall distributions of these parameters are clearly resolved, and are seen in this example to be

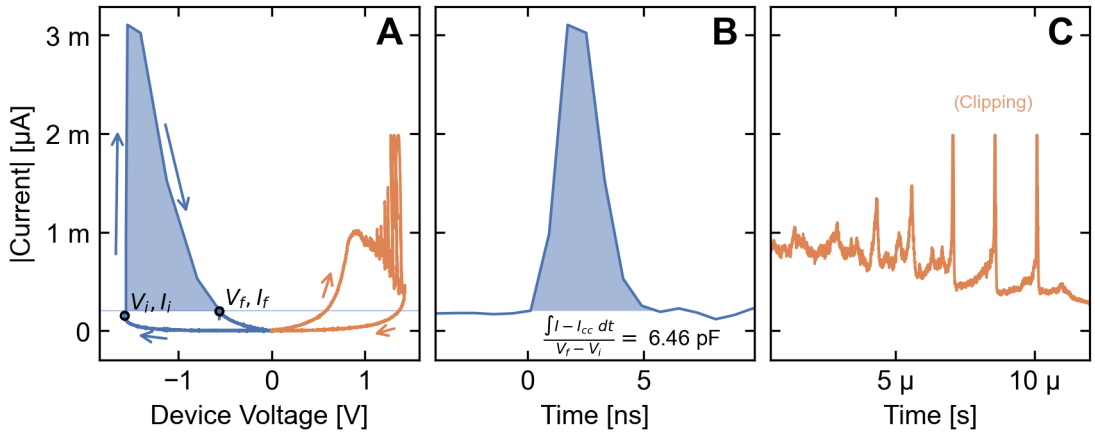

Figure 2.19: Electroforming and the following RESET process induced with a 100 μs triangular voltage sweep and a 200 μA CLA current bias. Device voltage and device current were each continuously sampled at 1.25 GHz. **(A)** Overshoot appears disproportionately large when plotted on the (I, V)-plane, but lasts only 5 ns or 0.005% of the total sweep duration. **(B)** The total area under the I vs. t curve during overshoot is consistent with the model of Eq. 2.5. **(C)** Spiking features are resolved during the first RESET which would not be visible in a standard quasistatic (DC) measurement. The spikes contain meaningful information, and could hypothetically be caused by a successive rupturing of a multi-filamentary formed state.

asymmetric and non-gaussian. Viewing the data on such a scale reveals that it is far from independently and identically distributed (i.i.d.) and is highly auto- and cross-correlated. Scatter plots of the switching parameters vs. past values (Fig. 2.23) reveal interesting inter-dependencies. In this particular dataset, the HRS is only very weakly correlated with with values of past and future parameters, while the other parameters show strong non-linear dependencies. One chain of correlations here that can be reasonably interpreted as causal is a high switching voltage → a high overshoot current → a low LRS → a high reset current. A statistical dataset of this type is analyzed further and used as input to a new device model described in Chapter 3.

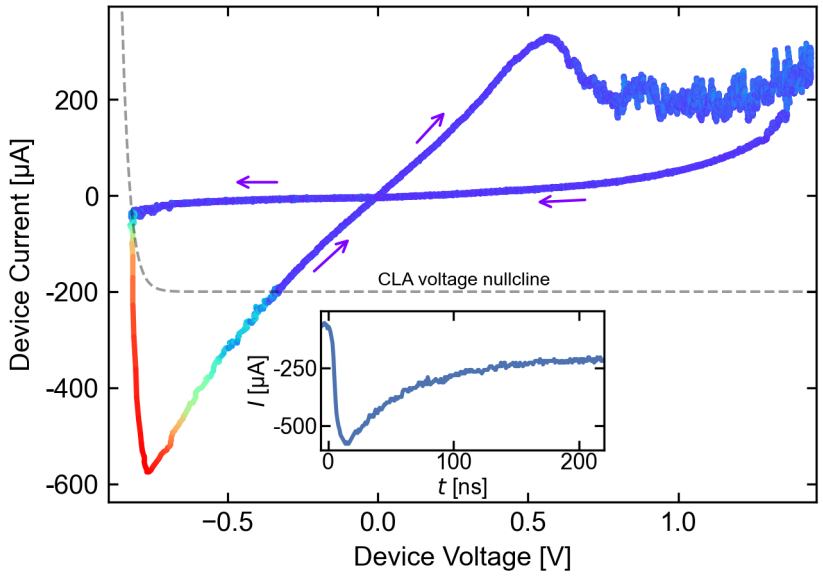

Figure 2.20: A single $(I_\mathrm{d}, V_\mathrm{d}, t)$ switching cycle displaying a particularly large overshoot, measured with a 1 ms triangular voltage sweep and a 200 µA CLA current bias. Color is a measure of the cartesian distance between consecutive datapoints on the plot, giving an indication of the speed along the measured trajectory. The action of the control circuit during overshoot is visualized by the voltage null-cline (load characteristic of the CLA). The inset shows the time dependence of the current during overshoot.

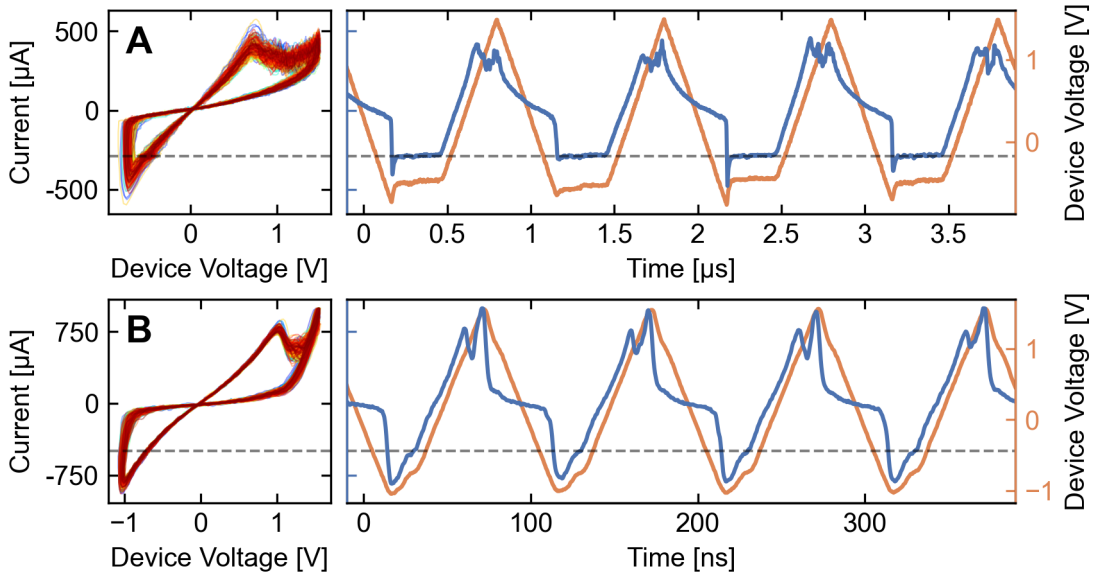

Figure 2.21: Demonstrations of continuous, mass collection of $(I_{\mathrm{d}}, V_{\mathrm{d}}, t)$ loops with CLA current control. **(A)** 10^5 loops captured in 100 ms (1 μs per cycle) with $I_{\mathrm{bias}} = 300$ μA and **(B)** 10^6 loops captured in 100 ms (100 ns per cycle) with $I_{\mathrm{bias}} = 500$ μA. The left frames show the projection of 300 consecutive loops on the $(I_{\mathrm{d}}, V_{\mathrm{d}})$-plane, and the right show the time dependence of the I_{d} and V_{d} signals. These traces show that even for cycles with period 100 ns, the current limiting function is still operational and acts to control the amount of resistive switching. The role of the short overshoot transients above I_{bias} is also clarified.

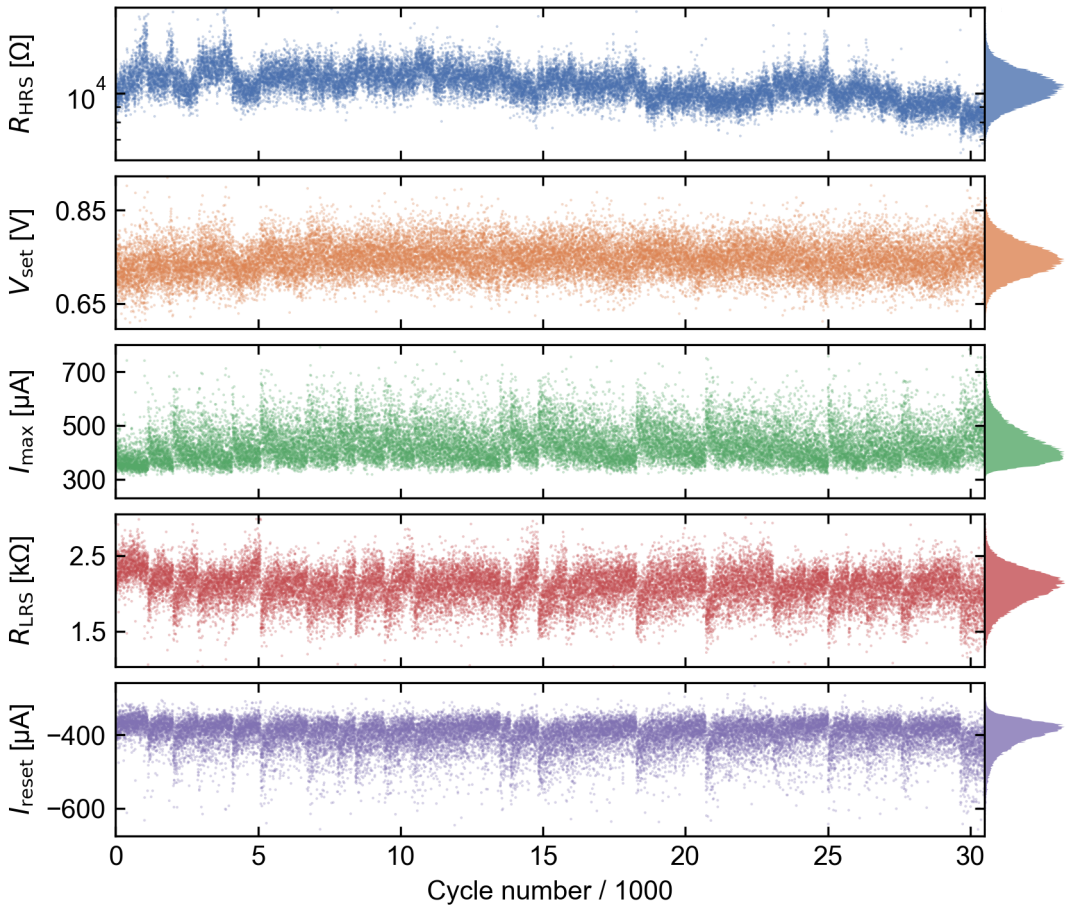

Figure 2.22: Switching parameter time series extracted from a segment of data (30k loops) from the measurement shown in 2.21(A). Each datapoint corresponds to a full (I_d, V_d, t) loop. Total histograms are shown on the right of each frame. The data are strongly auto- and cross-correlated and are asymmetrically distributed.

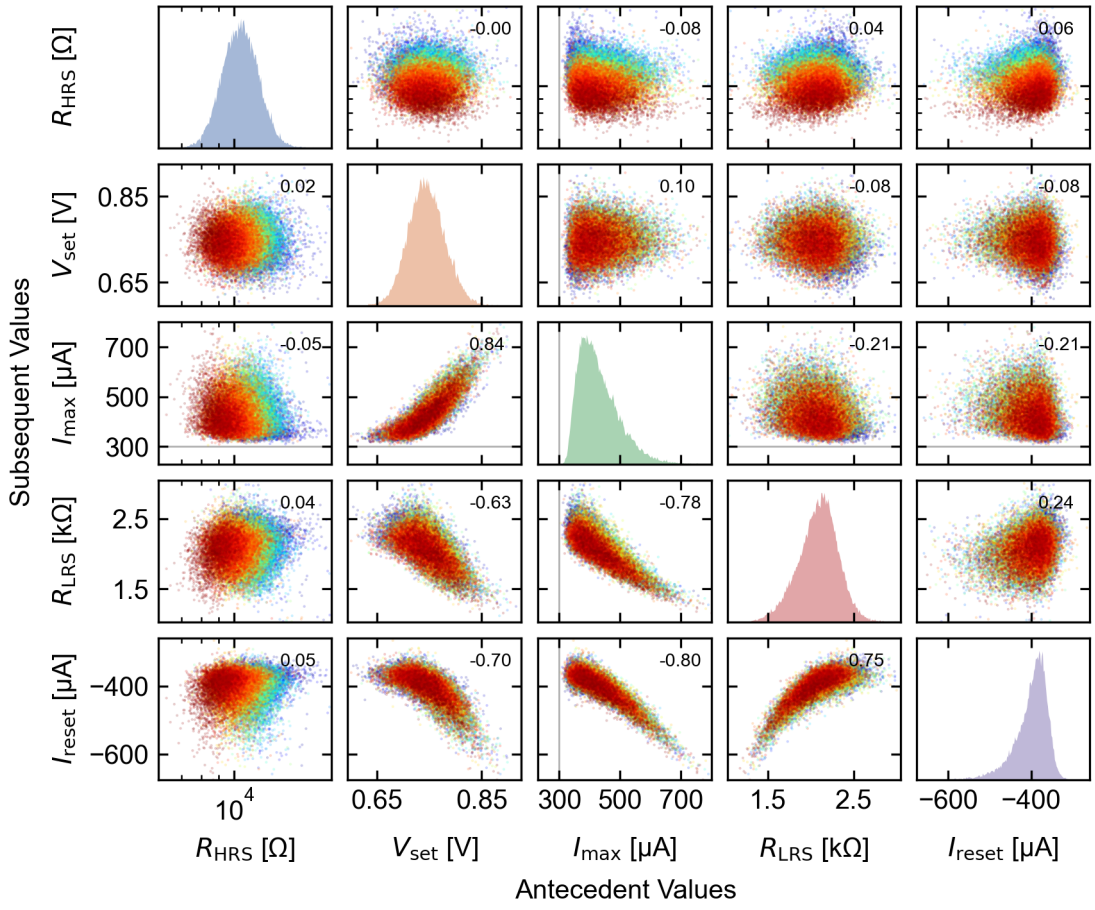

Figure 2.23: Scatter plots (off-diagonal subplots) and histograms (diagonal subplots) of switching parameters from a segment of data (30k loops) from the measurement shown in 2.21(A). Each parameter is compared with future and past values of each other parameter. The Spearman correlation coefficients for each parameter pair are marked on the top right of each subplot.

2.3 Discussion

In this chapter, two custom circuit designs were introduced that represent a significant step forward for external measurement of RS devices, relative to commercially available solutions. Both circuits function to stabilize or interrupt resistance transitions by a variable and programmable amount, and they both allow measurements at high speed using arbitrary waveforms. They provide the current-voltage feedback necessary to characterize NDR devices, while giving direct, high-speed measurements of both the device voltage and current vs. time, the significance of which has been commonly overlooked. Each circuit is deliberately designed to have a well understood electrical response, which is a basic requirement if one hopes to characterize or accurately simulate switching processes. Although there are certain overlapping applications, the two circuits are not redundant; they each provide different types of control.

The main advantage of a measurement using the digipot circuit is simplicity. A linear series resistance can be freely chosen within a certain range, there is only one AWG channel to control, and no non-linear calibration is involved. It should be used in cases when the impact of a linear load is of specific interest, or where the linearity simplifies data analysis. Besides some intrinsic downsides of measuring with an external series resistor, a limitation of the digipot circuit is that swapping the resistor value takes a certain amount of time ($\sim$100 µs), so measurements that require series resistance alterations need to incorporate this delay (though simple bipolar SET/RESET applications can use a diode in parallel with the digipot resistance [172]). Another limitation is on the bandwidth of voltage signal application to the device, as the digipot transfer function is well approximated as a first order RC filter. However, the bandwidth is above 10 MHz for a typical (lightly loaded) measurement, and in worst-case scenario still approaches 1 MHz, which is usually adequate. To maximize the bandwidth, one can choose the minimum amount of loading that will permit the intended measurement, set for example by the maximum expected value of NDR in the device at the frequency in question. Note that this filtering effect does not impact the bandwidth of the current measurement, which is limited instead by the 50 Ω oscilloscope channel (350 MHz here).

On the other hand, the action of the CLA is comparable to a standard lab bench power supply with its separate current and voltage dials (which represent limiting values), but with the dials replaced by high speed arbitrary waveforms. To accomplish this hybrid voltage-current sourcing function continuously at high speed and with very low parasitic capacitance, we use the (I, V) characteristic of a single BJT device in CB configuration. This effectively presents a non-linear load that transitions from very low resistance to very high resistance at the desired current level. The transition is not infinitely sharp, but is limited by the thermal voltage V_T and is the sharpest possible for a semiconductor device. This highly stable configuration lets us apply voltages to the DUT with much higher bandwidth than when using an external series resistor or FET device. The on-board current amplifier on the CLA achieves

a higher signal to noise ratio (SNR) than a 50 Ω shunt measurement, but as a trade-off the current measurement bandwidth of the specified setup is limited to around 10 MHz. There is room to improve this bandwidth by using alternate amplifying components, or by separately measuring the current on the opposite side of the device, as also demonstrated in this chapter.

Chapter 3

Fast stochastic modeling of synaptic arrays

Traditionally, electronic device modeling begins with a physical description of the materials and processes involved. In the case of ReRAM, the physical situation is immensely complicated with many degrees of freedom, and accurate modeling is a wide-scale and ongoing research undertaking. Efforts in this direction are motivated by advancing an understanding of physical and chemical dependencies that can in principle inform design choices on physically justified grounds. In the past decade, many different computational techniques have been employed to furnish device models, from ab initio density-functional theory (DFT), molecular dynamics (MD), kinetic Monte Carlo (KMC), finite element method (FEM), as well as ordinary differential equation (ODE) and differential algebraic equation (DAE) solvers [174–178]. The resulting models exist on a spectrum of physical abstraction, such that the cost of increasing computational speed is generally a trade-off in physical accuracy and detail [179].

Device models that naturally encompass stochasticity do so at the cost of complexity needed to compute the physical scenario in high detail. For example, atomistic KMC simulates switching processes with atomic precision and is inherently stochastic but requires hours of computation per cycle even for small individual cell volumes (e.g., $5\times5\times5$ nm^3 [180]). At the other end of the spectrum, dynamic models based on numerical solutions of ODE/DAE systems are designed to run significantly faster while sometimes aiming to remain physically realistic. However, their higher speed invariably comes at the cost of approximations, simplifications, and omissions of physical reality. Typically, device operation is distilled to a dynamical description of one or two state variables, such as a conducting filament length, radius, or a defect concentration.

Due in part to ambiguity in their high dimensional parameter space, a given ODE/DAE model encompasses a diverse range of possible cell behaviors and has the flexibility to approximately match measurement data [181,

182]. However, fitting the model to data is commonly an ad-hoc, manual, and/or unspecified procedure. Having dispensed with the atomistic sources of variability, ODE/DAE models are fully deterministic by default. Where stochasticity is required, it is accounted for by injecting noise into the state variables or parameters of the model [183–185]. Due to the unique experimental challenges posed by electrical measurement of ReRAM, the data used for fitting is not necessarily statistically sufficient nor measured under relevant electrical conditions and timescales. While these models can be tuned by hand to roughly match the dispersion observed in a measurement [186, 187], they generally fail to accurately reproduce the complex statistical properties of actual devices.

The main purpose of ODE/DAE device models is to be computationally efficient enough to support circuit simulation. Still, nonlinear solvers require many finely spaced timesteps and a considerable amount of total time to compute dynamical trajectories. Although they have been successfully used to demonstrate small scale circuitry such as logic elements and small crossbar arrays [188–191], benchmarks or indications of run time for ODE/DAE-based simulations have so far not been supplied. Except for extremely small ML model sizes on the order of 10^3 weights or below, demonstrations of network performance are expected to remain computationally intractable via conventional circuit simulation.

In this chapter, we address these device modeling challenges with a new type of generative model for arrays of artificial synapses. The main objective of the model is to accurately reproduce the statistical properties of fabricated devices while remaining computationally lightweight. Starting with newly available electrical measurement data as an input, this phenomenological model is systematically fit using a well defined statistical regression analysis. The exclusive use of easily computable analytical expressions provides close quantitative agreement with relevant experimental observation. Taking advantage of parallel resources on a modern CPU and GPU, we demonstrate the ability to simulate hundreds of millions of synaptic connections with over 10^8 weight updates per second. With its high throughput and low memory footprint, the model can be usefully employed to simulate large arrays of solid-state synapses for investigation of emerging NC concepts on a large scale.

3.1 Methods

In designing a stochastic model for synaptic arrays, we place high priority on speed and fitting accuracy. One of the beginning assumptions is that in every possible device state, the device current (I) can be represented by a linear mixture of two fixed polynomials in the applied voltage (V). These two polynomials, which are each estimated from a fit to measurement data, can be thought of as limiting cases for the highest possible high resistance state, $I_{\mathrm{HHRS}}(V)$, and lowest possible low resistance state, $I_{\mathrm{LLRS}}(V)$. The device current in all

possible resistance states is then given by

$$I(r, V) = rI_{\mathrm{HHRS}}(V) + (1 - r)I_{\mathrm{LLRS}}(V), \tag{3.1}$$

conveniently reducing the description of the conduction in the material to a single state variable $0 < r < 1$. This set of functions can be efficiently evaluated by Horner's algorithm and serve as a close enough approximation to the true non-linear conduction behavior for our purposes.

In ReRAM, the overall resistance state as well as the transition behavior is affected by a vast number of different possible configurations of ionic defects in the material, giving rise to the observed stochastic behavior and history dependence. Rather than attempting to describe the ionic transport physically, we turn instead to measurement data to directly provide the necessary statistical information. A discrete multivariate stochastic process based on a structural vector autoregression (SVAR) model is fit to the data and used to generate latent variables that guide the state evolution of simulated memory cells. As a cell is exposed to voltage signals, new terms of the SVAR model are realized by a sum of easily computable linear transformations of past states and pseudorandom vectors.

As an overview, the experimental and simulation approach that will be elaborated in this section can be shortly summarized as follows:

1. A fabricated ReRAM cell is experimentally driven through a large number of resistance cycles by applying a continuous periodic voltage signal while measuring the resulting current.
2. A time series of feature vectors, $\boldsymbol{x}_n$, composed of resistance values and switching threshold voltages, is extracted from each of the measured cycles.
3. A discrete stochastic process, $\boldsymbol{x}_n^*$, is constructed to enable generation of simulated feature vectors that reproduce the measured distributions as well as the long range correlation structure of $\boldsymbol{x}_n$.
4. An array of simulated cells are instantiated according to independent realizations of $\boldsymbol{x}_n^*$ to represent cycle-to-cycle variations, together with a random scaling vector $\boldsymbol{s}_m$ to represent device-to-device variations.
5. Two programming methods are exposed for each cell; one to apply voltages and another to make realistic current readouts. Applied voltages above the generated thresholds alter the device state, following an empirical structure which encodes the resistance transition behavior and allows access to a range of resistance states. Each voltage driven resistance cycle triggers the generation of new stochastic terms from $\boldsymbol{x}_n^*$, which govern the progression to future states.

3.1.1 Data collection

For the purposes of stochastic modeling, electrical measurement data is needed that capture relevant information about the internal state of a memory cell and its variation cycle-to-cycle (CtC) and device-to-device (DtD). However, ReRAM measurements performed at operational speed typically make exclusive use of rectangular voltage pulse sequences, which yield very little useful state information. On the other hand, measurements applying continuously swept voltage signals while sampling the resulting current are more suitable because much more information is collected each cycle, such as switching threshold voltages, current-voltage nonlinearity, resistance states, and transition behavior.

Conventionally, measurements employing voltage sweeps are carried out using the source measure units (SMUs) of commercial semiconductor parameter analyzers (SPAs). However, SMUs make heavy use of averaging to measure noisy signals at high resolution and thus sample too slowly to collect cycling data in a meaningful quantity. Furthermore, because two-terminal switching devices are prone to electrical instability and runaway transitions, voltage sweeping measurements usually require integrated current limiting transistors to avoid destruction or rapid degradation of the cell. This presents a significant fabrication overhead and limits the materials available for study. In light of these challenges, the input data for the present stochastic model was acquired using a custom measurement technique, introduced in detail in Section 2.2. The setup uses an external current-limiting amplifier circuit to allow for collection of sweeping measurements at over six orders of magnitude higher speeds than SMUs, while also eliminating the cumbersome requirement of on-chip current limiting.

The ReRAM cell used for measurement of cycling statistics was integrated in the back end of line of a 130 nm CMOS process, between M4 and M5 aluminum metal lines (Fig. 3.1). On M4, a damascene TiN via followed by a patterned TiN bottom electrode were processed, forming the inert electrode of the device. The memory stack was then deposited. First, 10 nm HfO_2 deposited by atomic layer deposition (using $HfCl_4$ and H_2O precursors) acts as the resistive switching layer [64]. Then, a 20 nm Ti scavenging layer was deposited by physical vapor deposition, allowing creation of oxygen vacancies within the HfO_2 during the memory operation. A 100 nm TiN top layer was used to cap the device. Deep ultraviolet photolithography and dry etching were used to pattern the memory dot, defining the active area. A SiN capping layer was used to isolate the memory from adjacent cells. Top vias were then opened by photolithography and dry etching in order to contact the memory dots. Finally, aluminum M5 was deposited and patterned to complete the process flow.

The measured device was electrically isolated with contact pads leading directly to the top and bottom device electrodes, with no access transistor or added series resistance. Using a fixed 100 μA current limit in the SET polarity, the pristine cell was electroformed by application of 100 μs duration triangular pulses with incrementally increasing amplitude until a current jump was

recorded near 3 V. For all subsequent cycling, a 1.5 V amplitude 10 kHz triangular waveform was applied. The cell was first exercised for 2.4×10^6 cycles before 10^6 additional cycles were collected for analysis. Current (I) and voltage (V) waveforms were simultaneously recorded with 8-bit resolution and with a sample rate of 1,042 samples per cycle. The measured current array was smoothed with a moving average filter to improve the quality of the raw data before further analysis. An adaptive rectangular window size was used to preserve current steps in the signal, with the maximum window size of 25 samples gradually reducing to a minimum of 3 samples at the pre-detected locations of SET transitions of each cycle. After smoothing, the contiguous I and V waveforms were split into indexable cycles at most positive value of the periodic applied voltage (see Fig. 3.2[A]).

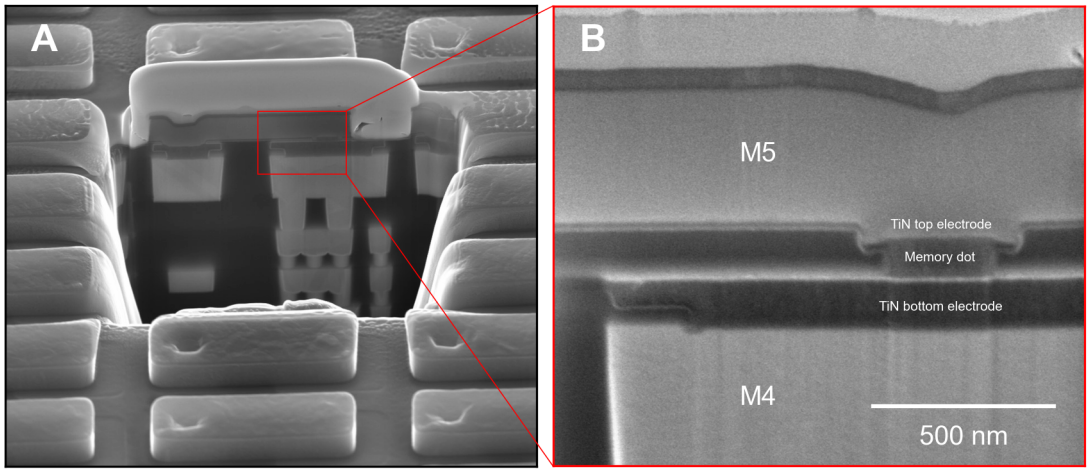

Figure 3.1: Scanning electron micrographs of the ReRAM cell design used for electrical measurement. **(A)** shows a cross-section of the cell, and **(B)** shows a zoom-in of the resistive memory between metalization layers M4 and M5.

Each cycle exhibits the following temporal sequence of states and events: a high resistance state (HRS), a transition (SET) out of the HRS into the following low resistance state (LRS), and finally another transition (RESET) into the next HRS. Current vs. voltage (I, V) plots for a subset of the collected cycles are shown in Fig. 3.2(B), which highlights the significant stochastic CtC variations. The observed characteristics are typical for ReRAM subjected to voltage controlled sweeps — on average, there is relatively higher voltage non-linearity in the HRS than in the LRS, and the SET transitions are abrupt with respect to the applied voltage, while the RESET transitions proceed relatively gradually over a voltage range of approximately 700 mV.

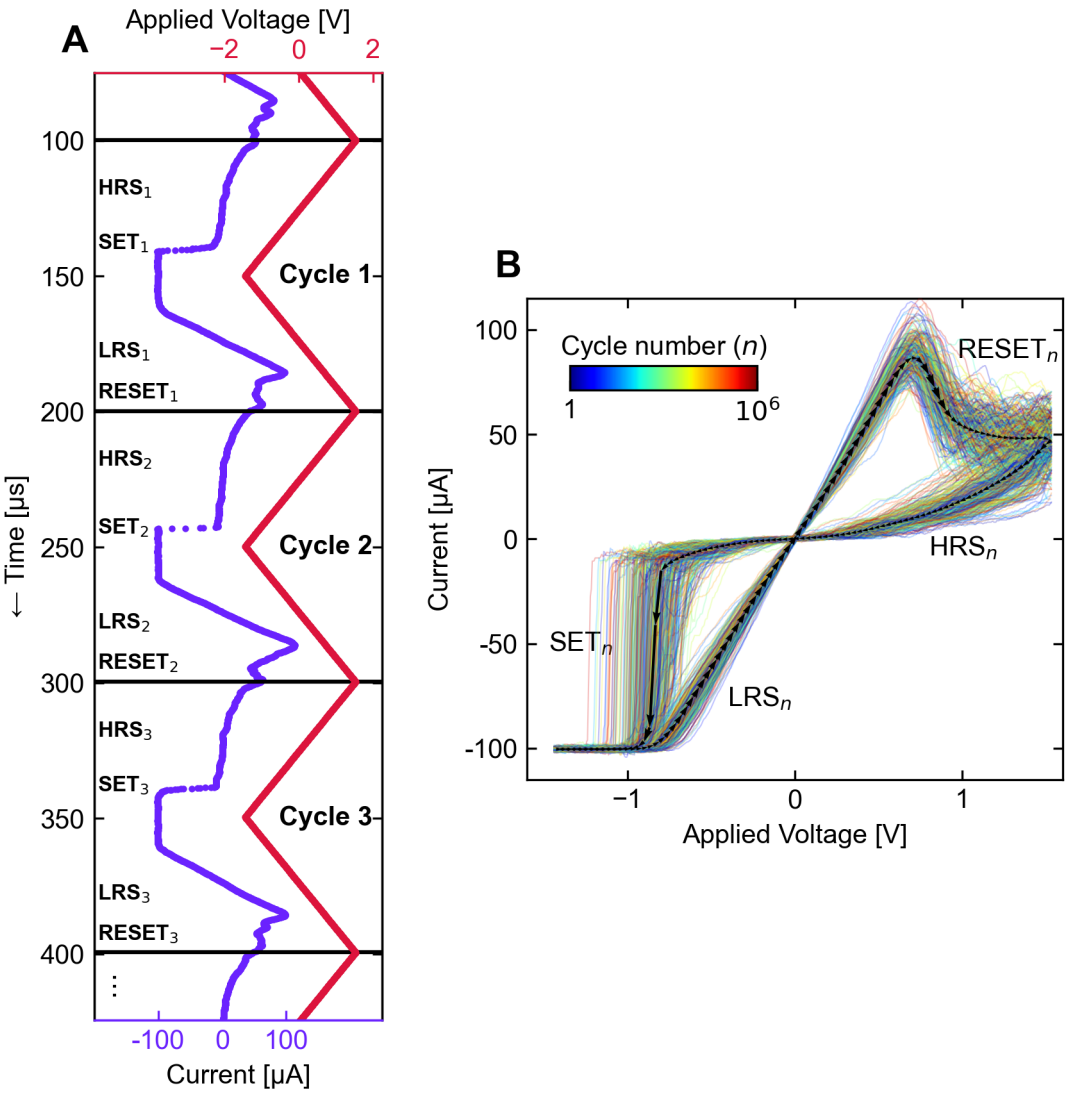

Figure 3.2: **(A)** The measured time dependence of I and V waveforms resulting from the ReRAM cycling experiment and used as input to the stochastic model. The waveforms are divided into 10^6 indexed cycles, the first three of which are shown. From this dataset, the periodic temporal sequence of the states and events of each cycle (HRS_n, SET_n, LRS_n, RESET_n) is extracted and subject to statistical modeling. **(B)** An I vs. V plot for 300 of the 10^6 measured cycles, showing significant statistical variation. The black arrowed path indicates the temporal direction of the measurement, following the average of 10^4 (I, V) curves whose SET voltage was within one percentile of the median.

3.1.2 Feature extraction

The full I, V cycling measurement just described consists of over 16 GB of numerical data and would not be practical to model on a point-by-point basis. Therefore, we aim to compress the dataset while retaining enough information such that the full (I, V) characteristics can be approximately reconstructed from the compressed representation. Accordingly, the full dataset is reduced to a vector time series of distinguishing features of each cycle. Four scalar features were chosen for extraction: the value of the HRS, $R_H[\Omega]$, the SET threshold voltage, $V_S[V]$, the value of the LRS, $R_L[\Omega]$, and the RESET voltage, $V_R[V]$. We denote the series as

$$\boldsymbol{x}_n = \begin{bmatrix} R_{H,n} \\ V_{S,n} \\ R_{L,n} \\ V_{R,n} \end{bmatrix} = \begin{bmatrix} R_H \\ V_S \\ R_L \\ V_R \end{bmatrix}_n , \tag{3.2}$$

where $n = \{1, 2, \ldots, 10^6\}$ is the set of cycle indices. The feature vector elements, whose precise definition follows, are chronologically ordered from top to bottom as they occur in the measurement dataset.

The SET voltage V_S, or the voltage where the cell resistance abruptly decreases, is extracted from each cycle as the absolute value of the linearly interpolated V corresponding to the first level crossing of $I = -50$ μA. The RESET voltage V_R, defined as the voltage where the reset process begins, is determined from the I datapoints by peak detection using simple comparison of neighboring samples. Here, only the increasing section of the voltage sweep with $V > 0$ is considered. The voltage corresponding to the first encountered peak with prominence ≥ 5 μA is taken as the RESET voltage. If no peak satisfies this criterion, the peak with maximum prominence is taken instead.

The device current for any static state is approximated in our model as a polynomial function of the applied voltage. The values of R_H and R_L are likewise extracted from least squares polynomial fits to appropriate subsets of the measured (I, V) data of each cycle. The HRS is fit with a 5th degree polynomial on the decreasing V sweep in the variable range $V_S + 0.1\ \text{V} \leq V \leq 1.5\ \text{V}$ and $-25\ \mu\text{A} \leq I \leq 80\ \mu\text{A}$, and the LRS is fit with a 3rd degree polynomial on the increasing part of the V sweep in the range $-0.7\ \text{V} \leq V \leq V_R - 0.05\ \text{V}$ and $-80\ \mu\text{A} \leq I \leq 120\ \mu\text{A}$. The fits are constrained such that the 0th degree coefficient equals 0 A, and the 1st degree coefficient is $\geq$1 nA/V. The values of R_H and R_L are then defined as the static resistance of the respective polynomials at a fixed voltage $V_0 = 200$ mV.

An overview of the result of this feature extraction is given in Fig. 3.3. The 10^6 cycles proceeded without significant long-term drift from the overall mean value,

$$\bar{x}_n = \begin{bmatrix} 166.5\ \text{k}\Omega \\ 0.85\ \text{V} \\ 8.2\ \text{k}\Omega \\ 0.72\ \text{V} \end{bmatrix}, \tag{3.3}$$

but with significant variations in each feature between cycles. A prominent characteristic of this data is that it is strongly correlated over long cycle ranges, as quantified in Fig. 3.11. The asymmetric marginal distributions for each of the features were very well resolved due to the large number of samples, and they did not accurately converge to any analytical probability density function (PDF) in common use, including the normal and log-normal.

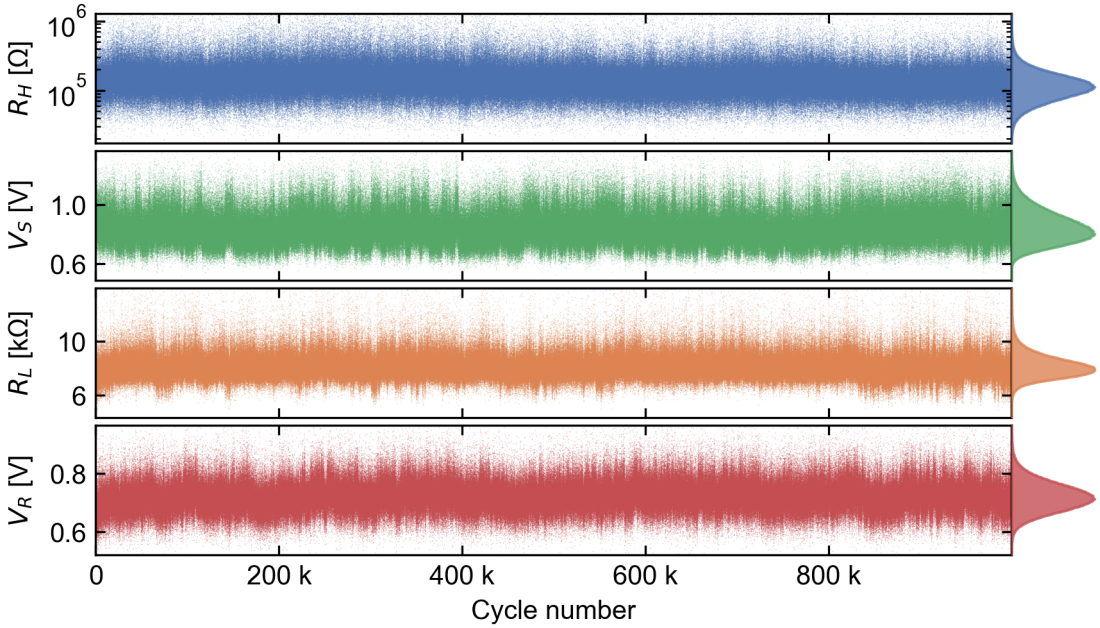

Figure 3.3: A view of the feature vector time series extracted from each of 10^6 measured (I, V) cycles. Each feature, which represents either a resistance state or a switching voltage, has its marginal histogram shown on the right.

3.1.3 Stochastic modeling

This section will introduce the statistical methods used to model the internal states of an array of synaptic ReRAM devices, including CtC and DtD variability effects. The handling of voltages applied to the cells as well as the simulation of realistic readouts of the resistance states will also be established. To help orient the reader, the overall structure of the generative model that will be described is provided in advance in Fig. 3.4.

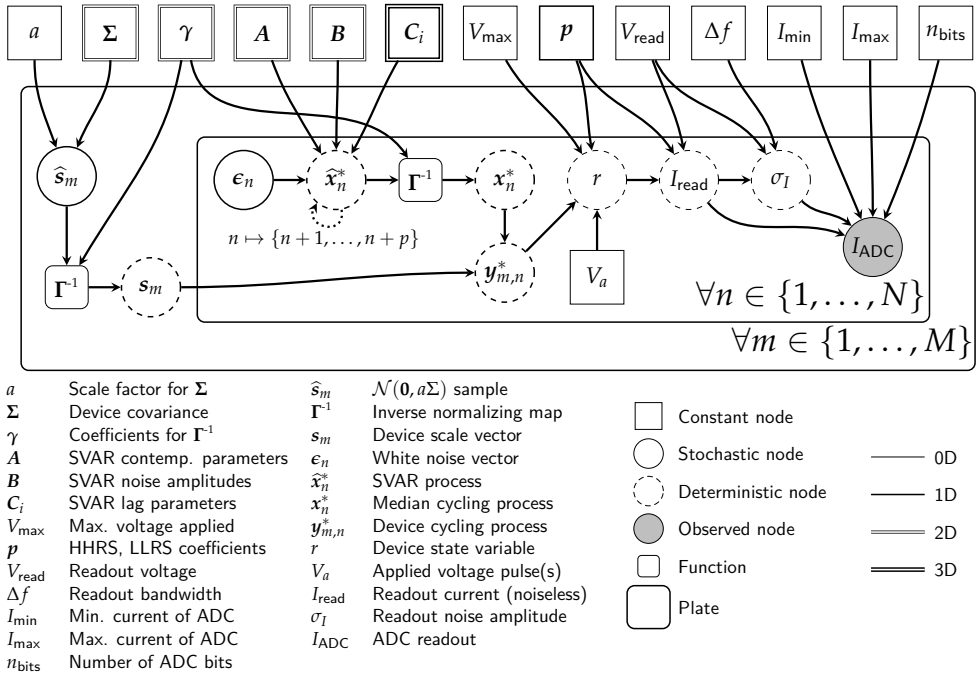

Figure 3.4: Graphical model depicting the relationships between all parameters and latent variables involved in the stochastic synapse model. Plate notation is used to represent N switching cycles of M devices, each yielding an observed readout current. The dotted recurrent arrow denotes a connection to each of the p following frames, as needed by the history dependent stochastic process.

Cycle-to-cycle variations

In seeking to represent the input time series $\boldsymbol{x}_n$ with a stochastic process, our main goals are to recreate the marginal distributions as well as the correlation structure of its vector components. To achieve the first goal with high generality, we use an approach based on transformation of the measured densities to and from the standard normal distribution $\mathcal{N}(0,1)$. This way, a single process can be used to achieve any set of marginals presented by the input data, with the relatively unrestrictive requirement that this base process generates normal marginals. Notationally, we define and apply an invertible, smooth mapping $\boldsymbol{\Gamma} : \mathbb{R}^4 \to \mathbb{R}^4$ that normalizes the marginal distributions of the vector components,

$$\boldsymbol{x}_n = \begin{bmatrix} R_H \\ V_S \\ R_L \\ V_R \end{bmatrix}_n \xrightarrow{\boldsymbol{\Gamma}} \begin{bmatrix} \widehat{R}_H \\ \widehat{V}_S \\ \widehat{R}_L \\ \widehat{V}_R \end{bmatrix}_n = \widehat{\boldsymbol{x}}_n, \tag{3.4}$$

where a hatted variable signifies that it is distributed as $\mathcal{N}(0,1)$. We then construct a base process $\widehat{x}^*_n$ whose marginals are normal, and finally transform its output back to the original data distributions via the inverse map $\boldsymbol{\Gamma}^{-1}$. The

overall process $\boldsymbol{x}_n^*$ is thus defined,

$$\widehat{\boldsymbol{x}}_n^* = \begin{bmatrix} \widehat{R}_H^* \\ \widehat{V}_S^* \\ \widehat{R}_L^* \\ \widehat{V}_R^* \end{bmatrix}_n \xrightarrow{\boldsymbol{\Gamma}^{-1}} \begin{bmatrix} R_H^* \\ V_S^* \\ R_L^* \\ V_R^* \end{bmatrix}_n = \boldsymbol{x}_n^*, \tag{3.5}$$

where a star indicates a generated random variable to distinguish from variables originating from measurement data.

This type of density transformation procedure is a widely used technique for working with arbitrary distributions, which finds application in a variety of fields and can be constructed in many different ways [192, 193]. While the transformation is trivially derived in the case where the target quantile function and its inverse are each analytically defined, we do not make this assumption in the present scenario. A simple numerical method in this case is a so-called quantile transform, where the input and output quantile functions are each discretely sampled and the transformation is defined through a direct map between bins or through interpolation. The main requirement for $\boldsymbol{\Gamma}$ in our model, however, is that its inverse (Eq. 3.5) is easy to evaluate without causing cache misses due to memory access, thus it is preferable to avoid referencing and interpolation of large look-up tables. The forward transformation (Eq. 3.4), on the other hand, only needs to be computed once for model fitting and is not used for the generating process. We therefore define $\boldsymbol{\Gamma}^{-1}$ as essentially a quantile transform, operating on each feature independently, that is evaluated from a fit of the quantiles to a specific analytic function. Namely,

$$\boldsymbol{\Gamma}^{-1}(\widehat{\boldsymbol{x}_n}) = \exp \begin{bmatrix} \gamma_1(\widehat{R}_{H,n}) \\ \gamma_2(\widehat{V}_{S,n}) \\ \gamma_3(\widehat{R}_{L,n}) \\ \gamma_4(\widehat{V}_{R,n}) \end{bmatrix} = \boldsymbol{x}_n, \tag{3.6}$$

where γ_1-γ_4 are each 5th degree polynomials, and the exponential function is applied element-wise. The coefficients of the polynomials are fit to standard normal quantiles vs. those of the respective (log) features, sampled at 500 equally spaced values between 0.01 and 0.99. The fitted polynomials are checked for monotonicity within four standard deviations above and below zero, and the forward transformation,

$$\boldsymbol{\Gamma}(\boldsymbol{x}_n) = \begin{bmatrix} \gamma_1^{-1}(\log R_{H,n}) \\ \gamma_2^{-1}(\log V_{S,n}) \\ \gamma_3^{-1}(\log R_{L,n}) \\ \gamma_4^{-1}(\log V_{R,n}) \end{bmatrix} = \widehat{\boldsymbol{x}}_n, \tag{3.7}$$

is computed using numerical inverse of the γ polynomials. A visualization of the function $\boldsymbol{\Gamma}$ as well as the marginal histograms corresponding to input series $\boldsymbol{x}_n$ and output series $\widehat{\boldsymbol{x}}_n$, are shown in Fig. 3.5.

With the input measurement data transformed into a normalized vector time series $\widehat{\boldsymbol{x}}_n$, a suitable stochastic process will be chosen for fitting. This process should serve as a useful approximation to the true physical mechanisms that generated the data, capturing the long-range correlation structure of the observed features. Time series analysis is broadly used across scientific and engineering domains, but despite its applicability to the rich statistical behavior displayed by resistive switching devices, device models have not yet widely employed dependent stochastic processes. Many models and analyses assume for convenience that features are independently and identically distributed according to a normal or lognormal PDF [183, 194]. However, there is not a strong theoretical basis for this assumption in a highly nonlinear and path-dependent system based on continuous evolution of conducting filaments. Dependent stochastic processes, on the other hand, more appropriately allow for a description of the dependence of future states on past states.

Simple models in the category of Markov chains have been considered as generating processes for memory cells. A rudimentary example is a 1-dimensional random walk process, where each future state is computed as a random additive perturbation on the previous state [185]. While random walk represents a reasonable short range approximation, it has the well known property that the expected absolute distance between the initial value and the Nth value is proportional to $\sqrt{N}$ for large N, causing the process to eventually drift to unphysical values without the use of artificial constraints.

Autoregressive (AR) models are simple univariate processes sharing some characteristics of random walk, but based additionally on a deterministic linear dependence on past observations. Each new term of an AR(p) (AR of order p) model is computed by linear combinations of p previous (lagged) values together with a noise term, producing processes that are wide-sense stationary and mean-reverting within suitable parameter ranges [195, 196]. The few times they have appeared in the literature, low order models like AR(1) and AR(2) were used to describe state variables independently (e.g. a sequence of high and/or low resistance states) [197, 198]. Here we pursue a more comprehensive statistical description of the interrelations between the different variables contained in the vectors $\widehat{\boldsymbol{x}}_n$ which takes into account long range correlations $p \gg 1$. This is enabled by using a VAR(p) model (vector AR of order p), which is the multivariate counterpart of the AR model applicable to discrete vector time series [195, 196].

We adopt a structural VAR (SVAR) formulation of the model, which is a factorization that makes the relationships between the contemporaneous (same index) variables explicit. The model has the form

$$\boldsymbol{A}\widehat{\boldsymbol{x}}_n^* = \sum_{i=1}^{p} \boldsymbol{C}_i \widehat{\boldsymbol{x}}_{n-i}^* + \boldsymbol{B}\boldsymbol{\epsilon}_n, \tag{3.8}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}_i$ are 4×4 matrices of model parameters, and $\boldsymbol{\epsilon}_n$ is a 4-dimensional standard white noise process. With this formulation we impose

a general structure of causal ordering for the generated random variables consistent with the chronological chain of measurement events. Within this structure, each variable may have a causal and deterministic effect on all future variables within range p, as visualized by the graph of Fig. 3.6. The size of these effects are all subject to fitting via the coefficients of the model. Constraints on the structural parameters,

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ A_{21} & 1 & 0 & 0 \\ A_{31} & A_{32} & 1 & 0 \\ A_{41} & A_{42} & A_{43} & 1 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} B_{11} & 0 & 0 & 0 \\ 0 & B_{22} & 0 & 0 \\ 0 & 0 & B_{33} & 0 \\ 0 & 0 & 0 & B_{44} \end{bmatrix} \tag{3.9}$$

enforce the desired causal structure while assuming an uncorrelated noise driving process. Model fitting was performed using the Python statsmodels package [199], wherein a VAR(p) model is first fit by ordinary least squares regression, and a maximum likelihood estimate is then used to determine the structural decomposition.

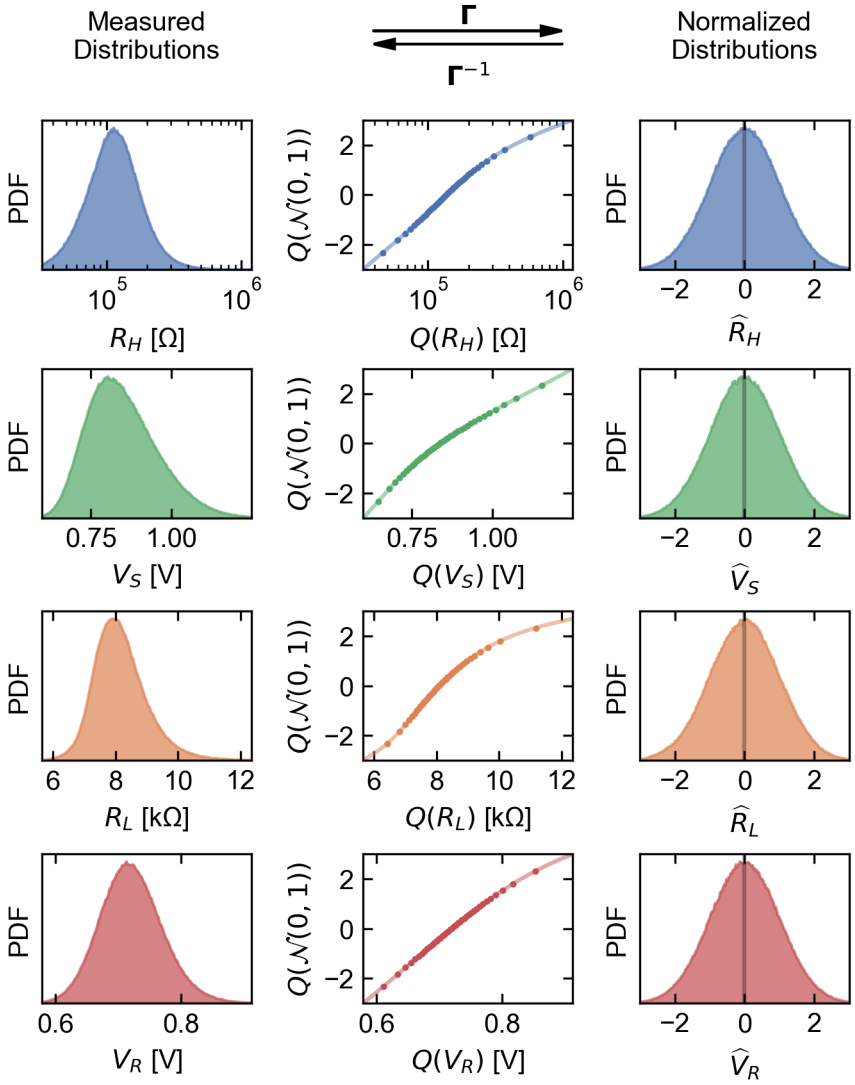

Figure 3.5: Visualization of the invertible normalizing transformation $\mathbf{\Gamma}$ that is applied to the measured feature vectors before fitting with a base stochastic process. The left column shows the marginal PDFs of the vector time series $\boldsymbol{x}_n$ extracted from measurement. The center column shows the input and output quantile-quantile plots with the fitted log-polynomial function used to transform the distributions (here, Q denotes the quantile function of its argument). The right column is the result of applying $\mathbf{\Gamma}$ to the input data, producing $\widehat{\boldsymbol{x}}_n$ whose elements are normally distributed.

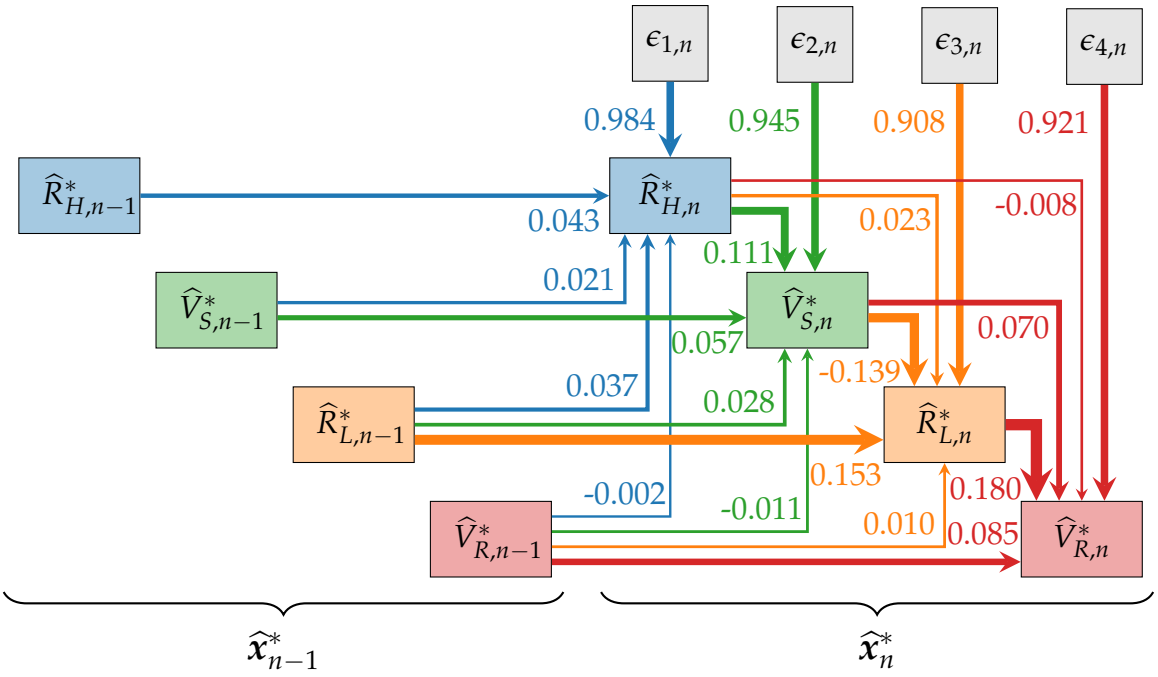

Figure 3.6: A weighted graph displaying the causal structure of the utilized SVAR(p) process, showing the nearest temporal contributions to realizations of the random vector $\widehat{\boldsymbol{x}}^*_n$. Arrow weights show the model parameters contained in $\boldsymbol{A}$, $\boldsymbol{B}$ and the upper triangular part of $\boldsymbol{C}_1$ when fit with $p = 100$. The actual SVAR(p) model uses many more connections $(16p + 10)$ than shown, so that each variable is impacted by all past values of all other variables within cycle range p.

Device-to-device variations

So far, we have only considered the statistical modeling of the cycling process of a single memory cell. However, the purpose of the presented model is to simultaneously simulate a large number of cells in a network. Individual memory devices on a wafer generally show statistical variations, mainly arising due to defects and non-uniformities in fabrication [74, 200]. These DtD variations depend strongly on the particular lithography processes and materials used. They can also originate from intrinsic factors and are influenced by conditions during the electroforming of each cell [201, 202]. Because of the potential positive or negative impact on network performance [200, 203], it is important for the model to account for the DtD variability.

The electrical effect of device variability is modeled with each cell using a modification of the same underlying SVAR cycling process. Device-specific processes are defined as members of a parametric family of processes, all based on element-wise scaling of $\boldsymbol{x}_n^*$, where the scaling factors are themselves random vectors. The specific process is denoted

$$\boldsymbol{y}_{m,n}^* = \boldsymbol{s}_m \odot \boldsymbol{x}_n^*, \tag{3.10}$$

where $m = \{1, 2, \ldots, M\}$ is the device index, $\odot$ is the Hadamard (element-wise) product, and $\boldsymbol{s}_m$ are 4×1 random vectors drawn from a fixed distribution at cell initialization.

The distribution of $\boldsymbol{s}_m$ is chosen so that the features of the median cycles of different devices are distributed and correlated in the same way as the measured cycling data $\boldsymbol{x}_n$. This choice reflects that the covariations of switching features DtD arise in the same physical system with causes and effects that are comparable to those of the CtC variations. To this end, random vectors $\widehat{\boldsymbol{s}}_m$ are drawn from a multivariate normal (MVN) distribution and $\boldsymbol{\Gamma}^{-1}$ is then reused to map them to the measured CtC distribution,

$$\boldsymbol{s}_m = \boldsymbol{\Gamma}^{-1}(\widehat{\boldsymbol{s}}_m) \oslash \boldsymbol{\Gamma}^{-1}(\boldsymbol{0}), \text{ where } \widehat{\boldsymbol{s}}_m \sim \mathcal{N}(\boldsymbol{0}, a\boldsymbol{\Sigma}). \tag{3.11}$$

Here, the denominator of the Hadamard division ($\oslash$) sets the median scale vector to the identity, $\boldsymbol{\Sigma} = \mathrm{cov}(\widehat{\boldsymbol{x}}_n)$ is the sample covariance of the normalized measurement data, and a is a free scalar parameter providing adaptability to different DtD covariance levels. A robust determination of a requires measurement of many switching cycles across a large number of devices of interest. Values in the range $a \in [1, 1.5]$ approximately correspond to published DtD measurement samples [74, 200], but improved processing and electroforming procedures may justify the use of $a < 1$.

Control logic

As components of a network, each simulated cell possesses a resistance state that encodes the weight of a connection. Voltage pulses directly applied to the cells are used to produce resistance state transitions to update the weights.

In this model, applied voltage pulses are distinguished only by a scalar amplitude V_a, whether they are in fact square waveforms or they have a more complex shape of an action potential. The duration of the pulses is assumed to be appropriately matched to the experimental timescale, such that a simulated voltage pulse of a given amplitude produces an effect comparable to the experimental voltage sweep at the instant it reaches that same amplitude. Possible state modifications in response to an input pulse is computed with respect to I, V sweeps that are reconstructed from each stochastic feature vector generated for each cycle as illustrated in Fig 3.7.

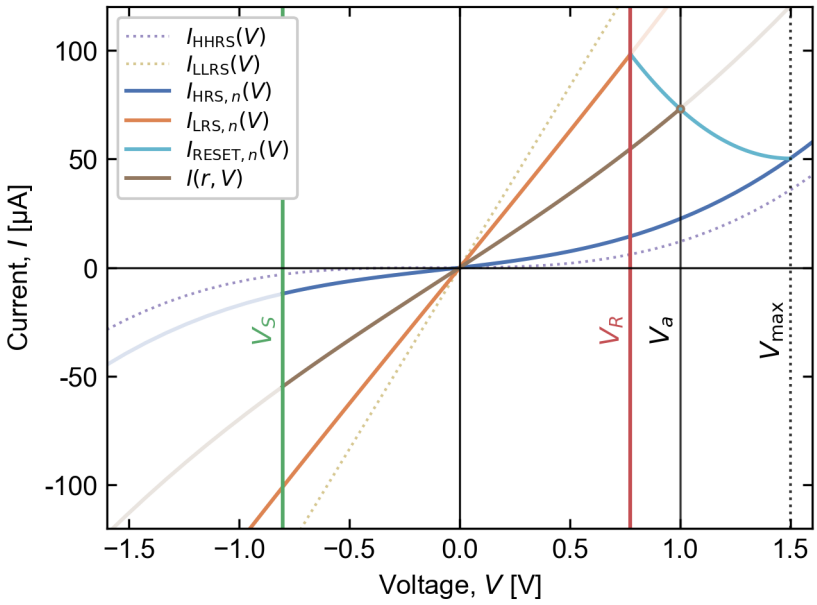

Figure 3.7: Conduction polynomials and threshold voltages allow reconstruction of (I, V) cycles from generated feature vectors. Simulated resistance switching is such that the conduction state $I(r, V)$ induced by an applied voltage V_a intersects the reconstructed cycle at $V = V_a$. For visual simplicity, the cycle shown begins and ends in the same HRS $(R_{H,n} = R_{H,n+1})$.

As previously specified in Eq. 3.1, every possible electrical state of a device is assumed to correspond to a polynomial $I(V)$ dependence parameterized by a state variable r. It is straightforward to calculate that the state variable for a curve passing through an arbitrary (I, V) point is uniquely given by the function

$$r(I, V) = \frac{I_{\mathrm{LLRS}}(V) - I}{I_{\mathrm{LLRS}}(V) - I_{\mathrm{HHRS}}(V)}. \tag{3.12}$$

Therefore the state variable corresponding to any static resistance level R (evaluated at V_0) can be calculated using

$$r(R) = \frac{I_{\mathrm{LLRS}}(V_0) - V_0 R^{-1}}{I_{\mathrm{LLRS}}(V_0) - I_{\mathrm{HHRS}}(V_0)}. \tag{3.13}$$

The $I(V)$ curves for the electrical states corresponding to each cycle's HRS and LRS, hereafter called $I_{\mathrm{HRS,n}}(V)$ and $I_{\mathrm{LRS,n}}(V)$, are defined according to equations (3.1) and (3.13) such that their static resistance equals the respective value of $R^*_{H,n}$ and $R^*_{L,n}$.

Transitions between the HRS and LRS states in response to an applied pulse amplitude V_a follow an empirically motivated structure, represented by the flow chart of Fig. 3.8. The SET transition for the nth cycle $\mathrm{HRS}_n \rightarrow \mathrm{LRS}_n$ may occur for negative voltage polarities and follows a simple threshold behavior, fully and instantaneously transitioning the first time a voltage pulse with amplitude $V_\mathrm{a} \leq V^*_{S,n}$ is applied. In contrast, the RESET transition $\mathrm{LRS}_n \rightarrow \mathrm{HRS}_{n+1}$ occurs gradually in the positive polarity with increasing V_a in the range $V^*_{R,n} < V_\mathrm{a} \leq V_\mathrm{max}$, where $V_\mathrm{max} = 1.5$ V is the maximum voltage applied in the voltage sweeping measurement. A transition curve $I_{\mathrm{RESET,n}}(V)$ is defined to connect the (I, V) points of the two limiting states where the RESET transition begins and ends. The functional form of the transition curve is chosen to be the parabola with boundary conditions

$$I_{\mathrm{RESET,n}}(V^*_{R,n}) = I_{\mathrm{LRS,n}}(V^*_{R,n}) \tag{3.14}$$

$$I_{\mathrm{RESET,n}}(V_\mathrm{max}) = I_{\mathrm{HRS,n+1}}(V_\mathrm{max}) \tag{3.15}$$

$$\left.\frac{dI_{\mathrm{RESET,n}}}{dV}\right|_{V=V_\mathrm{max}} = 0. \tag{3.16}$$

When a voltage pulse in the RESET range is applied, an intermediate resistance state (IRS) results which is calculated with reference to the transition curve such that $I(r, V_\mathrm{a}) = I_{\mathrm{RESET,n}}(V_\mathrm{a})$. Additional RESET pulses with larger amplitudes may be applied to incrementally increase the cell resistance, with HRS_{n+1} being reached only if $V_\mathrm{a} \geq V_\mathrm{max}$, after which no further RESET switching is possible for the nth cycle. After either partial or full RESET, the resistance may only decrease again by entering the following LRS_{n+1} with a voltage pulse meeting the SET criterion $V_\mathrm{a} \leq V^*_{S,n+1}$.

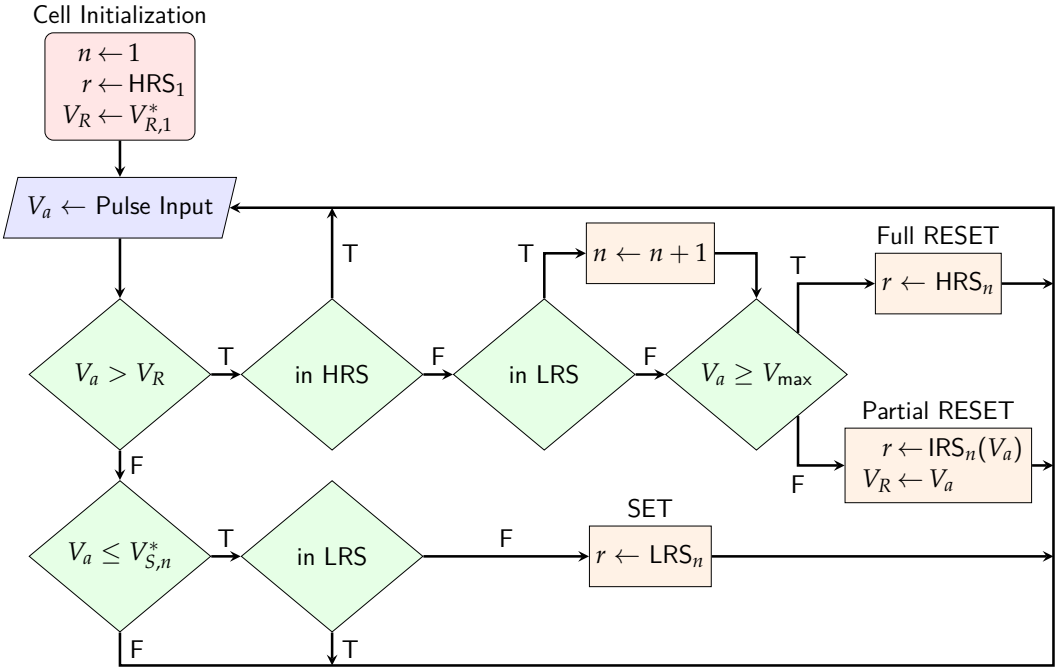

Figure 3.8: Logical flow chart showing how applied voltage pulses affect the state of each cell during simulation. Following the experimental observations, SET processes always occur abruptly below a threshold voltage, while partial switching is induced for a range of RESET voltages, with intermediate states bounded for cycle n by resistance values between $R_{L,n}$ and $R_{H,n+1}$. As resistance cycling progresses, later terms of the stochastic driving process are used for limiting resistance states and threshold voltages. Pulse amplitudes not producing a state change are efficiently disregarded.

Readout

Simulated current measurements (readouts) for each individual cell can be generated given an arbitrary readout voltage input V_{read}. The noise-free current level simply corresponds to evaluation of $I(r, V_{\text{read}})$ for each cell. In any real system, however, current readouts are accompanied by measurement noise, which may impact system performance and even present a fundamental bottleneck. Furthermore, in digital systems current readouts are converted to finite resolution by analog to digital converters (ADCs). Due to constraints of power consumption and chip area, ADC resolution is often limited such that digitization is the dominant contributor to the total noise [204]. Many additional noise sources can be considered, such as the Johnson-Nyquist and shot noise, which represent a lower bound of the noise amplitude impacting all systems.

To account for measurement noise, each individual current readout includes an additive noise contribution drawn from a normal distribution. The

noise amplitude is approximated from the Nyquist and Schottky formulas,

$$\sigma_I = \sqrt{\frac{4k_B T I_{\text{read}} \Delta f}{V_{\text{read}}} + 2qI_{\text{read}}\Delta f}, \tag{3.17}$$

where Δf is the noise equivalent bandwidth, k_B is the Boltzmann constant, $T = 300$ K is the temperature, q is the electron charge, I_{read} is the noiseless current readout, and V_{read} is the voltage used for readout. The total current is then ideally digitized with an adjustable resolution n_{bits} between adjustable minimum I_{min} and maximum I_{max} current levels.

3.1.4 Program implementation

To facilitate investigations of neuromorphic systems, model implementations designed to simulate arrays of devices were developed in the Julia programming language. Julia is a modern high-level language that is focused on performance and provides an advanced ML and scientific computing ecosystem. Julia programs compile to efficient native code for many platforms via the LLVM compiler infrastructure, and a cursory analysis indicated that single threaded CPU performance of a Julia implementation is up to 5,000 times faster than a Python implementation. Furthermore, as modern computational resources are highly parallel, Julia's support for CPU multi-threading and GPU programming through CUDA.jl [205] is an important advantage.

All model parameters corresponding to the characterized device characterized in this work, including different possible SVAR model orders, $p \in [1, 200]$, are stored in a binary file which is read in by the program at startup. Each instantiated cell stores state information and p cycles of history using primarily 32-bit floating point numbers. The total memory footprint grows linearly with the chosen model order and is approximately $16p + 56$ bytes per cell. A reduced form VAR process is used to compute realizations of $\boldsymbol{x}_n^*$, which are lazily evaluated along with the parabolic transition polynomials if and when they are needed. The majority of the necessary runtime computations are formulated as matrix multiplications, which are heavily optimized operations across many different contexts.

The present release contains two model implementations in order to suit a wide variety of computing platforms and use cases [206]. The first is a CPU optimized version wherein the cells of an array are individually addressable for read/write operations. These operations are naturally parallelized for multicore processors by partitioning the cells and assigning each partition to independent threads of execution. The second implementation is a GPU accelerated version compatible with CUDA capable GPUs. This version uses a vectorized data structure and parallel array abstractions to take advantage of the implicit parallelism programming model of CUDA.jl. Here, all defined cells are always accessed simultaneously, with each read/write operation employing optimized linear algebra GPU kernels. While the GPU implementation

integrates well with other ML components residing in GPU shared memory and achieves higher throughput per cell for large parallel operations, the CPU implementation obtains higher update rates for sparse operations commonly encountered in large-scale models [207, 208].

3.2 Simulation results

As shown visually in the scatterplot of Fig. 3.9, the stochastic process $\boldsymbol{x}_n^*$ generates data that closely resemble the measurement data $\boldsymbol{x}_n$. The generated distributions match the empirical distributions so closely that it is difficult to visualize their difference. The Wasserstein metric is a distance function defined between probability distributions that can be used to quantify a small discrepancy [209]. The first Wasserstein distance was calculated element-wise and averaged across 100 realizations of $\boldsymbol{x}_n^*$ with length 10^6. The result,

$$\overline{\boldsymbol{W}}_1(\boldsymbol{x}_n, \boldsymbol{x}_n^*) = \begin{bmatrix} 5146\ \Omega \\ 937\ \mu\text{V} \\ 20\ \Omega \\ 356\ \mu\text{V} \end{bmatrix}, \tag{3.18}$$

is much smaller than the mean feature vector, $\bar{\boldsymbol{x}}_n$ (Eq. 3.3), and independent of the chosen model order. This shows that the goal of reproducing the measurement distributions is well achieved for the input dataset by using the described method of probability density transformation.

Simulations of full (I, V) cycling measurements (Fig. 3.10[A]) show close similarity with the measurement data of Fig. 3.2. Multi-resistance-level capability is also demonstrated by a similar simulation involving partial RESET operations by changing the maximum voltage applied (Fig. 3.10[B]). The dependence of the resulting HRS value on the applied voltage reproduces a nonlinear characteristic comparable to experimental findings [210, 211].

While a full structural analysis of the fitted SVAR(p) model parameters $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}_i)$ will not be presented here, a few aspects are worthy of note. For the fit corresponding to the particular device and measurement described here, the white noise terms are by far the dominant contributors to all four modeled features. The contemporaneous terms $(\boldsymbol{A})$ and first order $(\boldsymbol{C}_1)$ terms are the next most significant, which indicates that the most recent cell history is most relevant for generating the proceeding states. Nevertheless, input data correlations persist for many cycles, and the generating process $\boldsymbol{x}_n^*$ successfully reproduces the overall correlation structure of the data up to at least p cycle lags, as shown in detail in Fig. 3.11.

Although no physical effects were explicitly put into the model definition, it is important to recognize that they are quantitatively captured and put into a useful statistical context by the SVAR model fitting procedure. The model weights contained in $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}_i$ quantify deterministic relationships between past and future variables even in the presence of large random fluctuations. As

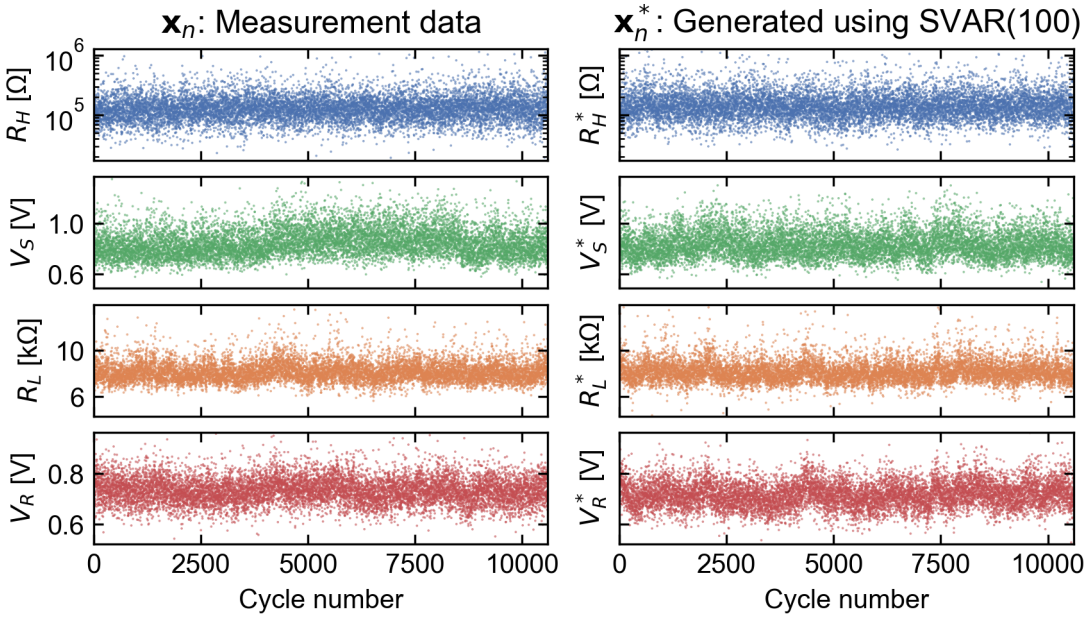

Figure 3.9: Comparison of feature time series extracted from measurement data and those generated by the SVAR-based model. The compared features converge to effectively equivalent distributions and the short-range behavior is qualitatively similar across thousands of cycles.

seen in the graph of Fig. 3.6, the four strongest coefficients in the fitted model correspond to the relationships

$$\widehat{R}^*_{H,n} \xrightarrow{0.111} \widehat{V}^*_{S,n}, \tag{3.19}$$

$$\widehat{V}^*_{S,n} \xrightarrow{-0.139} \widehat{R}^*_{L,n}, \tag{3.20}$$

$$\widehat{R}^*_{L,n-1} \xrightarrow{0.153} \widehat{R}^*_{L,n}, \tag{3.21}$$

$$\widehat{R}^*_{L,n} \xrightarrow{0.180} \widehat{V}^*_{R,n}. \tag{3.22}$$

Comparable relationships between switching variables have been identified and discussed in physics-based models and simulations as well as in experimental studies involving various materials [107, 114, 212–216]. According to relation 3.19, larger starting HRS values tend to contribute to a higher SET voltage, which is a well known effect due to a reduced driving force for ionic motion at a given applied voltage, as a larger HRS gives both reduced power dissipation as well as a reduced electric field in a thicker insulating gap. The subsequent LRS is strongly affected by the SET voltage (relation 3.20). This can be attributed to the runaway nature of the SET transition and a higher voltage initial condition, and is also connected with the dynamics of the current limiting circuitry [163]. The LRS value is also strongly correlated with the value of the previous LRS (relation 3.21), because of the influence of the residual filamentary structure from the previous cycle [217]. Lastly, relation 3.22 indicates that higher LRS values tend to have larger reset voltages, which has to do with

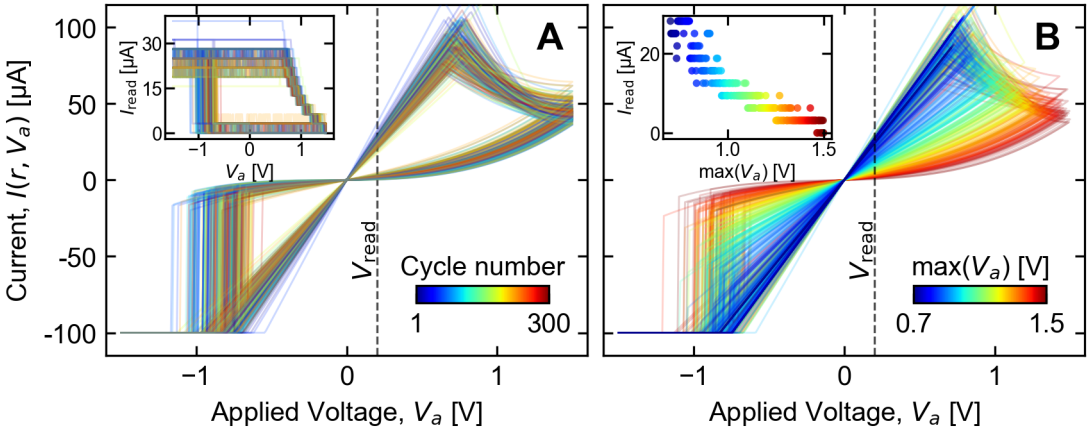

Figure 3.10: Two example simulations involving repeated cycling of a single device. Voltage pulse sequences were applied with varying amplitude following a triangular envelope, and the (I, V) characteristic of the each cycle is plotted in a different color. Subplot **(A)** shows 300 consecutive cycles between the full voltage range ± 1.5 V, with a readout performed after every pulse (inset). Subplot **(B)** demonstrates multilevel capability with 300 cycles between -1.5 V and maximum voltage that increases each cycle, from 0.7 V to 1.5 V. Readouts following each cycle are shown in the inset. In each case, readouts were simulated using a fixed $V_{\text{read}} = 200$ mV, including noise and 4-bit quantization between $I_{\text{min}} = 0$ μA and $I_{\text{max}} = 40$ μA.

a balance of factors influencing filament dissolution, including temperature and drift. This balance depends on the cell materials, operating regime, and internal series resistance [218].

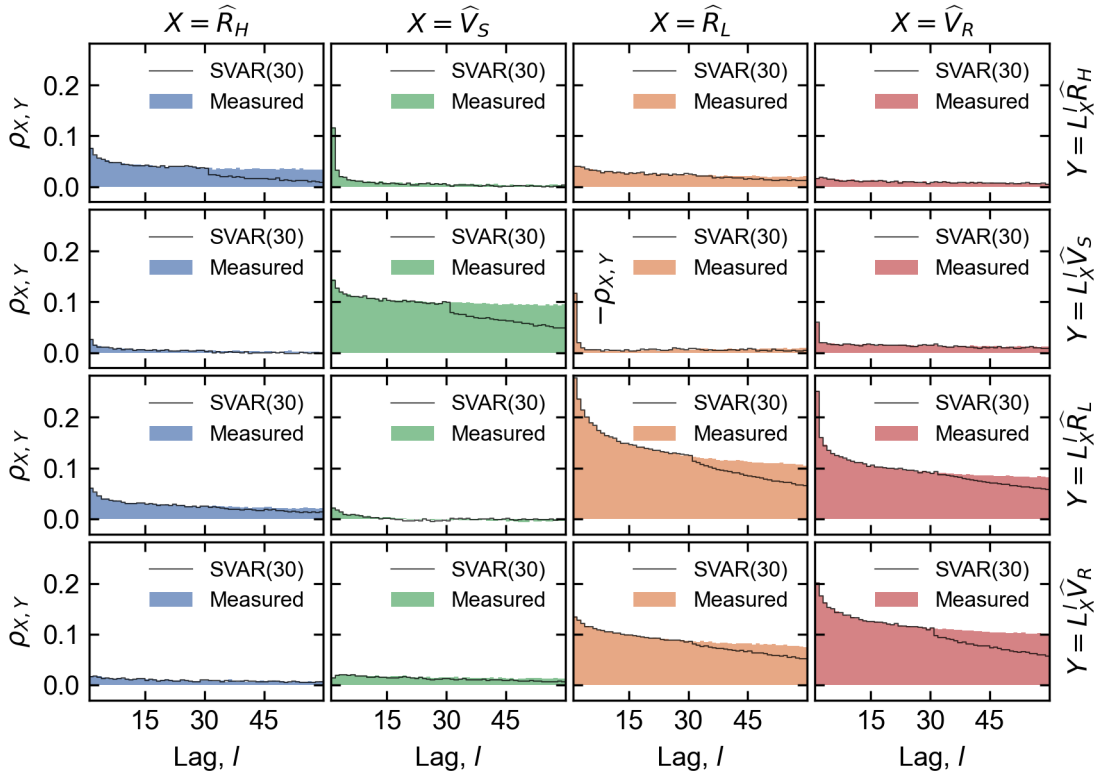

Figure 3.11: Auto- and cross-correlations of the normalized feature vector components, showing the Pearson coefficients $\rho_{X,Y}$ of the variables specified in the subplot columns X and rows Y as a function of lag l. Row variables are lagged with respect to column variables, as denoted by the lag operators L_X. A comparison between measurement data and data generated from SVAR(30) shows extremely close agreement up to cycle range 30. For lags larger than the chosen model order, some of the correlations of $\widehat{x}^*$ decay more quickly than $\widehat{x}$.

3.3 Benchmarks

As a benchmark of the throughput of write operations, repeated resistance cycling was induced on arrays of simulated cells under varying conditions. In each case, voltage pulse sequences to be applied to all defined cells were generated prior to the benchmarks, consisting of amplitudes ± 1.5 V with alternating polarity. Defined as such, every pulse drives each cell through a transition into its next HRS or LRS. The read operation was benchmarked separately under equivalent conditions, reading out the entire array using a fixed readout voltage of $V_{\text{read}} = 0.2\ V$.

The CPU benchmark was performed on using an Intel Xeon Silver 4116 CPU, varying the cell array size M, the order of the VAR process p, as well as the number of threads used to perform the operations in parallel. The resulting read/write throughputs are summarized in Fig. 3.12. Write throughputs were obtained in the range $10^7 - 10^8$ operations per second (OPS), or between $10 - 100$ ns per individual write operation. Read operations were approximately

an order of magnitude faster than writes, with $10^8 - 10^9$ OPS or $1 - 10$ ns per read operation. Due to the size of necessary matrix multiplications, increasing the VAR order p incurs a cost of write throughput, with a $p = 100$ model running approximately $2.4\times$ slower than one with $p = 10$. The read operation, in contrast, shows a negligible dependence on the VAR order.

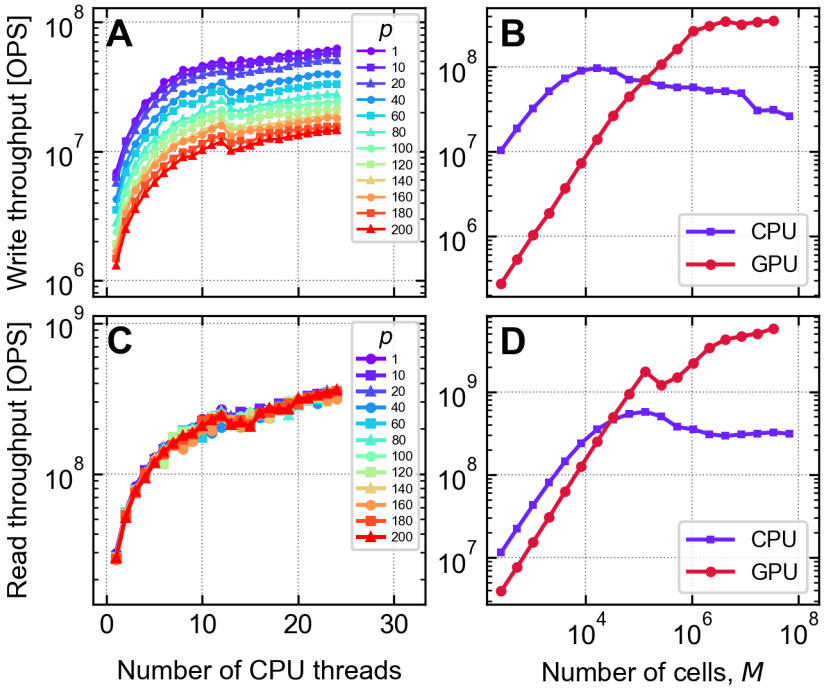

Figure 3.12: Benchmarks of the read/write operation throughput per cell of the Julia model implementations. In **(A)** and **(C)**, an array of 2^{20} ($\sim 10^6$) cells are simulated on the CPU as a function of number of parallel threads spawned, and the VAR model order p. In **(B)** and **(D)**, the CPU (24 threads) and GPU implementations are benchmarked versus the array size M, with $p = 10$.

The GPU accelerated version was benchmarked in an analogous way, using the same host machine with an NVIDIA TITAN RTX GPU device. The results are shown in dependence of the cell array size M in Fig. 3.12(B,D). The GPU implementation overtakes the CPU above $M = 10^5$ parallel operations where the entire array is updated, and achieves an order of magnitude faster performance for large arrays with $M > 10^6$. However, CPU throughput is applicable to subsets of the array, and may retain an advantage for sparse operations.

3.4 Discussion

In order to assess the potential of emerging synaptic devices, new lightweight and accurate device models are needed to constitute the millions/billions of weights used in modern machine learning (ML) models. Candidate memory

cells such as ReRAM are highly non-linear stochastic devices with complex internal states and history dependence, all of which needs to be explicitly taken into account. In this work we introduced an efficient generative model for large synaptic arrays, which closely reproduces the statistical behavior of real devices.

Taking advantage of a recently developed electrical measurement technique [163], we systematically fit the model to a dataset that is dense in relevant information about the device state evolution. Together with this new kind of measurement, our modeling approach helps complete a neuromorphic design feedback loop by defining a programmatic connection from the measured behavior of a fabricated device under the intended operating conditions directly to fitted model parameters. Probability density transformation of the underlying SVAR stochastic process gives the model the power to accurately reproduce nearly arbitrary distribution shapes and covariance structures across the switching cycles and across the separate devices. These features enable evaluation of network performance while automatically adapting to a wide variety of possible future device designs.

We provide parallelized implementations for both CPU and GPU, where up to 15 million cells per gigabyte of available memory can be simulated at once, and benchmarks show write throughputs above three hundred million weight updates per second. As a point of reference, this throughput exceeds the pixel rate of a 30 frames per second video stream at 4K resolution (3840×2160 pixels). Realistic current readouts including digitization and noise were also benchmarked, and are approximately an order of magnitude faster than weight updates. While speeds can be expected to improve with future optimizations, these benchmarks give a basis for estimating the scope of applicability of the model to ML tasks.

The implementation and the general concept of this model are naturally extendable. Although model parameters were adapted here to a specific HfO_2-based ReRAM device, the method is applicable to a variety of other types of stochastic memory cells such as PCM, MRAM, etc. Four specific switching features were chosen in this demonstration to reconstruct (I, V) cycling behavior, but additional switching parameters can also be extracted from measurements and accommodated within this framework. Ideally informed by statistical measurement data, different functional forms, transition behaviors, time dependence, and underlying stochastic processes can each be substituted. Fitting may also be performed with respect to the output of physics-based simulations, thereby establishing an indirect link to physical parameters while achieving much higher computational speed. With these considerations, the model represents a flexible foundation for implementing large-scale neuromorphic simulations that incorporate realistic device behavior.

Chapter 4

Mott-oxide neuronal nano-devices

The graphics marked by a copyright in this chapter are adapted, with permission, from Ref. [44] and Ref. [45].

4.1 Thin film deposition

Synthesis of pure-phase $(V_{1-x}Cr_x)_2O_3$ thin films is a challenging materials engineering problem, and is the subject of previous work by Jonathan A. J. Rupp. The deposition of the $(V_{1-x}Cr_x)_2O_3$ films discussed in this chapter are one of the products of this prior work. For completeness, the deposition conditions parameters will be briefly recounted here, but the interested reader can refer to other works which have described various aspects of the material synthesis and optimization in much greater detail [219–222].

Depositions of mixed vanadium/chromium oxide thin films were carried out using reactive RF magnetron sputtering from alloyed $V_{1-x}Cr_x$ sputtering targets with $x = 0.0$, 0.05, and 0.15. One inch diameter targets were used, with a target to substrate distance of 5 cm and an incident angle of 15°. With an applied sputtering power of 50 W and with the substrate temperature held at 600 °C, polycrystalline layers were deposited on various substrates including platinum, titanium nitride, silicon nitride membranes, and silicon membranes. Nominally amorphous films were also deposited at room temperature (20 °C) under otherwise identical conditions as the polycrystalline films.

As part of the material optimization, the oxygen partial pressure was adjusted to obtain the targetted corundum phase and $(V_{1-x}Cr_x)_2O_3$ sesquioxide stoichiometry. The base pressure was held below 1 nbar and the working pressure was fixed at 10 μbar, and the oxygen partial pressure was controlled by simultaneous flow from two separate gas sources; The first containing pure Argon and the second a mixture of oxygen/argon in the ratio 1/99. Crystallographic phases were determined for using X-ray diffractometry in grazing incidence mode (GI-XRD) with a X'Pert PRO powder diffractometer (PANalytical). Film stoichiometry was evaluated by transport measurements through the low temperature V_2O_3 metal-insulator phase transition, the steepness of which is known to depend very sensitively on the defect concentration in the sub 1% range [223]. Optimized conditions were obtained with the mass flow controllers for the two gases set to 94 sccm and 6 sccm, respectively.

For the determination of thickness and calculation of deposition rate, X-ray reflectometry data were acquired and fitting was performed by PANalytical's "X'pert Reflectometry". Using the optimized deposition conditions, film thicknesses between 90 nm to 5 nm were deposited within a few minutes for each of the substrates and each of the chromium concentrations. The desired pure-phase corundum V_2O_3 crystallinity was confirmed by GI-XRD down to 30 nm thickness, but was not well resolved by this technique for thicknesses below 30 nm.

To avoid post-oxidation or degradation before or during X-ray characterization, all $(V_{1-x}Cr_x)_2O_3$ films examined by X-ray were capped *in situ* with a 2 nm tantalum film.

4.2 Film characterization with X-ray absorption

While the bulk properties of the $(V_{1-x}Cr_x)_2O_3$ system are relatively well studied and understood [89, 224], properties of thin films are less known because of the extra experimental challenge involved. Due to the small length scale, the composition and crystal configuration in films are difficult to characterize, especially at extremely low thicknesses below 10 nm. In thin films, phase coexistence, off-stoichiometry, and interface effects such as strain and diffusion all conspire to impact the metal-insulator phase transition behavior. X-ray absorption spectroscopy (XAS) is a widely used technique that measures excitations of atomic core levels to probe the electronic structure of a sample; it is useful for chemical speciation, and is applicable to extremely thin material layers. In particular, X-ray absorption near edge structure (XANES) analysis provides information about the valence and coordination environment of a specific element in the sample. The energies and intensities of vanadium K-edge absorption features were found to be correlated with the vanadium oxidation state in bulk and powdered oxides [225–227], but so far this technique has not been applied to thin $(V_{1-x}Cr_x)_2O_3$ film samples.

Here, we show XANES characterization for the various $(V_{1-x}Cr_x)_2O_3$ films described in the previous section. XANES measurements were performed at the microXAS X-ray beamline at the Swiss Light Source facility (Paul Scherrer Institute, Switzerland). Hard X-rays from the synchrotron were monochromatized to the Cr and V K-edges using a double Si crystal and focused to a spot size of approximately 1.5 μm (horizontal)$\times$1 μm (vertical) using an elliptically shaped Rh-coated mirror pair in the Kirkpatrick-Baez (KB) geometry. Incident X-ray intensities were measured with a He-flushed ion chamber before reaching the sample. The photon energy was calibrated with respect to the K-edge of a reference vanadium foil, with the first detected peak of its derivative spectrum defined as 5465 eV.

Reference powders of binary oxides (V_2O_3, VO_2, V_2O_5) were obtained from Sigma-Aldrich (analytical grade). To prevent oxidation, the powders were handled and stored in an inert environment prior to measurement. Each reference material was encapsulated in Kapton tape compartments in order to

be measured in transmission mode. While in transit to the beamline, reference samples were packaged individually in aluminized gas barrier bags, along with RP agent (500 mL) from Mitsubishi Gas Chemical Company, Inc., which absorb moisture, oxygen, and corrosive gases.

Reference powder samples were measured in transmission mode, and thin film specimens were measured in fluorescence mode using a Si drift detector. Absorption and fluorescence spectra were acquired for incident photon energies between 5330 eV to 5800 eV for the V K-edge and between 5850 eV to 6400 eV for the Cr K-edge. The photon energy was scanned with a varying step size with a minimum step size of 0.2 eV near the main absorption edges. The scanning procedure was repeated four times for the V edge and ten times for the Cr edge, to provide a higher SNR by averaging over the repeated scans. The partial fluorescence yield was extracted from each of the XRF spectra through numerical integration between 4780 eV to 5110 eV for the V K_α peak and 5290 eV to 5550 eV for the Cr K_α peak. These integration windows were carefully selected to avoid overlap with the elastic scattering peaks. A correction was made to the Cr K_α (5414 eV) intensity to account for the overlapping the V K_β line (5427 eV). This background was subtracted using the extracted V K_α intensity, assuming a fixed intensity ratio of the V K_α to V K_β of ten. All fluorescence yields were normalized by the incoming beam intensity and corrected for detector dead time. Spurious peaks in the XANES spectra, arising from Bragg peaks intruding into the V and Cr K_α integration windows, were manually pruned from the dataset.

All XANES spectra for references and thin films were normalized by the same procedure and using the same parameters. The normalization procedure first set the pre-edge baseline to zero by subtraction of a linear fit within a fixed pre-edge energy range. The post-edge was then normalized to unity through division by a line which was fit within a fixed post-edge energy range. For the V K-edge, the respective fit ranges were 5389 eV to 5455 eV and 5671 eV to 5768 eV. For the Cr K-edge, they were 5911 eV to 5979 eV, and 6241 eV to 6360 eV. As shown in Fig. 4.1, absorption spectra for both K-edges were clearly resolved all the way down to the weakest signal case (5 nm film thickness and 5% Cr content). Between specimens, main edge energy differences on the order of 1 eV, as well as strong intensity variations of the pre-edge features were detected.

Main edge energies were extracted from the spectra after normalization by interpolation to the 0.5 absorption level crossing. The extracted edge energies were seen to increase with the cation oxidation state for both V and Cr oxide reference compounds, in agreement with the trends previously reported [225–227]. The extracted edge energies of the film specimens were compared with the reference materials to judge oxidation state and identify potential intermixing of the many possible VOx phases (Fig. 4.2). The edge energies for all film samples were distributed closest to the values of reference materials V_2O_3 and Cr_2O_3, showing no clear deviations from the intended material phase $(V_{1-x}Cr_x)_2O_3$ with the utilized variations in deposition parameters. There

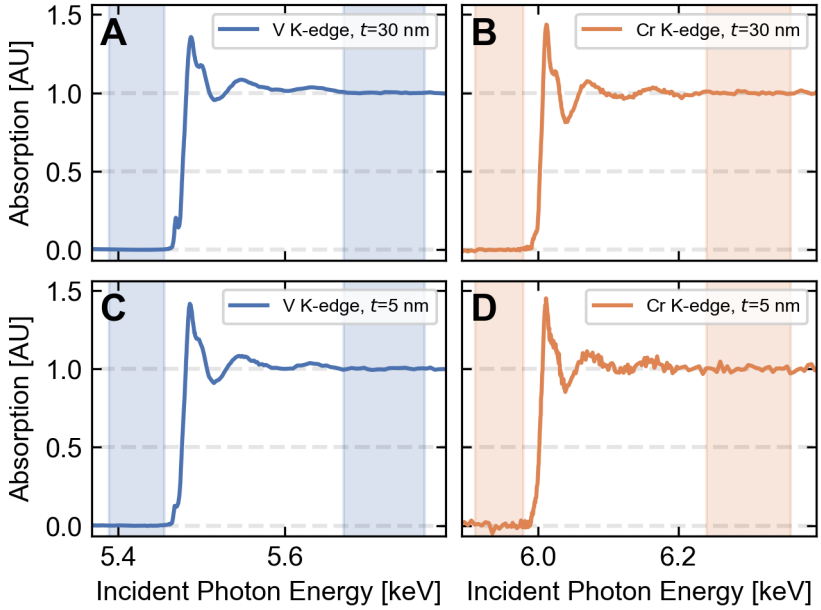

Figure 4.1: Normalized absorption spectra (from fluorescence yield) for the K edges of V **(A, C)** and Cr **(B, D)** for 30 nm **(A, B)** and 5 nm **(C, D)** thin films, each with 5% Cr content. Good signal quality was obtained even for the Cr K edge of 5 nm films. The energy ranges used for normalization are indicated by the colored transparent bands.

is a trend of increasing energy for both the V and Cr edges with film thickness, suggesting a possible reducing effect of the Si membrane substrate on the oxide film near the interface. Strain effects induced by lattice parameter mismatch could also play a role here, with such effects being more prominent for the thinner films.

Even more pronounced than the main edge energy shifts, the thin film XANES spectra also showed strong variations in the pre-edge peaks when different deposition parameters were used (Fig. 4.3). This pre-edge corresponds to a 1s $\rightarrow$ 3d electronic transition, which is formally dipole forbidden but becomes allowed due to a slight deviation from octahedral symmetry of the V coordination environment [225–227]. A triplet structure arises here due to the crystal field splitting of the ground state. These structures are very sensitive to the stoichiometry, the crystal phase, and can also be related to the Mott insulating state [88, 228].

The intensity of the pre-edge peak in the VOx system is suggested in the literature to be a more sensitive indicator of the V oxidation state than the main K-edge energy shifts. This was studied by several groups through reference powder measurements of compounds with different oxidation states (e.g. VO, VO_2, V_2O_3, V_2O_5) obtained from a commercial source (Alfa Aesar) [225–227]. However, a comparison of the literature spectra reveals a quantitative discrepancy of the pre-edge intensities both between the studies and with the

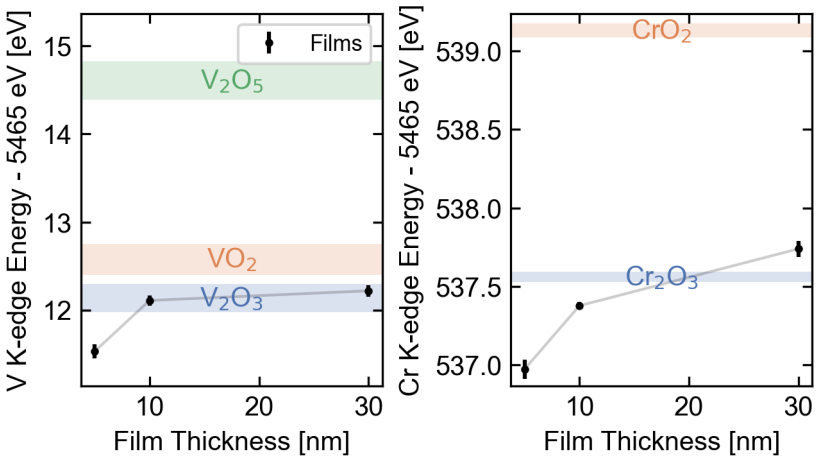

Figure 4.2: Variation of the measured V and Cr main K-edge energy vs. film thickness. The shown films were deposited at temperature 600 °C on 10 µm Si membrane substrates and contain 5% Cr. Colored bands indicate the measured edge energy ($\pm 1\sigma$) for the reference oxide powders. A trend of increasing edge energy with film thickness suggests an impact of the substrate interface.

present study (Fig. 4.4). With V_2O_5 being the most thermodynamically stable oxide, powders of lower V oxides tend to slowly oxidize in atmosphere. Therefore, larger pre-peak intensities could indicate varying levels of contamination with V_2O_5 occuring due to air contact during sample preparation and transport. Until the true reference spectra are clarified, the present discrepency confounds a complete evaluation of the pre-peaks in the XANES spectra of the thin film samples.

In summary, we have successfully collected XANES spectra of ultra-thin $(V_{1-x}Cr_x)_2O_3$ films deposited under a variety of conditions, which are the first of their kind reported for this material system. The spectra show strong variations in the pre-edge peak structure, which contain information about the vanadium valence state and its coordination geometry. We demonstrated that the pre-edge region is very sensitive to the stoichiometry and could be better probe for the valence states than the main edge energy. We also find a strong influence of the substrate on the pre-edge that will require further investigation. We expect that combining X-ray microscopy and electrical measurements will allow a way to bridge the chemical microstructures with the functional properties of devices.

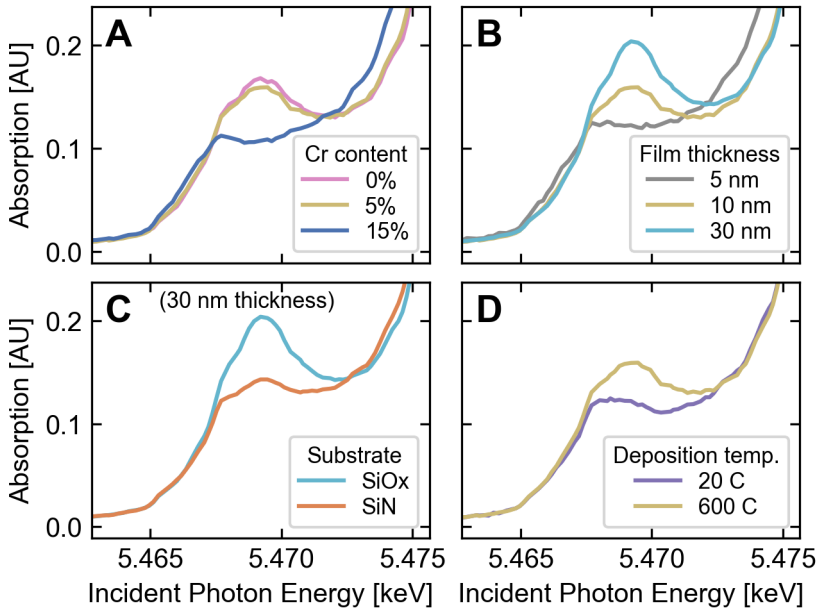

Figure 4.3: Pre-edge (V-K) absorption peaks of the thin film specimens. Peak intensities are strongly affected by variations of deposition parameters, **(A)** Cr content **(B)** film thickness **(C)** substrate **(D)** deposition temperature. Each subplot shows the effect of a parameter variation from the reference condition of 10 nm thickness, 10 µm SiOx substrate, 600 °C deposition temperature, and 5% Cr concentration.

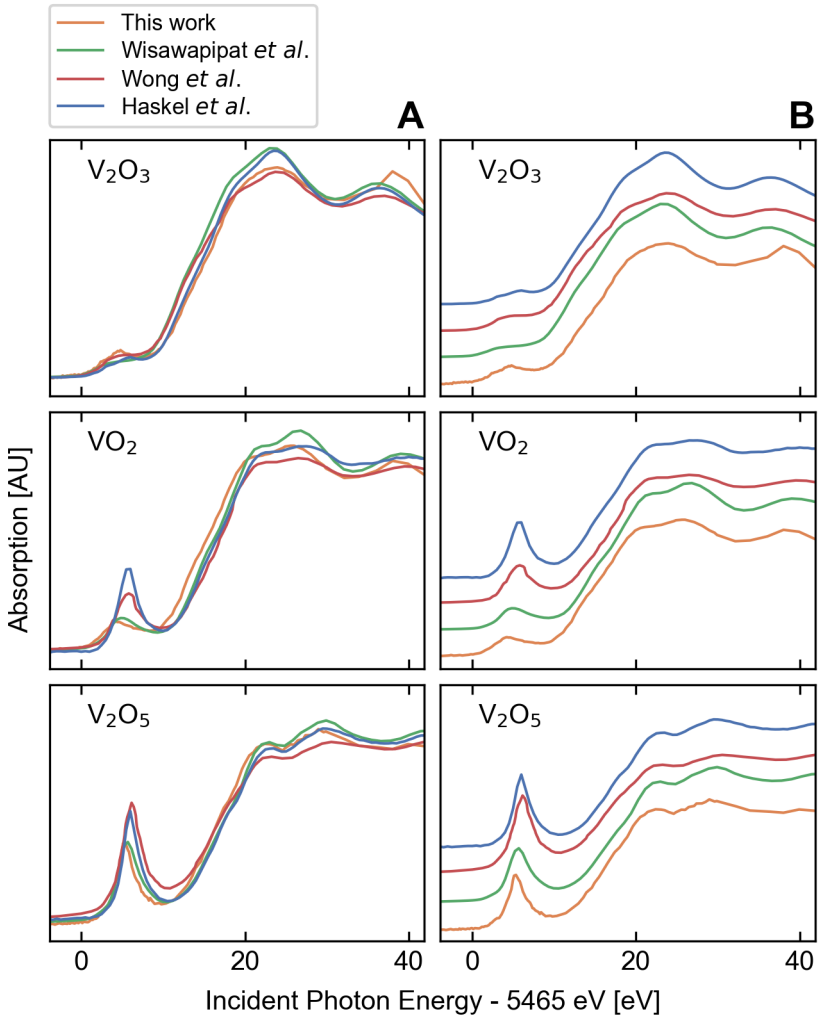

Figure 4.4: Absorption spectra of vanadium oxide reference powders measured in this work (V_2O_3, VO_2, V_2O_5) compared with literature references. Strong discrepancies in pre-peak intensities as well as edge energies are displayed. Column **(A)** shows the overlapped spectra and column **(B)** shows the same spectra with relative offsets.

4.3 Nanodevice fabrication

For all electrical investigations of resistance switching phenomena covered in this chapter, $(V_{1-x}Cr_x)_2O_3$-based nano-scale metal-insulator-metal (MIM) devices were prepared using the same deposition parameters as the films used in the XANES study in Section 4.2. The devices consist of full oxide films contacted from below by square TiN vias and capped with 30 nm Pt top electrodes. The effective device widths, determined by the dimensions of the TiN via, varied across the dies between 120 nm and 500 nm, and the deposited oxide thickness ranged from 90 nm to 5 nm. A device cross-section for a 90 nm oxide thickness is shown in the SEM micrograph of Fig. 4.5.

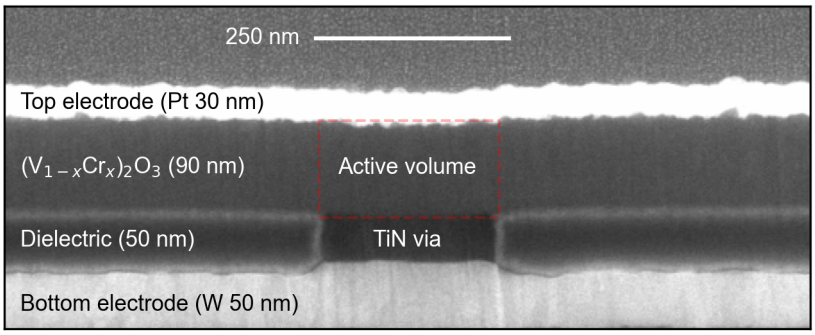

Figure 4.5: Cross-sectional SEM of a $(V_{0.85}Cr_{0.15})_2O_3$ nanodevice, using a 90 nm thick oxide layer, 30 nm top electrode (Pt), and a 250 nm × 250 nm bottom point contact (TiN). The active volume of the cell is indicated with a dashed red outline.

Each individual device has independent top and bottom electrode contacts routed to 100×100 µm probing pads, both electrically isolated from the surrounding film and substrate. The pads have an estimated capacitance 65 fF to the substrate, and the capacitance of the devices themselves is much smaller, with a maximum estimated value of

$$C_d = \frac{\epsilon_0 \epsilon_r w_{ox}^2}{t_{ox}} \approx 9 \text{ fF}, \tag{4.1}$$

assuming relative permittivity $\epsilon_r = 20$, device width $w_{ox} = 500$ nm, and device thickness $t_{ox} = 5$ nm.

In addition to isolated devices with direct electrode contacts, a subset of devices also included integrated serpentine series resistors which can be used to control (I, V) trajectories during switching events. With very little parasitic capacitance between the on-chip resistors and the memory cells, devices are well protected from unwanted current overshoots and their NDR characteristics can be interrogated with minimal risk of degradation. The mean calibrated values of the resistances used were 150 Ω, 2,164 Ω, 8,197 Ω, and 12,857 Ω, each with a ±3% relative error as determined by a measurement of 18 samples of each nominal resistor size.

Ultrathin (< 10 nm) functional layers are particularly important for nanoelectronic applications. The thinnest fabricated devices (5 nm) were characterized by transmission electron microscopy (TEM), as summarized in Fig. 4.6. Material composition was confirmed using electron energy loss spectroscopy (EELS), revealing a uniform device thickness and high quality interfaces with the electrodes.

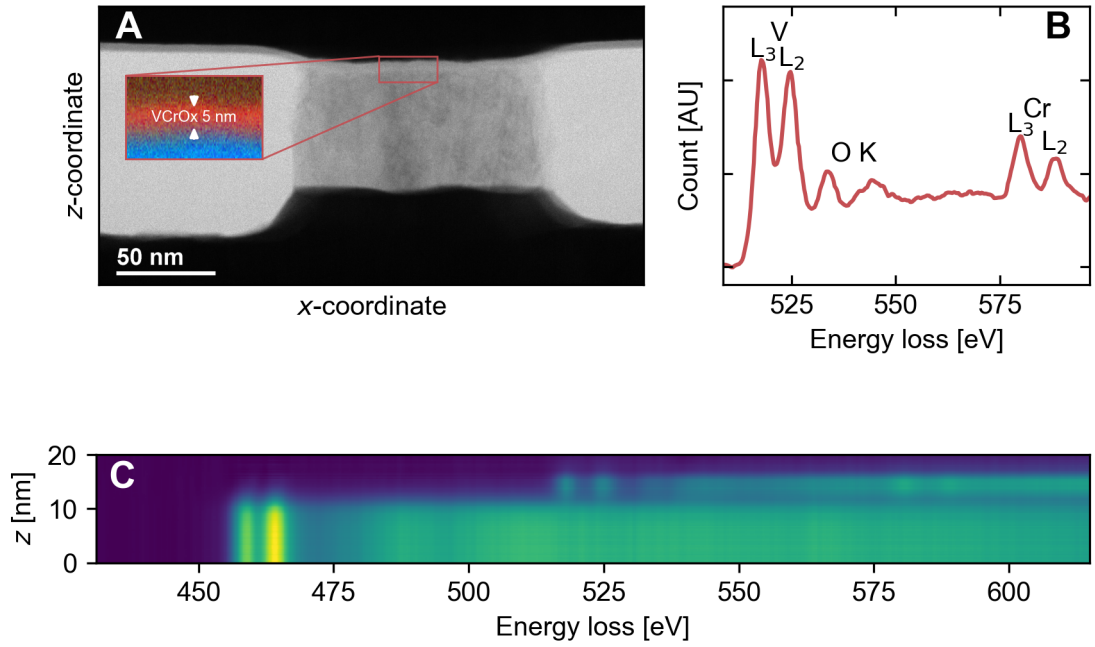

Figure 4.6: A cross-section of an ultra-thin $(V_{0.85}Cr_{0.15})_2O_3$ device (5 nm × 120 nm × 120 nm) characterized by TEM. An overview **(A)** shows the location of the thin active layer (inset). The energy loss spectrum of the 5 nm oxide layer **(B)** clearly shows the L-edges of vanadium and chromium and K-edge of oxygen. Spatially resolved EELS **(C)** shows a sharp material interface with the bottom (TiN) and top (Pt) electrodes.

4.4 Electrical characterization

4.4.1 The pristine state

Initial resistivity analysis

Following fabrication of rectangular nanodevices of varying dimensions, measurements of their initial resistances were made in order to verify electrical uniformity. For a statistical view, an automatic wafer probing system was used to measure all the devices across the different samples. High resolution subthreshold (I, V) characteristics were recorded with voltage sweeps between ± 100 mV using a direct connection of the DUT to an Agilent 4155C SPA. The device resistance was defined as the inverse slope of a least-squares fit line to the (I, V) data for $|V| \leq 100$ mV. The result, shown in Fig. 4.7, indicates that

the conductivity of the films is uniform in both area and thickness, following Pouillet's law for the resistance of a uniform medium,

$$R = \frac{\rho t_{ox}}{w_{ox}^2}, \tag{4.2}$$

where ρ is the sample resistivity, t_{ox} is the oxide thickness and w_{ox} is the device width. On average, standard deviation of the log resistance was 5.5% of the mean value for each respective device dimension. The thickness scaling suggests that the conduction mechanism is bulk-limited with no major electrical influence of the interface down to 10 nm thickness. This finding contrasts with the interface limited conduction commonly seen in MIM stacks, such as VCM cells, where lowering of a Schottky barrier is a central component in the theory of resistive switching [229, 230].

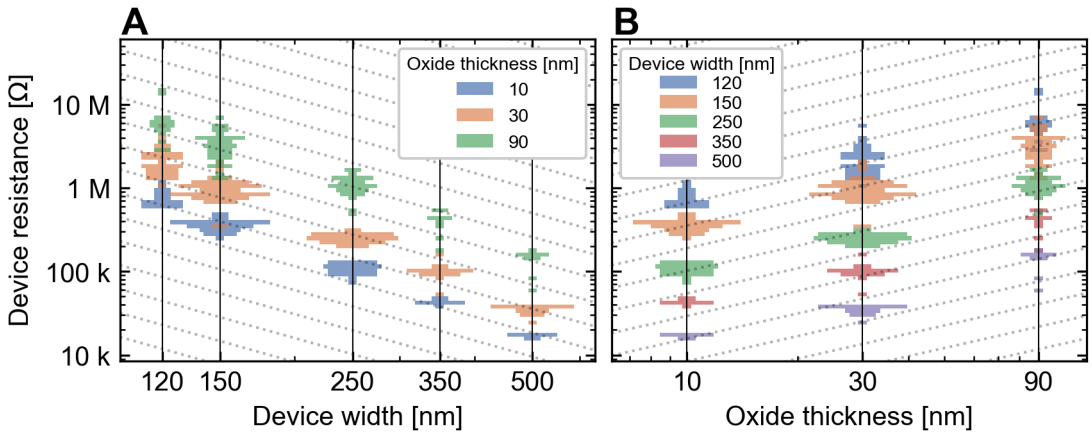

Figure 4.7: Violin plots showing resistance measurements for (as deposited) $(V_{0.85}Cr_{0.15})_2O_3$ devices as a function of **(A)** device width and **(B)** oxide thickness. The vertical (and mirrored) histograms each show the measured distributions of 10 to 50 duplicate devices with five different widths and three different thicknesses. The average log device resistance scales in proportion to oxide thickness and in inverse proportion to device area, following Pouillet's law.

Resistivity values for all measured devices were calculated using equation 4.2 and compared with two literature references (Fig. 4.8). Kuwamoto et al. [224] studied single $(V_{1-x}Cr_x)_2O_3$ crystals, which show a steep resistivity change at the phase transition boundary near $x = 1\%$. Homm et al. [83] measured thin $(V_{1-x}Cr_x)_2O_3$ films (60-80 nm) grown by molecular beam epitaxy on sapphire substrates and annealed in vacuum. Here, the loss of the sharp MIT was attributed to strain due to the substrate lattice mismatch.

The measured resistivities for the devices in our study were between one to two orders of magnitude higher than those reported for epitaxial thin films, and were closer to the bulk crystal values. The cause of this was not determined conclusively, but could be explained by a number of factors, such as different strain conditions, disorder, or slight differences in stoichiometry. For

amorphous devices with $x = 15\%$, measured resistivities were spread across four orders of magnitude, and were not uniform in film thickness. This non-uniformity could be due to a formation of an interfacial conduction barrier with these deposition conditions, in contrast to all others.

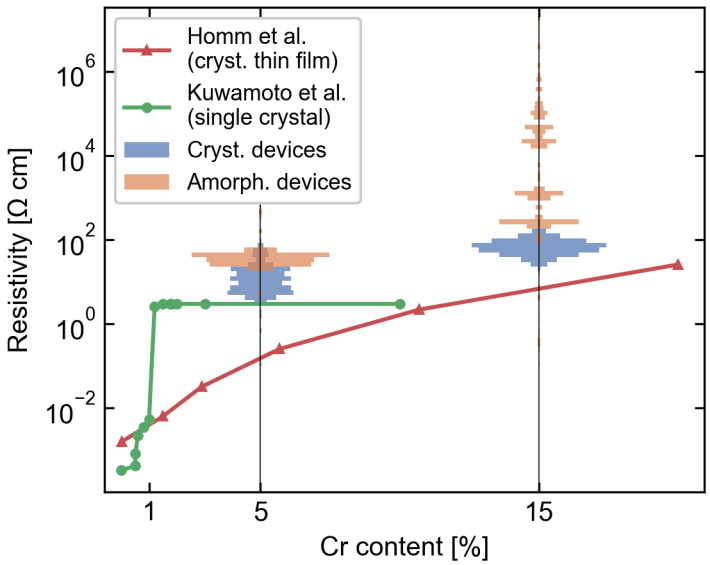

Figure 4.8: Measured resistivity distributions of $(V_{1-x}Cr_x)_2O_3$ devices (120 nm to 500 nm width) vs. chromium concentration, x. Values for crystalline and amorphous films are compared with values extracted from references [83, 224].

Temperature dependence of the conductivity

Determining the physical mechanism of electron transport is important for understanding and modeling device behavior. In particular, the impact of elevated temperatures on the conductivity is relevant for the operation of nanodevices, where Joule heating often plays a strong role. However, the theoretical mechanisms of conduction are diverse and difficult to accurately predict from first principles. Unfortunately, they are also not straightforward to ascertain from electrical measurements. Once device current is measured as a function of temperature and/or voltage, identification of a conduction mechanism amounts to a curve fitting exercise, with many candidate analytical models, and potentially several acting in parallel or in series [127]. As free parameters amass, fits become overdetermined, and good fits do not necessarily imply a physically correct description.

With the above in mind, we collected multivariate conduction data in order identify a consistent conduction model. A $(V_{0.85}Cr_{0.15})_2O_3$-based device with width 500 nm and thickness 90 nm was measured with a Keithley 2636B SMU voltage source with a 2 kΩ external resistor in series. Device current was measured as a function of both voltage applied to the top Pt electrode and of ambient temperature from 305 K to 425 K, as controlled by a custom-made

temperature stage. Device voltage was calculated by subtracting the series resistance voltage drop $I_d R_s$ from the sourced voltage. A strong temperature dependence was observed, as shown in the Arrhenius plot of Fig. 4.9.

Most dielectric conduction models, such as interface limited Schottky emission and bulk limited Poole-Frenkel (P-F) emission, have the thermal activation of carriers in common. This leads to a strong exponential temperature dependence of the current,

$$I \propto \exp\left(\frac{-E_B}{k_B T}\right). \tag{4.3}$$

In these models, applied voltage generally acts to reduce the effective energy barrier [127]. However, when the fitted Arrhenius slopes of our measurement data are attributed to energy barrier changes, E_B *increased* from 215 meV to 275 meV as device voltage increased from 0.20 V to 1.25 V. This is a first indication that self-heating needs to be taken into account, and the changing slopes can be more accurately ascribed to an increasing device temperature with applied power.

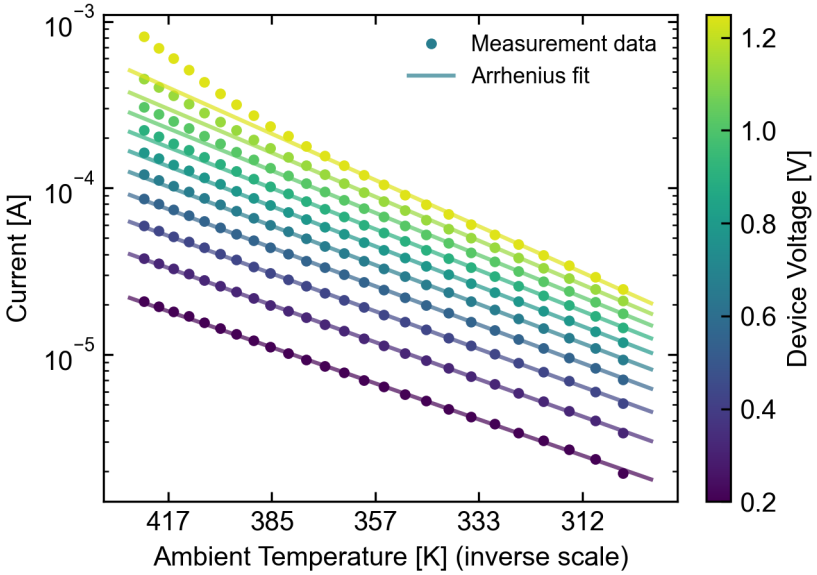

Figure 4.9: Arrhenius plot for temperature dependent conduction measurements. At low voltages and temperatures, the trends closely follow a simple thermal activation behavior with energy barriers between 215 meV and 275 meV.

Close agreement with the measurement data was found using a simple empirical relation

$$I(V,T) = aV \exp\left(\frac{-b}{T}\right) \exp\left(c\sqrt{V}\right), \tag{4.4}$$

where a, b, and c are free parameters. A thermal resistance approximation,

$$T = T_0 + R_{th} I V, \tag{4.5}$$

Table 4.1: Parameters for the current vs. voltage vs. ambient temperature fit.

Symbol	Value	Unit
a	3.433×10^{-2}	S
b	2.626×10^{3}	K
c	9.694×10^{-1}	$V^{-1/2}$
R_{th}	1.112×10^{5}	$K\,W^{-1}$

is used to calculate the cell temperature relative to the ambient level T_0, and R_{th} is also considered a free parameter. This conduction model was fit to all the multivariate measurement data simultaneously using the least squares method, and the result is shown in Fig. 4.10. As the equation system has no analytical solution, it was solved numerically inside of the fitting routine. The values of the best fit parameters are given in Table 4.1.

Although the data is displayed on a so-called P-F plot, the functional form differs from P-F conduction and has separable voltage and temperature contributions – a constant energy barrier is used with no voltage-induced lowering. A possible physical interpretation of the separate components is described in the supplementary material of [231] by a combination of thermal activation and trap assisted tunneling, though this could be further complicated by electron-electron interactions in the correlated oxide.

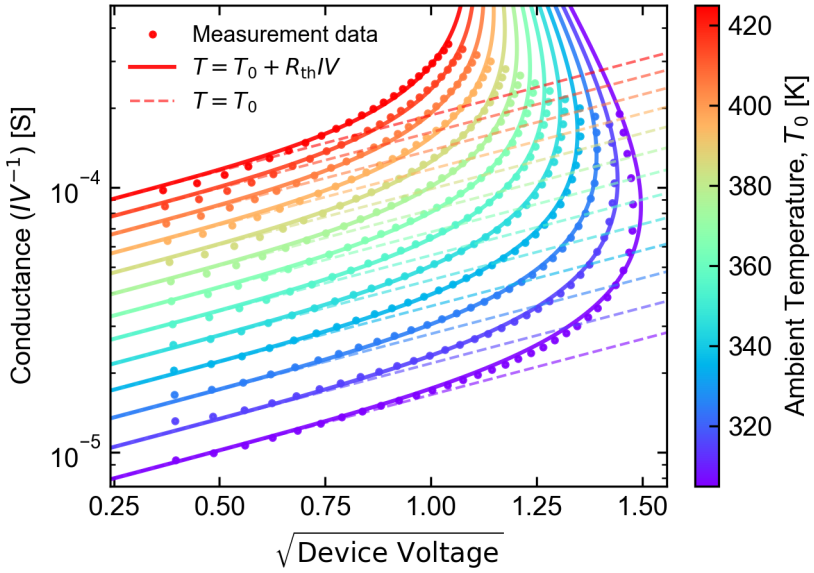

Figure 4.10: Temperature dependent conduction measurements of a $500\times500\times90$ nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ device fit with the empirical model of Eqs. 4.4 and 4.5. Dashed lines show the conduction equation without taking Joule heating into account, which highlights its importance for fitting the higher power regions of the data.

S-type negative differential resistance

As seen in the temperature dependent measurements of Fig. 4.10, the fabricated $(V_{1-x}Cr_x)_2O_3$ nanodevices show an NDR effect in the initial as-deposited state. These smooth NDR characteristics were observed in both polarities when measured with sufficiently large external series resistances and sufficiently slow (> 1 ms) voltage sweeps. The reversible NDR curves are retraced on the positive and negative measurement sweep, and their shape depended strongly on the device dimensions (Fig. 4.11). This type of NDR is classified as S-type, although only the bottom part of the S shape was reversibly observable. This was in part due to a low internal series resistance of the test devices (recall Fig. 1.12) but also because a degradation of the dielectric generally occured before a possible return to positive differential resistance at elevated currents (above ∼1 mA).

When the same cells are controlled with a different measurement circuit, it is possible to observe a threshold switching (TS) effect [37, 160]. As a point of comparison with typical quasistatic sweeping measurements, a device was measured with a direct SMU connection and different sourcing modes (Fig. 4.12). Placing an external 200 Ω series resistance near the sample was necessary to observe complete reversibility of the TS effect, due to the slow reaction of the SMU current compliance. Devices with smaller widths (< 250 nm) were less likely to endure even a single cycle of this quasistatic TS without irreversibly altering the state of the device.

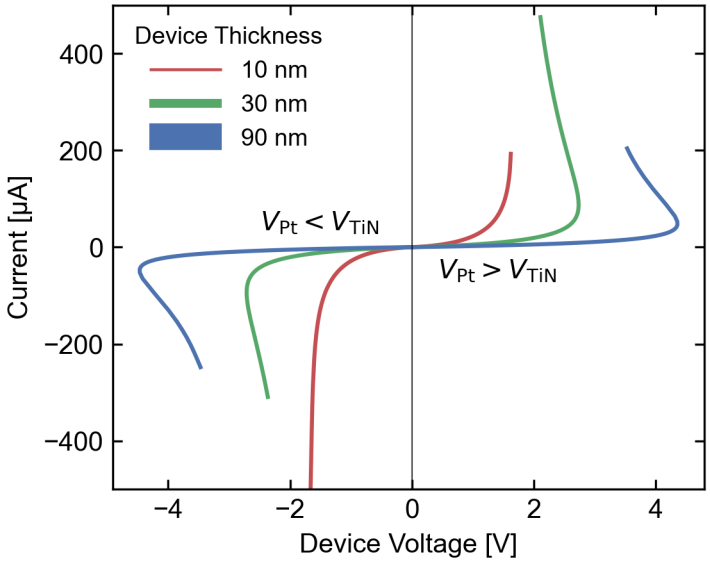

Figure 4.11: Bipolar NDR characteristics of $(V_{0.85}Cr_{0.15})_2O_3$ devices with width 250 nm and with three different thicknesses, measured using 8.2 kΩ on-chip series resistances. The device thickness has a large effect on the NDR shape, and very little difference is measured between the two voltage polarities. © 2018 IEEE.

Beyond the possibility of destructive current overshoots that are not directly detectable by the instrument, there are several pitfalls to be aware of when considering TS measurements of this standard type. Although a connecting line is plotted between each consecutive measurement data point, no datapoints are collected by the instrument during the threshold transitions because of the low (≤ 50 Hz) sampling frequency of the SMU. Furthermore, due to an unfortunate convention, the voltage values plotted for the voltage controlled curves are the programmed source voltage, not the voltage actually output by the SMU or seen at the DUT electrodes. This means that every TS datapoint shown in Fig. 4.12 that does not coincide with the device's NDR curve is in fact fictional, and does not correspond to any change in the external biasing conditions. Plotted in this way, there is an appearance of hysteresis with respect to the applied voltage, even if the device itself is not internally hysteretic. In reality, the device settles into the same set of (I, V) steady states in both the current and voltage controlled cases, and these steady states do not depend on the history but only on the present biasing condition. Therefore, we conclude that the TS and NDR effects both arise here due to the same continuous and non-hysteretic material mechanism.

The specially designed circuits discussed in Chapter 2 extend the options for non-destructive evaluation of NDR materials, and give access to a much wider range of experimental timescales. While TS is barely observable in our devices with an SMU measurement, high TS endurance is measurable using the CLA circuit, with negligible degradation over 2.5 billion cycles (see Fig. 4.13). A general weakness of TS measurements is that they do not

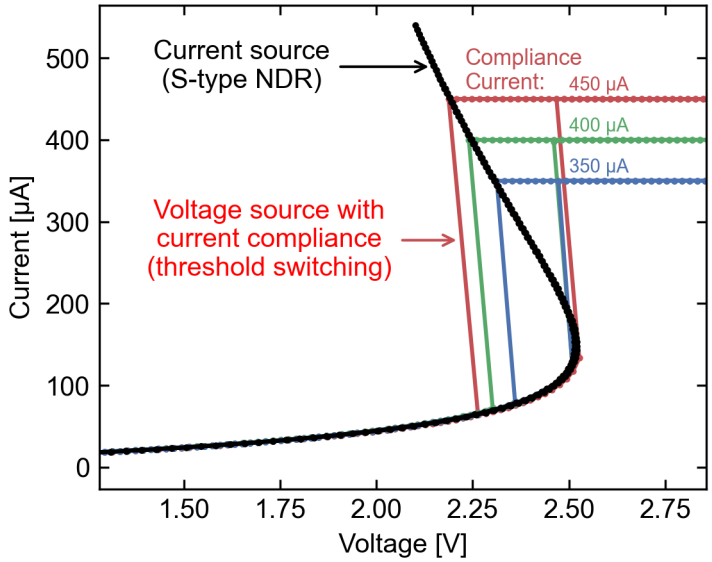

Figure 4.12: A 500×500×90 nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ device (as deposited) measured with a Keithley 2636B SMU configured as a voltage source and as a current source. In each case, an increasing and decreasing sweep were both recorded. In voltage sourcing mode with the current compliance function, typical TS behavior is measured. In current sourcing mode, a smooth reversible NDR characteristic is measured. © 2018 IEEE.

adequately interrogate the nature of the resistance transition, because the intermediate states are rendered unstable by the driving circuitry. Different physical causes of conductivity changes present themselves very similarly in TS measurements; they simply appear as rapid jumps in current. On the other hand, the digipot circuit introduced in Section 2.1 allows a convenient characterization of the device NDR characteristics (as seen in Fig. 2.6[A]). The NDR curve gives the set of fixed points of the volatile device state on the (I, V) plane, which is fundamental to understanding an overall measurement circuit as a dynamical system. Whereas switching trajectories that stray from the steady state NDR curve are highly dependent on the driving circuitry and signals, the NDR curve can be regarded as a purely device property that can be directly measured electrically in appropriate circumstances.

A study of the dependencies of the NDR curves is particularly insightful for gaining a physical understanding of the system. First, we look at the dependence of the device geometry on the shape of the NDR curve. The curves for the as-deposited devices were measured for each cell size, which varied in both thickness and in width (Fig. 4.14). This was done using external series resistance values that varied between 1 kΩ and 20 kΩ, and triangular voltage pulses lasting between 1 ms and 10 ms. To successfully collect this data, the series resistance and the timescale of the measurements had to be carefully selected for each device size. For small device widths ($\leq$ 150 nm), the cells could not indefinitely sustain steady-state biasing before a separate non-volatile process (to be further discussed) began to degrade the cell. At the same time, the

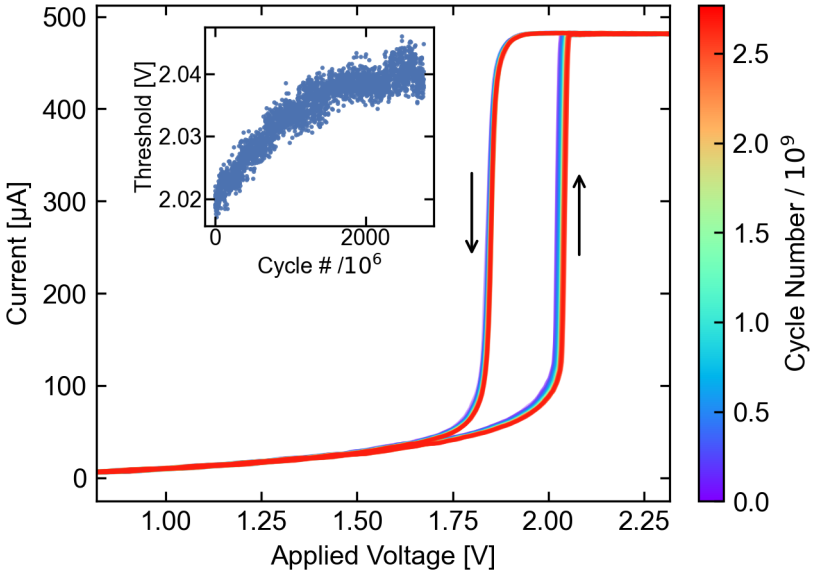

Figure 4.13: Repeated threshold switching of a $500\times500\times90$ nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ device using the CLA circuitry with current limit 480 μA. Cycles were controlled with a 1 MHz triangle wave, with 1 ms duration measurements (shown) interleaved every millionth cycle. Very little device degradation was observed, and the threshold voltage drifted by only 20 mV after over 2.5 billion cycles.

sweeps needed to be slow enough to not depart dynamically from the NDR curves we intended to measure.

In the device size dependent data, the scaling of the voltage at the point of NDR onset ($dV/dI = 0$) rules out the possibility that the devices switch due to an effect that occurs, for example, at a constant threshold electric field value. There is, however, an interesting trend in the total power ($P = IV$) at this NDR "knee". The knee power increases in proportion to the width of the device, but does not depend strongly on the thickness of the device. This observed power scaling hints at the involvement of thermal effects in the switching mechanism. We will return to this data in section 4.5, where a physical model is used to explain the NDR knee dependency and to make predictions for device behavior at even smaller dimensions.

According to the foregoing discussion, and in light of the electro-thermal theory presented in Section 1.4, there is already a strong indication for Joule-heating as the cause of NDR in the pristine $(V_{1-x}Cr_x)_2O_3$ device state. As a further direct test of this, we measured the effect of the top electrode (TE) thickness on the NDR curve, with the expectation that the TE thickness only impacts the overall thermal resistance of the cell and has no other confounding effect. We designed the experiment to examine the effect of TE thickness in isolation, leaving the initial oxide/TE interface intact. This was done with a sequence of NDR curve measurements of the same device between successive room temperature depositions of 30 nm Pt layers on top of the existing TE.

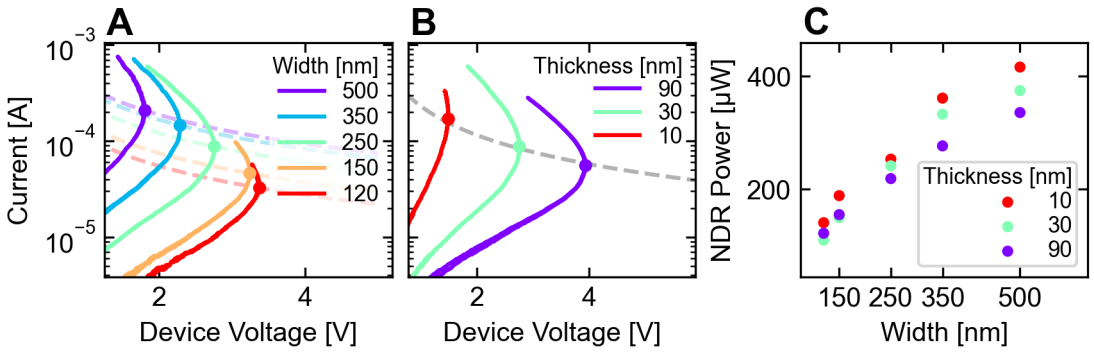

Figure 4.14: Measurements of the NDR characteristics of $(V_{0.85}Cr_{0.15})_2O_3$ nanodevices with varying **(A)** device widths and **(B)** device thicknesses. Dashed lines of constant power intersect each point of NDR onset. **(C)** The power dissipated by the device at the NDR knee depends strongly on the device width, but only weakly on the device thickness.

The results, shown in Fig. 4.15, show an increase of the NDR onset power from 328 μW to 456 μW as the TE thickness was increased from 30 nm to 90 nm. The size of this effect is beyond what could be expected from a variation in electrical resistance of the electrode lead, measured to be below 100 Ω for a 30 nm electrode thickness, and clearly points to heating as a relevant factor for the NDR mechanism.

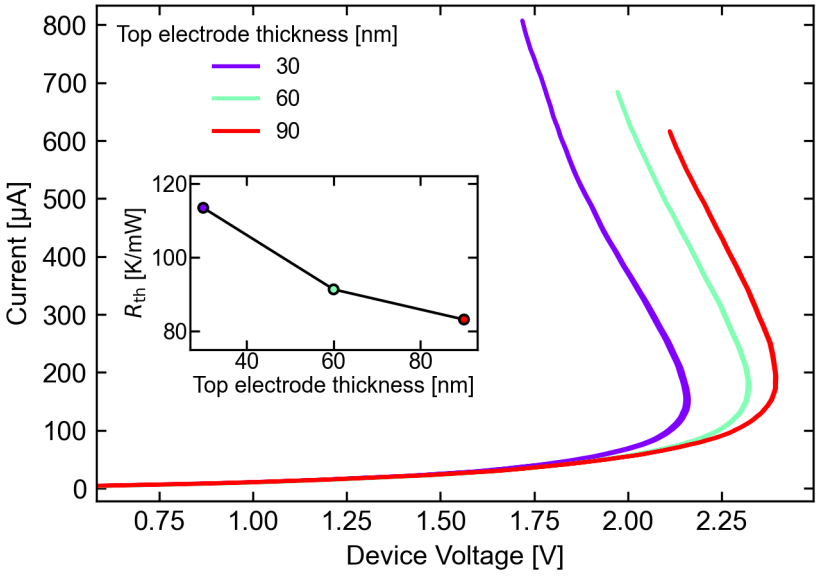

Figure 4.15: Variations in the measured NDR characteristics of the same $(V_{0.85}Cr_{0.15})_2O_3$ device when additional Pt layers were added to the top electrode. Due to the heat sinking effect of the top electrode, the power at NDR onset and the fitted R_{th} value increased with the electrode thickness. The oxide thickness was 30 nm and device width was 350 nm. Measurements were performed with 10 ms voltage sweeps and a 2 kΩ external series resistance.

4.4.2 Non-volatile effects

All insulating films break down when electrically stressed beyond certain limits. Typically through ionic migration or other reconfiguration of the insulating material, initially well insulating devices can reach new states that possess different properties, accompanied by non-volatile (NV) changes in the overall resistance level. The specific state arrived in after initial breakdown generally depends on how the measurement circuit reacts to the rapidly changing device resistance.

As an extreme example, when electroforming and measuring the cells with a quasistatic sweeping procedure using an SMU with current compliance, bipolar resistance switching is typically induced (Fig. 4.16). With this kind of measurement, it is unlikely to observe the purely volatile TS regime in the pristine state, and the measured (I, V) traces instead resemble those seen in a variety of other material systems that operate according to the VCM mechanism. This behavior is presumably connected to ionic defects being created and released into the oxide film during the electroforming current overshoot transient. As seen in the measurements of Fig. 4.16, subsequent resistive switching operations are also strongly associated with these transients, and only sporadically slow transitions are spared from experiencing large overshoot currents.

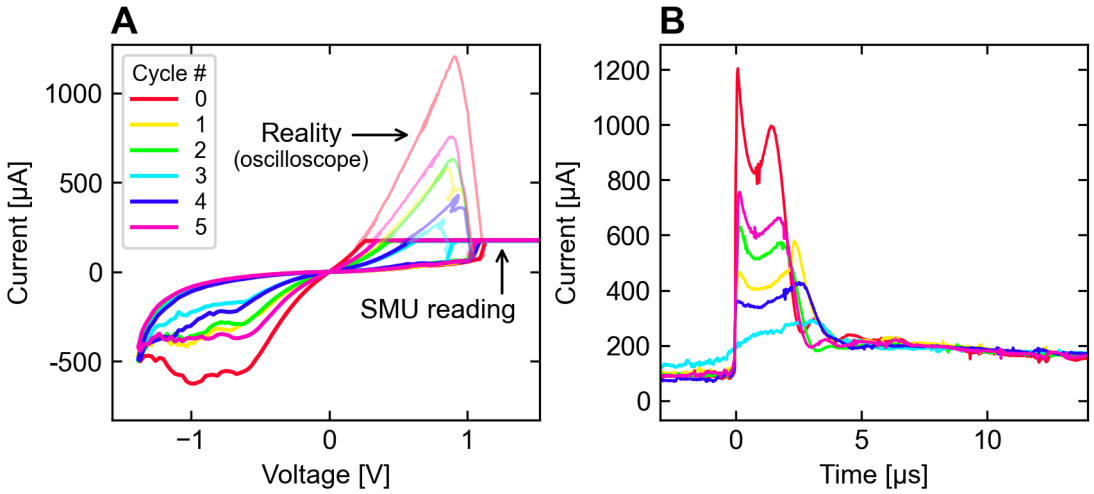

Figure 4.16: Bipolar resistive switching recorded in a $150\times150\times30$ nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ device using direct SMU voltage sourcing with current compliance 175 μA. **(A)** Data for six switching cycles as returned by the SMU, contrasted with high frequency oscilloscope samples of the same SET transitions. **(B)** The corresponding current vs. time transients induced in the measurement. Monitoring the SMU measurement with an oscilloscope reveals strong role that overshoot currents play in the switching process.

As we have seen that the material system supports both volatile and non-volatile resistive switching components, natural questions are: what are their roles in the overall switching process, when does each come into play, and what is happening physically inside the device? Electrical measurements can provide valuable insights to help answer these questions, but require careful attention to the measurement circuitry, to which specific signals are being measured, and to how they are sampled and plotted.

To get a closer look at the impact of current overshoot transients during runaway switching from the pristine device state, we can use circuitry to explicitly control and measure the amount of overshoot and examine its effect. To do this, we used the CLA circuit described in Section 2.2, which has a well known response (see Eq. 2.5), while incrementally adding capacitance in parallel to the DUT. The capacitance was controlled through connection of a rotary variable capacitor between the CLA input node and ground, and adjusting its setting in between measurements. The results, plotted in Fig. 4.17, show that reversible threshold switching was only observable when using ≤ 85 pF of added capacitance. Adding 105 pF to the measurement circuit, approximately equivalent to a 1 m coaxial cable, led instead to a NV conductance transition under otherwise equivalent conditions. While the standard TS plot of Fig. 4.17(A) gives little insight into the cause of the difference in switching behavior, the sub-1 μs transients shown in Fig. 4.17(B) reveal current spikes and oscillations following TS, with increasing amplitude according to the added capacitance value. These spikes, not present at the lowest capacitance setting,

eventually exceed the dielectric strength of the oxide and lead to rapid decomposition of the insulating layer.

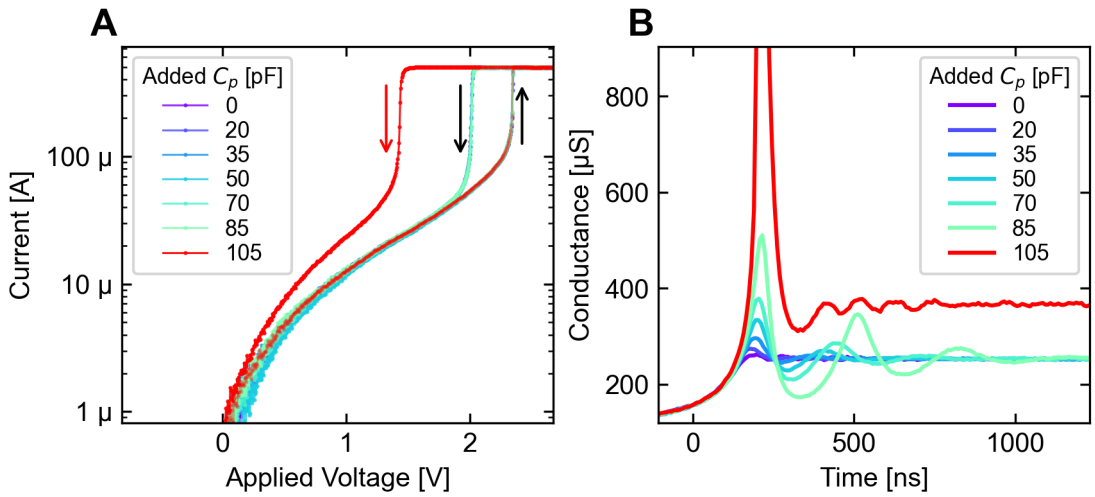

Figure 4.17: By adding capacitance to the measurement circuit, volatile TS eventually gives way to NV resistive switching in a $(V_{0.85}Cr_{0.15})_2O_3$ nanodevice ($350\times350\times30$ nm^3). The measurements used triangular voltage sweeps (3 V) with 1 ms duration and CLA current bias 500 µA. **(A)** A standard (I, V) plot retraces the same reversible TS curve for $C_p \leq 85$ pF, but a NV event is recorded with $C_p = 105$ pF. **(B)** As the capacitance increases, high speed device conductance measurements at the switching threshold show oscillatory signals with increasing amplitudes. With $C_p = 105$ pF, the initial current spike is followed by a permanant shift in device conductance.

Even with a fixed, low value of parasitic capacitance, differences in measurement procedure can also give rise to sudden NV events. An important instance of this is the effect of the applied voltage sweep rate in a series resistance control circuit. For sufficiently large R_s and sufficiently low sweep rates, the steady-state NDR curve can be retraced on the increasing and decreasing voltage sweep. However, due to the limited dynamical speed of the conductance change in the switching material, faster sweep rates lead to a rate-dependent hysteresis that brings the device (I, V) switching trajectory progressively further from the NDR curve as the state variable lags the input driving signal (see Fig. 4.18). As the cell is pushed further out of equilibrium by higher sweep rates, eventually a spiking threshold can be crossed. If the induced spike is large enough, it is accompanied by a rapid NV resistive switching process, resulting in a decreased resistance state. The (I, V) trace of this newly created state is similar to the previous example, reached by increasing the parasitic capacitance during threshold switching. This interesting, counter-intuitive result of increasing the sweep rate implies that characterizations using fast driving signals such as square pulse waveforms (where the sweep rate approaches infinity) are prone to cause NV events, even where a sweep to the same amplitude with 1,000$\times$ longer duration does not.

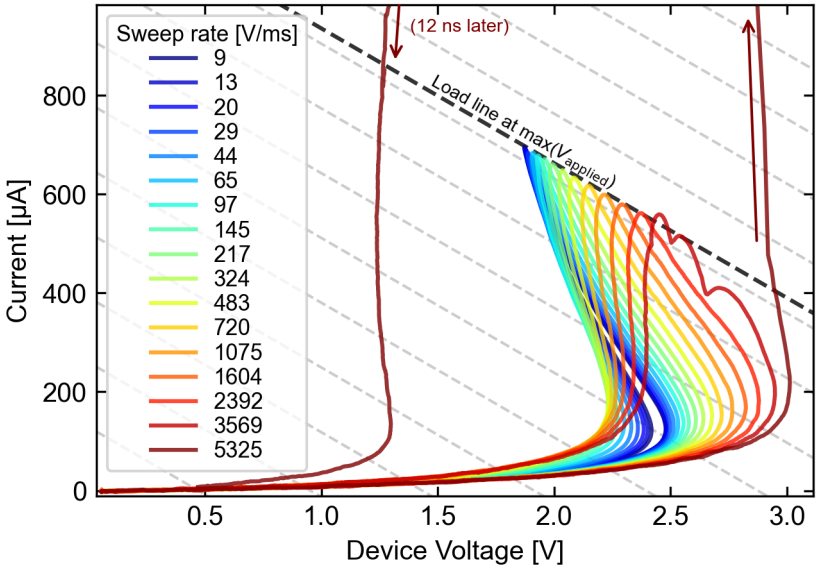

Figure 4.18: Voltage sweeping measurements of an as-deposited $(V_{0.85}Cr_{0.15})_2O_3$ device ($350\times350\times30$ nm^3) using the digipot characterization circuit with $R_s = 3.7$ kΩ and 4.5 V triangular voltage waveforms. Sweeping to the same maximum applied voltage, device (I, V) trajectories for low voltage sweep rates (ms timescale) trace the characteristic NDR curve, while trajectories for faster sweep rates diverge from it and eventually trigger a runaway transition (μs timescale). A major departure of the trajectory from the measurement load line shows that a ~10 ns current spike originated from discharge of the parasitic capacitance.

In addition to sudden NV events that were shown to coincide with transient spikes in the device current, there is another distinct way NV changes can be induced in the samples. This second mode can be achieved by extended stable biasing of a device near or above the NDR threshold, and contrasts with the former mode in that the NV shifts occur over a relatively long timescale and allow a gradual transition to lower resistance levels. To control the gradual transition mode above the NDR knee, biasing needs to be done either with high enough series resistance to suppress runaway load-line transitions, or using a current source with low parasitic capacitance. By continuously sweeping the bias at low speed and repetitively retracing NDR curves, we demonstrate this gradual NV switching mode for different starting film thicknesses (Fig. 4.19). The gradual resistance switching observed here is comparable to a constant current stress degradation, frequently studied in context of dielectric breakdown. However, capturing the full (I, V) curves during the process provides additional information about the nature of the state transition.

The device states that follow sudden or gradual NV transitions typically possess a modified NDR characteristic, distinct from the original NDR curve of the pristine state. By comparing these curves for the sudden and gradual cases, it is apparent that these modes are distinguished not only by their

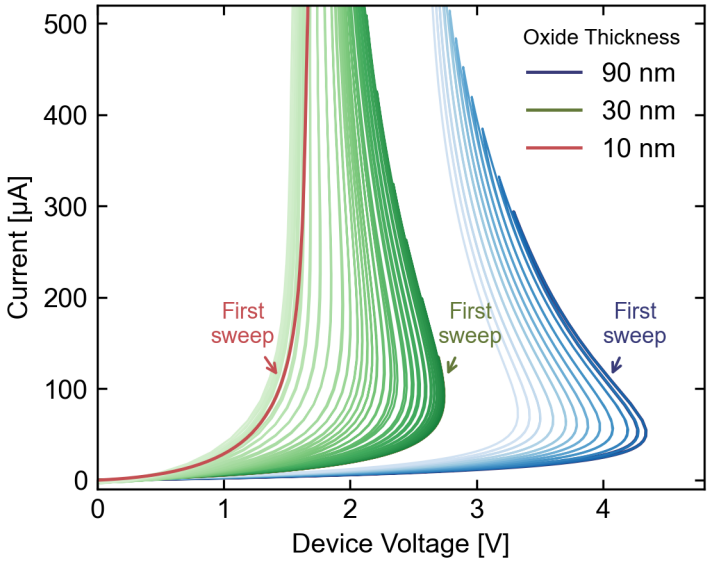

Figure 4.19: Gradual NV switching of $(V_{0.85}Cr_{0.15})_2O_3$ devices (device width 250 nm) induced by slow sweeping measurements along the NDR curves. The sweeps, lasting three seconds each, were performed with 8.2 kΩ series resistance and with different device thicknesses. Each repeated sweep causes a small NV modification of the NDR curve — an effect which is qualitatively similar to a reduction of the starting oxide thickness.

speed, but they also produce notably different states. In each case, the modified NDR curves have progressively lower NDR threshold voltages and higher sub-threshold leakage currents. However, for the sudden case, the device power at the new NDR onset point is significantly reduced relative to the pristine state (visible in Fig. 4.17 and 4.18) whereas for the gradual case, the NDR onset power remains relatively constant as the device switches. According to the interpretation to be further discussed in the modeling section 4.5.2, the appearance of these different NDR curve shapes can be attributed to a difference in the spatial profile of the conductivity redistribution in the two cases. Here, we initially point out experimental clues that the sudden NV events are associated with filamentary conduction paths while the gradual NV switching results in the growth of a relatively uniform high conductivity layer across the entire device area.

With the initial device resistance inversely proportional to device area (recall Fig. 4.7), the standard electrical evidence for filamentation is the production of formed states across devices of different areas showing resistance levels independent of area [232, 233]. This highly simplified picture rests on the assumption that filamentary growth (and thus the final resistance state) is not influenced in any way by the total device area. However, area dependence of the formed state can enter in various ways such as through its impact on the conditions required to initiate the forming process [70], by its connection to the energy stored in the device self-capacitance [110], and by conduction through

the residual unformed area of the cell. The analysis is also complicated by the possibility of creating multiple filaments [234], and the fractal dimension of the percolation process [235]. Nevertheless, sub-threshold resistance measurements before and after electroforming can be used to give experimental support for the presence of a filament in the device.

As a first approximation, consider the total resistance of memory cell (assumed uniform in thickness) if a filament of area A_2 is embedded inside a larger device area A_1,

$$R_{\text{formed}} = \left(\frac{\rho_1 t_{\text{ox}}}{A_1 - A_2}\right) \parallel \left(\frac{\rho_2 t_{\text{ox}}}{A_2}\right) = \frac{\rho_1 \rho_2 t_{\text{ox}}}{A_2 \rho_1 + (A_1 - A_2)\rho_2}, \tag{4.6}$$

where ρ_1 and ρ_2 are the resistivities of the pristine material and of the filament, respectively. As shown in Fig. 4.20, measurements of the sub-threshold resistance of $(V_{0.85}Cr_{0.15})_2O_3$ cells before and after a voltage-controlled and current-limited electroforming step closely follows this functional form of area dependence. This suggests that filamentary conduction paths were created as a consequence of the sudden NV resistance switching induced by this unstable measurement configuration.

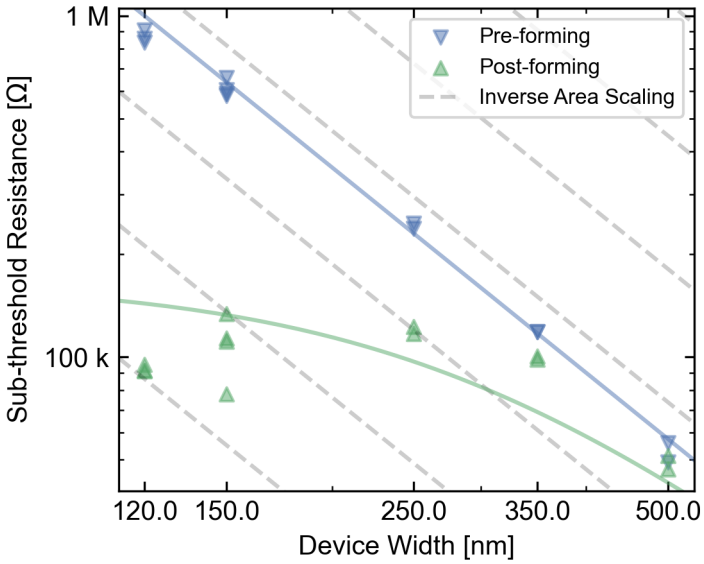

Figure 4.20: Sub-threshold (500 mV) resistance levels of $(V_{0.85}Cr_{0.15})_2O_3$ devices (30 nm thickness) before and after a sudden forming event induced using the CLA current limiting circuit. The device area dependence of the resistance after forming shows a filamentary signature, approximately following Eq. 4.6.

For the case of the gradual NV process, a similar analysis based on the area dependence of the resistance levels is not possible, as there are no two distinct before/after states to be identified for comparison. Instead, switching continuously progresses as the bias level or the biasing time is increased. However, the effect of this gradual process on the NDR curve is informative when compared

between different device thicknesses. Starting from a relatively higher thickness, the initial NDR curve eventually develops an NDR signature comparable to a pristine film with lower thickness (Fig. 4.19). A simple explanation for this NDR curve evolution is that a uniform, low resistivity layer grows gradually across the whole device area. This result indicates that a control circuit enabling stable biasing near the device NDR curve leads to quite different conductivity distributions after electroforming than in the unstable case, and the question of electronic stability could mean the difference between filamentary and non-filamentary NV switching in the same device material system.

4.4.3 Switching dynamics

The speed of conductance switching is clearly important for device applications, due to its implications for memory read/write times and energies and its impact on the overall performance of physical neural networks. Switching speed measurements can also give clues about the physical mechanisms involved in the switching process. As seen previously in Fig. 4.18, $(V_{1-x}Cr_x)_2O_3$ devices in the pristine state show strong differences in the measured (I, V) trajectories when operated on millisecond and microsecond timescales. Generally, switching dynamics are instead studied by application of square voltage pulses directly to the cell of interest. In these measurements, a delay time before a relatively shorter transition time is seen very broadly across many different classes of resistive switching devices [171, 236]. Measurements and simulations of time delay behavior in different types of electro-thermal switches when subject to voltage pulses have been reported [39, 40], and the specific shape of transition has even been used as evidence for the type switching mechanism [237].

Here we show that for the same type of device, several distinct dynamic behaviors can be induced by voltage pulses, depending on the driving signal levels and circuitry external to the cell. The variety of dynamical behaviors are consistent with and largely predictable by a simple electro-thermal model as elaborated in the following section (4.5). We also show a large reduction in the characteristic switching times of formed device state relative to the pristine, which is later understood to originate from switching in a decreased material volume with a reduced heat capacity.

The result of a direct voltage pulsing experiment on a pristine $(V_{0.85}Cr_{0.15})_2O_3$ cell is summarized in Fig. 4.21. The measured switching delay times span four orders of magnitude from 100 ns to 1 ms across only a 100 mV range of pulse amplitudes. The transition times were approximately one order of magnitude lower than the delay times for each pulse amplitude. The CLA (section 2.2) was used here to halt the runaway transition at 480 μA, but its effect in the circuit is negligible when the device current is significantly below the limit. However, the specific configuration of the CLA during the measurement limited the rise time of current readout to around 100 ns, and the transition times are expected to be even faster as the voltage amplitude is further increased.

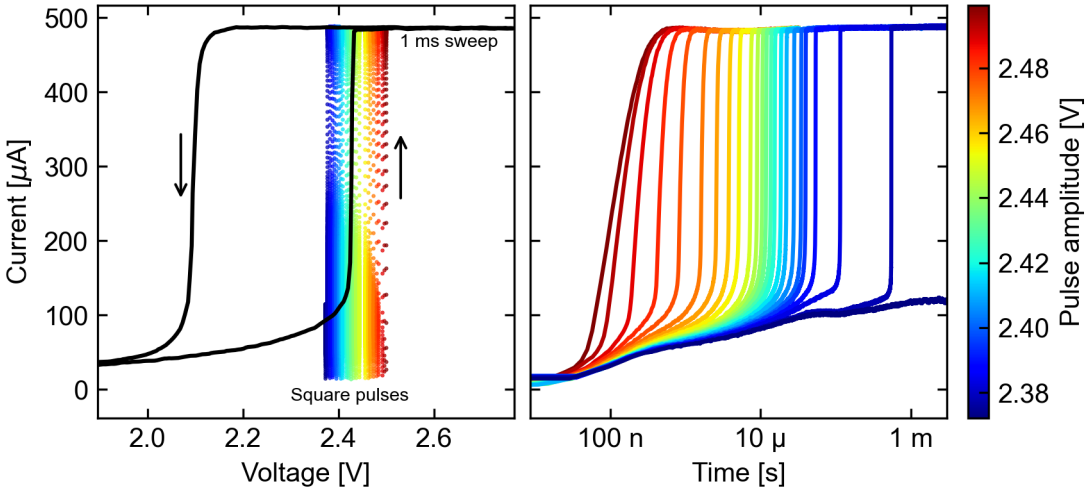

Figure 4.21: Switching speed (current vs. time) measurements for a $500\times500\times90$ nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ device subject to square voltage pulses with different amplitudes. Current is limited to 480 μA by the CLA probing circuit. A 1 ms duration triangle sweep measurement is also superimposed on the left plot for reference. A switching delay time is observed that sensitively depends on the voltage amplitude near the threshold voltage.

By measuring a similar cell through an on-chip series resistance, as shown in Fig. 4.22(A,B,C), we demonstrate that the switching delay behavior seen previously in Fig. 4.21 is a matter of the "angle of attack" of the load line to the cell's NDR curve, which is also anticipated by the electro-thermal theory (recall Fig. 1.14). In contrast to the delay-transition behavior seen *without* added series resistance, the conductance transition signals *with* added series resistance follow an approximately exponential dependence in time, with time constants independent of the pulse amplitude. This shows that the delay-transition behavior is specific to using a control circuit with current-voltage feedback that constrains the switching trajectories to first pass nearby, and then away from the NDR condition where the the conductance change is slowed. By approaching the NDR curve with a trajectory that moves monotonically nearer to it, as in the shown series resistance case, the initial switching delay is eliminated.

An important experimental result also shown in Fig. 4.22 is that after electroforming, the timescale of switching is dramatically reduced relative to switching in the pristine device state. Using voltage pulses, formed devices typically support threshold switching on the 10 ns scale or below. Another way to compare the switching timescales before and after forming is by looking into the sweep response with a series resistor, as shown in Fig. 4.23. For a 100 ns sweep, the (I, V) trace opens up slightly as the device current begins to lag the applied voltage, but the dynamic threshold voltage is less than 100 mV larger than the NDR knee voltage and similar currents are reached at the maximal point of the sweep. This should be compared with sweep rate dependence of a pristine device, previously shown in Fig. 4.18, where a dramatic departure from the NDR curve was induced by an order of

magnitude lower sweep rate.

Threshold switching is not only much faster in formed devices than before forming, but it also remains very stable and repeatable. The increased speed of switching makes it practical to test the endurance beyond the level shown for the pre-formed device (recall Fig. 4.13). In [44], we showed an example of such a formed device enduring over 10^{12} cycles with little sign of degradation, and less than 5% variability and drift in the threshold voltage. However, as with any endurance-type measurement, this specific number of endured cycles is expected to strongly vary with the type of control exerted on the cell [64]. While the endurance measurement is a promising result, it can not be considered a fixed property of the $(V_{0.85}Cr_{0.15})_2O_3$ material.

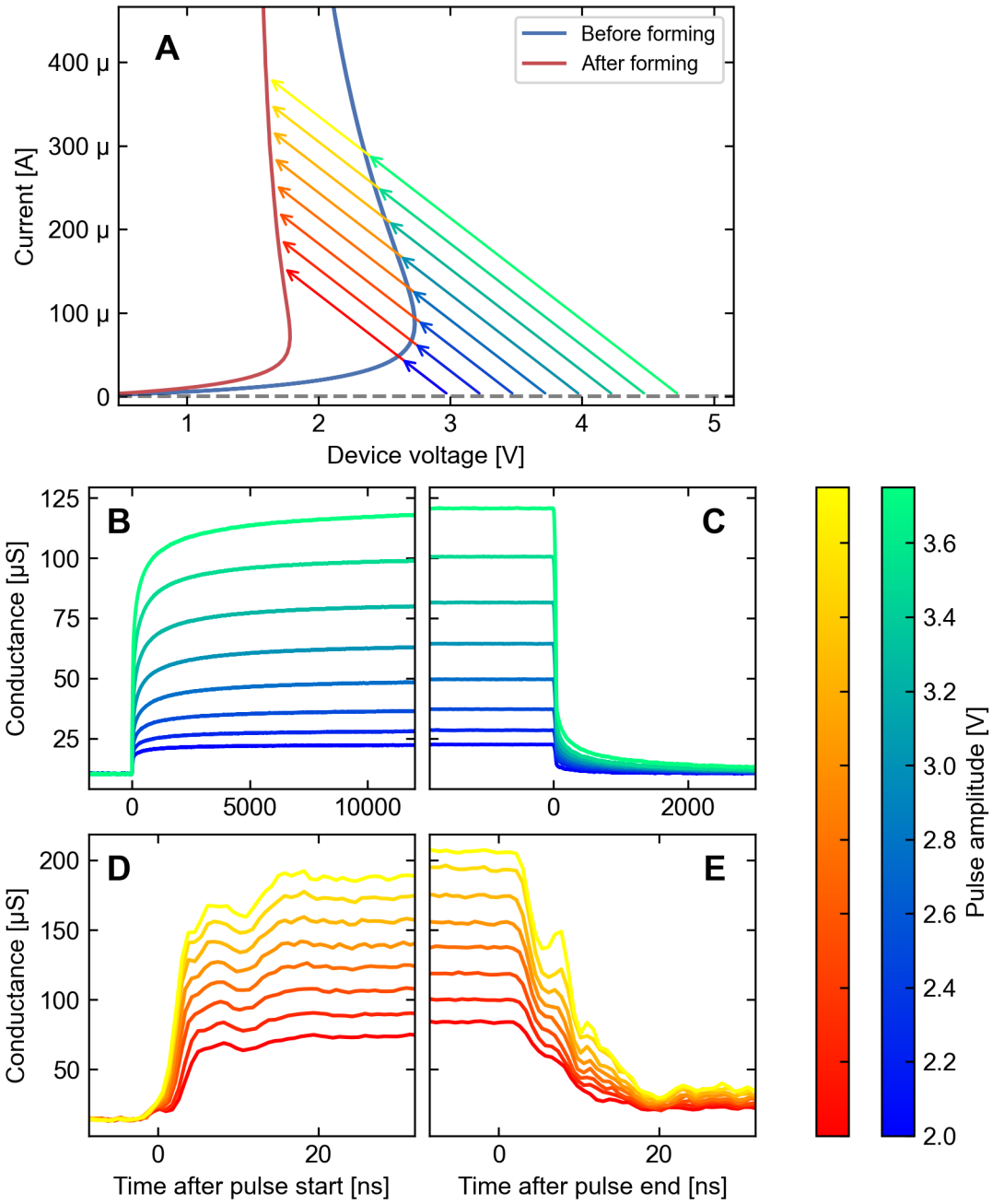

Figure 4.22: Measurements of the conductance switching dynamics of a 250×250×30 nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ device using an 8.2 kΩ on-chip series resistance before **(B)**,**(C)** and after **(D)**, **(E)** (abrupt-type) electro-forming. **(A)** shows the (I, V) switching trajectories enforced by the series resistance circuit, each approaching the respective (pristine, formed) measured NDR characteristics. The same set of voltage pulse amplitudes were used in both cases, with the square pulse waveforms beginning and ending at 1 V to allow continuous conductance measurements. The transitions effectively have a time constant that does not depend on the pulse amplitude, but that is dramatically reduced after forming.

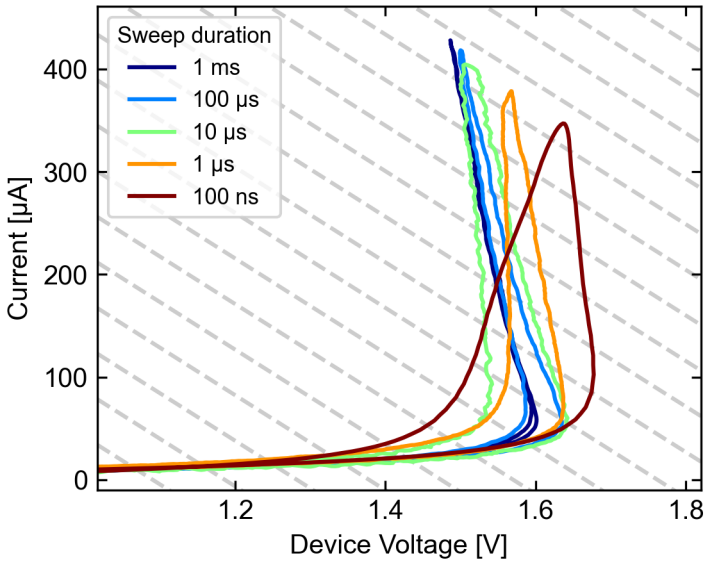

Figure 4.23: The effect of voltage sweep rate on the measured switching trajectories in a 250×250×30 nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ device with on-chip series resistance 2.2 kΩ after abrupt-type electroforming. Threshold switching persists even at 100 ns timescales. © 2018 IEEE.

4.4.4 Oscillatory/spiking behavior

Due to the reduced switching time and energy after electroforming, formed cells are prone to oscillation during (I, V) measurement, especially when controlled with external circuitry (as opposed to on-chip resistors). These nonlinear relaxation oscillations arise because of the stray capacitance in parallel with the cell, which can store enough energy to significantly heat the cell even for small values below 10 pF. One of the conditions to observe the oscillations is for the cell to be biased with a fixed point in a sub-range of the NDR region, which requires that the capacitor/device combination is driven through a large enough series resistance, or by a current source. The biasing range where free-running oscillations can be induced depends on the value of the parallel capacitance C_p (recall Fig. 1.17), and for any sufficiently large C_p, the corresponding limit cycle always encircles the fixed point that would otherwise be a steady state.

Oscillations were experimentally induced in electroformed $(V_{0.85}Cr_{0.15})_2O_3$ devices by applying voltage pulses and using an external series resistance (Fig. 4.24). As the critical biasing amplitude for sustained oscillation is approached from below, the measured (I, V) trajectories spiral into the fixed point and eventually settle there. Applying voltages nearer to the border of stability, or the Hopf-bifurcation level, the transient oscillatory signal lasts longer and is highly susceptible to noise from thermal fluctuations. Finally, when the pulse amplitude is in a suitable range, limit cycles are reached after a period ($\sim$10 µs) of relaxation from the initial spiking transient.

Experimental (I, V) limit cycles for different bias levels were measured and are shown in Fig. 4.25. These show an oscillation frequency and amplitude that depend on the voltage bias level, the parallel capacitance and the series resistance. While the frequency only increases with the applied voltage, the amplitude increases to a maximum level in the middle of the unstable biasing range, before reducing again and eventually reaching another non-oscillatory stable point further up the NDR curve. Oscillations frequencies were observed in this system above 300 MHz when using a discrete surface mount series resistor to further minimize the parasitic capacitance to a sub-picofarad level. In Section 4.5.3, an electro-thermal model provides very close agreement to the measured voltage dependence of the oscillation amplitude and frequency.

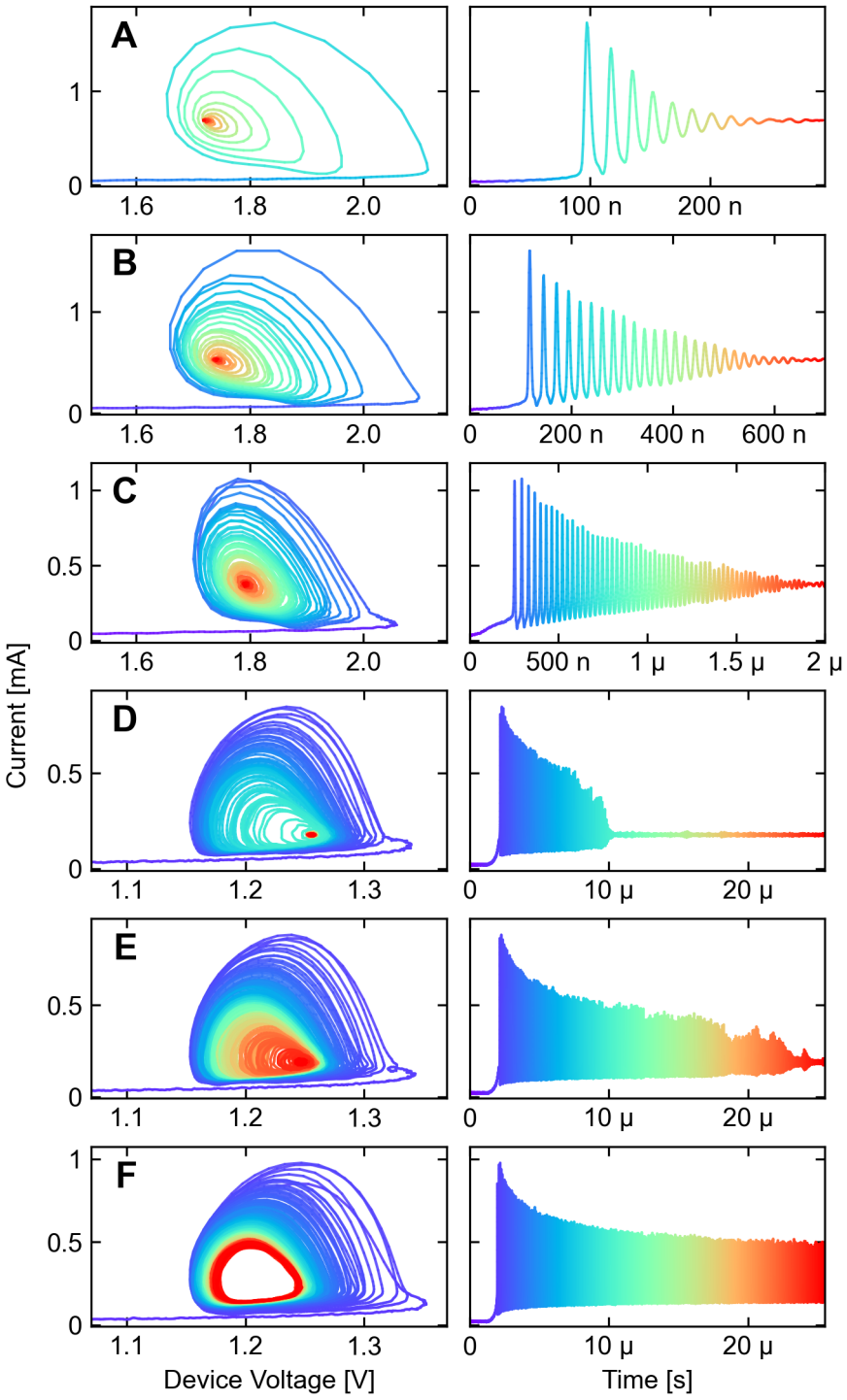

Figure 4.24: Oscillatory transients measured in formed $(V_{0.85}Cr_{0.15})_2O_3$ devices in response to square voltage pulses (100 ns rise time) with increasing amplitude. The progression from subplots **(A)** to **(F)** represents an increasing voltage pulse amplitude, where larger amplitudes lead to longer decaying transients before stable limit cycles **(F)** are eventually reached. Data is shown from two separate devices, both with 30 nm thickness: **(A, B, C)** with width 250 nm, R_s=1.1 kΩ, and **(D, E, F)** with width 150 nm, R_s=4.5 kΩ.

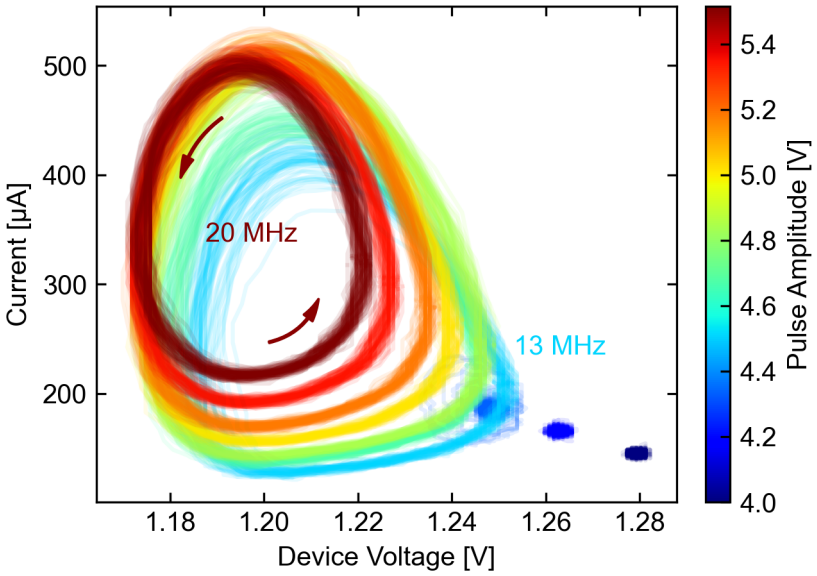

Figure 4.25: Current-voltage limit cycles reached in a formed $150\times150\times30$ nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ device during voltage pulses of different amplitude. Measurements were made with the digipot circuit with series resistance setting 4.5 kΩ. Oscillation frequencies were in the 10 MHz range and increased with the applied voltage.

4.5 Electro-thermal modeling

4.5.1 Scaling of the NDR steady state

In this work, devices were fabricated down to 120×120×5 nm^3 size, making them the smallest nano-devices of $(V_{1-x}Cr_x)_2O_3$ reported to date. However, memory and neuromorphic applications will require even smaller dimensions to achieve high density and 3D integration. Thus, it is important to study the impact of scaling on the device operation. Our fabricated devices varied in size in two dimensions (Fig. 4.26), and we observed experimentally that reducing the width and thickness had roughly opposite effects on the NDR knee voltage (Fig. 4.14). In these samples, we see leakage currents in the 10 µA range, a figure which should be reduced for lower power consumption. While it is expected that scaling down the device area will also reduce the leakage, it is questionable whether it is also possible to maintain or improve the overall device functionality through scaling. Here we analyze the collected data to build an electro-thermal model that includes size dependencies in order to predict the scaled behavior.

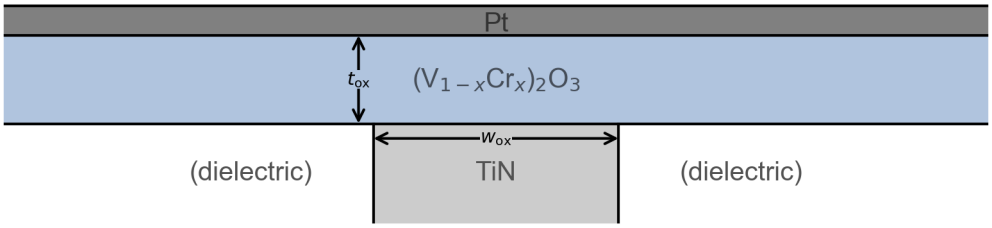

Figure 4.26: Diagram showing the varied dimensions of the thin film nano-devices: the thickness (t_{ox}) and the width (w_{ox}). Devices have a square cross-section and the width of the third dimension is also w_{ox} (not drawn).

In section 4.4.1, the device size dependence on the current was verified by resistivity measurements of many devices across the wafers (Fig. 4.7). We found that the resistance at low voltage was proportional to oxide thickness and inversely proportional to the device area. Furthermore, we obtained very close agreement to temperature dependent measurement data using the empirical relation given in Eq. 4.4. These findings justify the use of a bulk limited conduction model where microscopic current density is given by

$$j(E,T) = \alpha E \exp\left(\frac{-b}{T}\right) \exp\left(\gamma\sqrt{E}\right), \tag{4.7}$$

and the total macroscopic device current scales as

$$I(V,T,t_{ox},w_{ox}) = \alpha w_{ox}^2 \frac{V}{t_{ox}} \exp\left(\frac{-b}{T}\right) \exp\left(\gamma\sqrt{\frac{V}{t_{ox}}}\right), \tag{4.8}$$

where α, b, and γ are fit parameters, and w_{ox}, t_{ox} are the device width and thickness.

Considering the thermal part of the model, we use a thermal resistance approximation for the steady-state temperature,

$$T = T_0 + R_{th}(w_{ox}, t_{ox})IV. \tag{4.9}$$

Different scaling behaviors of the thermal resistance are possible depending on the relative thermal conductivities and thermal boundary resistances of the various materials and interfaces present. It was noted in the experimental section 4.4.1 that the NDR knee power as the device dimensions varied mainly depended on the device width, with a much lower contribution from the thickness. Gibson et al [161] derived that when the functional form of the conduction equation has separable V and T components, as is the case in Eq. 4.8, the electro-thermal NDR knee always occurs at the same temperature difference ΔT above the ambient T_0, and ΔT depends only on the activation energy b. This makes the knee power an experimentally accessible indication of thermal resistance, and shows that it mainly scales with the device width in our case.

Another way to analyze the measured NDR characteristics is to explicitly fit the curves to equations 4.8 and 4.9 with R_{th} as a free parameter. The results of independent fits (Fig. 4.27) show that the equations fit each of the NDR curves remarkably well and also produce an approximately linear dependence of the thermal resistance on the inverse of the device width. An interpretation of this type of width scaling and the relative absence of thickness dependence is that the heat losses out of the cell are dominated by the electrode contacts, and the lateral loss through the sides of the low thermal conductivity oxide itself is negligible. This reasoning also explains why changing the top electrode thickness had a strong effect on the NDR curve (shown previously in Fig. 4.15). The scaling factor $1/w_{ox}$ approximates the spreading resistance and holds for contacts to semi-infinite media as well as finite plates [238–241]. This scaling behavior of R_{th} has been assumed in other works [140, 242], though different scaling behaviors are sometimes used in other situations [107, 212, 243, 244].

To fit the entire size-dependent dataset with single shared parameter set, we use the approximation

$$R_{th}(w_{ox}, t_{ox}) = R_{th,0} + \frac{R_{th,w}}{w_{ox}} \tag{4.10}$$

where the parameters $R_{th,0}$ and $R_{th,w}$ are subject to fitting. To account for a slight thickness gradient or interfacial layer in the deposited oxide material, we also allow a small linear variation of the γ parameter with thickness,

$$\gamma(t_{ox}) = \gamma_0 + \gamma_t t_{ox}, \tag{4.11}$$

where γ_0 and γ_t take the place of γ as fit parameters. This adjustment is meant to improve the fit to the experimental data at large thicknesses, and is not extrapolated further in the following scaling projection.

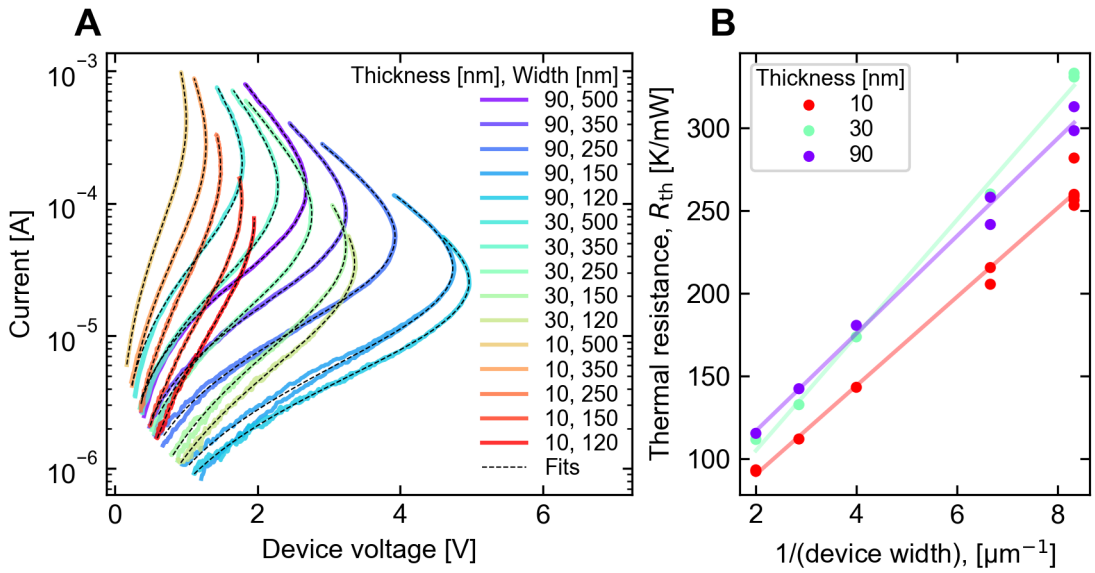

Figure 4.27: **(A)** Measurements and fits of NDR characteristics for $(V_{0.85}Cr_{0.15})_2O_3$ devices with different widths (120 nm to 500 nm) and thicknesses (10 nm to 90 nm). Fits to each curve are independent to accommodate device-to-device variability. **(B)** shows the R_{th} values corresponding to the best fit, which grows linearly with $1/w_{ox}$.

The complete set of fitted parameter values for the NDR scaling model are given in Table 4.2.

Using the parameters that fit our NDR measurements, we varied the device dimensions in our model to obtain the projected behavior for scaled devices (Fig. 4.28[A]). We found a generally favorable scaling properties, with the half-threshold leakage dropping to the 100 nA level while the threshold voltage stays below 3 V for a $10\times10\times10$ nm^3 device. Increasing the thermal isolation by reducing the top electrode thickness (Fig. 4.28[B]) leads to a strong reduction in the NDR onset voltage, giving additional possibilities for controlling device operation via tuning of the thermal environment.

Table 4.2: Parameters for the NDR scaling model.

Symbol	Value	Unit
α	7.378×10^4	S m^{-1}
b	3.423×10^3	K
γ_0	1.382×10^{-4}	m$^{1/2}$ V$^{-1/2}$
γ_t	2.424×10^3	m$^{-1/2}$ V$^{-1/2}$
$R_{th,0}$	4.495×10^5	K W^{-1}
$R_{th,w}$	2.235×10^{-2}	m K W^{-1}

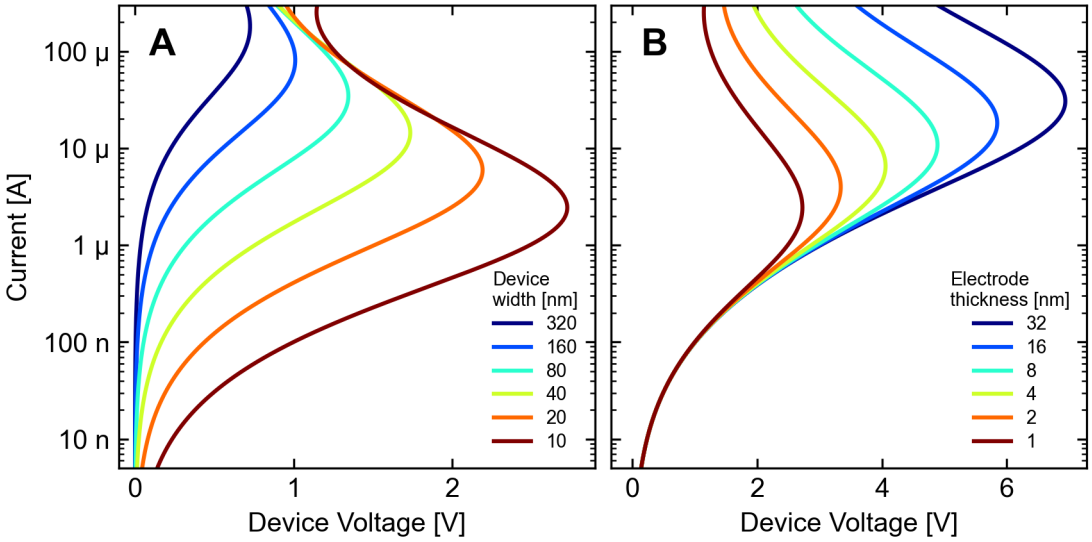

Figure 4.28: Simulated effect of scaling on the NDR curves of $(V_{0.85}Cr_{0.15})_2O_3$ devices. In **(A)**, the width of a device with 10 nm thickness is scaled. Leakage currents are drastically reduced while the NDR onset voltage remains in a usable range. In **(B)**, the effect of reducing the electrode thickness, or equivalently reducing the heat dissipation in a $10\times10\times10$ nm^3 device.

4.5.2 Non-volatile model

In section 4.4.2, a variety of electrical situations were identified to induce non-volatile (NV) shifts in the (I, V) characteristics of $(V_{1-x}Cr_x)_2O_3$ cells. The NV effects apparently set in when the device already has its conductivity affected by elevated temperature and electrical conditions are such that critical levels of device voltage and current are exceeded for a sufficient amount of time.

Experimentally, we have demonstrated two distinct NV switching modes:

1. a "sudden" mode resulting from electrical instability that is associated with production of filamentary conduction paths, and
2. a "gradual" mode operating at long timescales (above approximately 1 ms) with stable biasing, which appears to reduce conductivity uniformly across the whole device area.

When efforts are made to quickly interrupt the NV process, intermediate states can be produced that also have a reversible (volatile) regime with electrothermal character. In progressive stages of the process, modified NDR curves are measurable that are different from the NDR curve of the pristine state. Here, we present a simple model for the electrical effect of the NV process to be compared with measurement data to further support the idea that different filament sizes can be created in the same devices using the two different NV modes. Using this model, we show a readily measurable fingerprint in the

NDR curves of formed devices that is indicative of the filamentary radius, and give further implications for scaled devices.

Many mechanisms for local conductivity changes resistive switches exist and have been extensively classified depending on the electrodes and the switching material [16]. For the purposes of the following argument we remain indifferent about the physical origin of the conductivity changes during forming and only consider the effect of an assumed filament size and shape on the volatile device operation. One possibility is that after forming, a complete filament is newly created that extends through the entire oxide thickness. This secondary phase or reduced oxide could have volatile switching properties that were not already present in the pristine device, such as an electronic IMT, and similar experimental findings in related materials (NbOx, VOx, TaOx) have been interpreted along these lines [90, 245–247]. However, we can also consider what happens when a conductive filament incompletely bridges the oxide thickness. Whatever its physical origin, it is reasonable to expect that if a filament continuously grows from one electrode to the other and a limitation is placed on the energy delivered during the process, a "gap" of the original switching material remains before the oxide is eventually bridged.

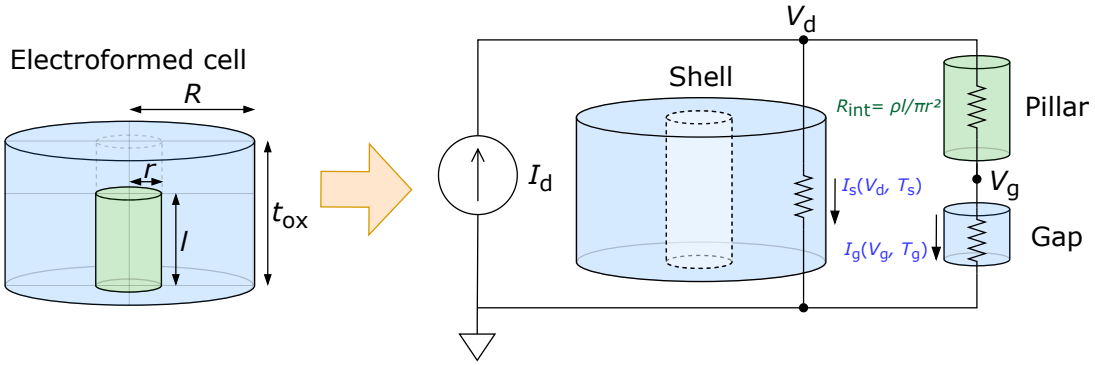

Figure 4.29: A simple lumped-element model for non-volatile states in a $(V_{1-x}Cr_x)_2O_3$ cell. In a formed cell, three circuit elements consisting of subvolumes of the cell are isolated and considered separately. A resistive "pillar" with a certain radius, uniformly high conductivity, and no temperature dependence partially bridges the oxide thickness and confines a new switching volume in the "gap". Two separate volumes of as-deposited material (the "shell" and "gap") each have their own temperature states T_s and T_g.

The conductance distribution after forming could have a variety of shapes in the radial or axial dimensions, but for convenience and without much loss of generality, we suppose a cylindrical conductive "pillar" is formed in the center of the device leaving a gap between it and one of the electrodes. We assume that inside the pillar, the resistivity has a uniformly low value ρ, and that it has no state variable and does not participate in switching. The effect of this inactive pillar is a substitution of a small internal series resistance and to confine a new effective switching volume composed of the same material as

the pristine device. A diagram and circuit approximation for the formed cell with a specific pillar length l and radius r embedded in a device with thickness t_{ox} and radius R is shown in Fig. 4.29.

The equation system that implicitly identifies the steady state of the composite system driven by a current source (I_d) consists of the circuit equations from Kirchhoff's laws,

$$I_d = I_s + I_g, \tag{4.12}$$

$$V_g = V_d - I_g \frac{\rho l}{A_g}, \tag{4.13}$$

and the conduction equations for the subvolumes,

$$I_s = \alpha A_s \frac{V_d}{t_{ox}} \exp\left(\frac{-b}{T_s}\right) \exp\left(\gamma \sqrt{\frac{V_d}{t_{ox}}}\right), \tag{4.14}$$

$$I_g = \alpha A_g \frac{V_g}{t_g} \exp\left(\frac{-b}{T_g}\right) \exp\left(\gamma \sqrt{\frac{V_g}{t_g}}\right), \tag{4.15}$$

$$T_s = T_0 + R_{th,s} I_s V_d, \tag{4.16}$$

$$T_g = T_s + R_{th,g} I_g V_g, \tag{4.17}$$

$$R_{th,s} = \frac{R_{th,w}}{R\sqrt{\pi}}, \tag{4.18}$$

$$R_{th,g} = \frac{R_{th,w}}{r\sqrt{\pi}}, \tag{4.19}$$

which are found by applying the scaling equations 4.8 and 4.9. Here, the subscripts $_s$ and $_g$ refer to the variable belonging to the shell and gap volumes, respectively, and the geometrical factors are defined as

$$A_s = \pi(R^2 - r^2), \tag{4.20}$$

$$A_g = \pi r^2, \tag{4.21}$$

$$t_g = t_{ox} - l. \tag{4.22}$$

Solving the above equation system numerically for varying pillar dimensions l and r (Fig. 4.30), we find that formed steady states have an NDR region for all but the thinnest gap lengths below ~5 nm, and the NDR onset voltage always reduces relative to the pristine state. The main difference in the calculated NDR curves is that for small pillar radius ratios r/R, the current and power at NDR onset reduces as the pillar length increases, whereas for $r/R = 1$, the NDR power remains constant. This suggests that the power shift at the NDR knee is a measurable indication of the degree of filamentation in a formed cell. Comparing with the measurements after non-volatile shifts in Section 4.4.2, this simple model shows that the differences in the shapes of the NDR curves in the gradual and sudden forming cases can be attributed to different radii of the formed pillars.

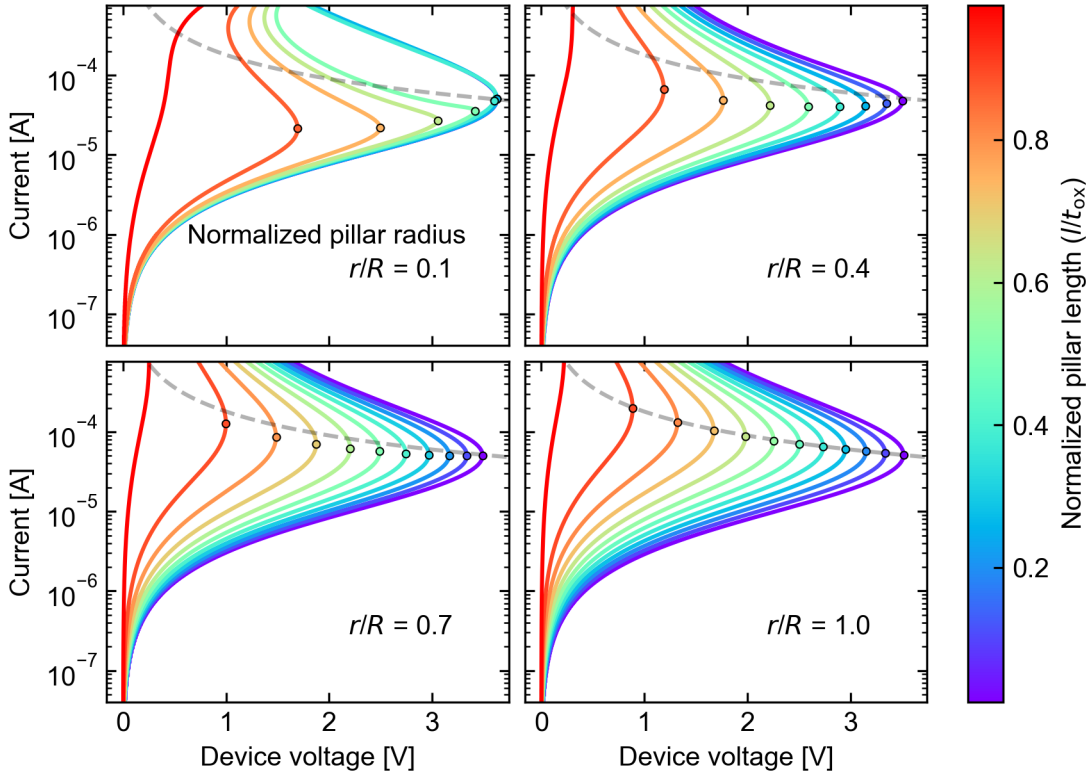

Figure 4.30: The simulated effect of the gradual growth of a conductive pillar of different radii, r, inside an electro-thermal cell of radius R. For small r/R, the NDR power (marked for $l = 0$ with a dashed line) reduces as the pillar grows, but for $r = R$, the NDR power stays the same. Here, the pillar resistivity was given an arbitrary value of $\rho = 10^{-5}$ Ωm.

In Section 4.4.3 we demonstrated faster dynamics of a filamentary formed state relative to the pristine state. This finding is also explained by the pillar-gap model when considering as an approximation that the heat capacity of each region scales with volume. The dynamical equations for the rates of temperature change for the shell and gap are

$$C_{\mathrm{th,s}}\frac{dT_\mathrm{s}}{dt} = I_\mathrm{s}V_\mathrm{d} - \frac{T_\mathrm{s}-T_0}{R_{\mathrm{th,s}}}, \tag{4.23}$$

$$C_{\mathrm{th,g}}\frac{dT_\mathrm{g}}{dt} = I_\mathrm{g}V_\mathrm{g} - \frac{T_\mathrm{g}-T_\mathrm{s}}{R_{\mathrm{th,g}}}, \tag{4.24}$$

$$C_{\mathrm{th,s}} = C_{\mathrm{th,v}}A_\mathrm{s}t_\mathrm{ox}, \tag{4.25}$$

$$C_{\mathrm{th,g}} = C_{\mathrm{th,v}}A_\mathrm{g}t_\mathrm{g}, \tag{4.26}$$

where $C_{\mathrm{th,v}}$ is the volumetric heat capacity of the oxide material. As the gap becomes thinner, the shell is less involved in the switching and the relevant thermal time constant is $R_{\mathrm{th,g}}C_{\mathrm{th,g}}$, which scales down with the gap size as rt_g. As we have seen in Section 1.4.2, the speed of switching in an electrothermal

cell can span orders of magnitude because of circuit constraints on the heating efficiency, but in each case the total heat capacity acts to directly scale the heating speed, and therefore the switching speed.

Insofar as the assumptions of the non-volatile model presented here reflect reality, there are also interesting implications for scaled devices. When the gap after forming is small, its contribution dominates the shape of the (I, V) curves; the shell only contributes to the total leakage and does not heat up significantly. Therefore, highly scaled cells can be expected to have similar characteristics as measured in formed cells but with lower leakage, and importantly with no forming step required [44]. Curve fitting (I, V) measurement data to the model provided here could in principle provide an estimated size and shape of the gap volume in question, but due to the various assumptions made, this approach may have limited accuracy. Nonetheless, we expect the NDR power in formed devices to be correlated with the filament sizes, making relative comparisons possible.

4.5.3 Oscillatory model

For realistic investigations of neural networks incorporating compact oscillatory and/or spiking devices, a model is needed that closely matches the measurement data of real devices. For this purpose, an oscillatory dataset was collected using a $150\times150\times30$ nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ device with the voltage applied to the top electrode through an external 10 kΩ series resistance soldered at the base of a non-coaxial probe needle. The cell was electroformed through a circuit instability (saddle-node bifurcation) during an initial 1 ms sweep to 4 V, and the oscillatory current signal was measured in the formed state with a bandwidth of 350 MHz and sample rate of 2.5 GHz while applying a second 1 ms sweep to 3.5 V. In this dataset, oscillations were recorded for applied voltages in the range 2.5 V to 3.5 V. As the voltage increased, the oscillation frequencies increased from 30 MHz to 70 MHz and the current amplitudes also increased from 200 μA to 600 μA.

For fitting the measurement data, we adopt a slightly simplified version of the pillar/gap model from the previous section where heating in the pristine shell region is ignored and replaced by a static resistor with value R_p. A circuit schematic approximating the measurement setup is shown in Fig. 4.31. In this measurement, we assume that C_p is dominated by the probe needle interconnect and not the device itself, so that I_C is transmitted back to the AWG source without being measured at the opposite device terminal. The device current signal, measured through a scope channel's 50 Ω shunt to ground, is denoted I_d.

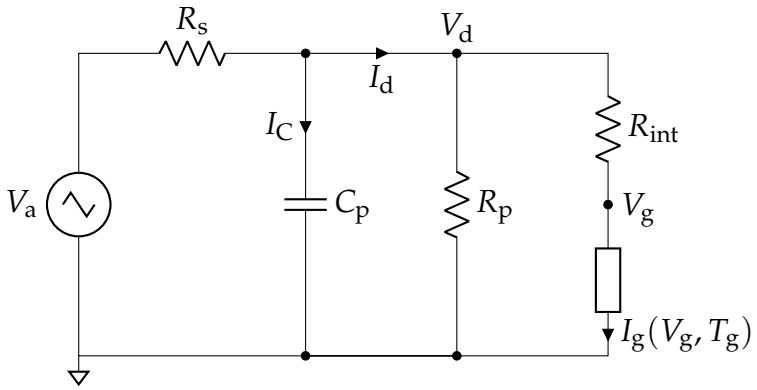

Figure 4.31: Electronic circuit model used for fitting oscillatory measurement data.

The differential-algebraic system of equations corresponding to the electrothermal circuit is

$$\begin{cases} C_{th}\dfrac{dT}{dt} = I_g V_g - \dfrac{T_g - T_0}{R_{th}} \\ C_p\dfrac{dV_d}{dt} = \dfrac{V_a - V_d}{R_s} - I_d \\ I_d = \dfrac{V_d}{R_p} + I_g \\ I_g = \alpha\pi r^2 \dfrac{V_g}{t_g}\exp\left(\dfrac{-b}{T}\right)\exp\left(\gamma\sqrt{\dfrac{V_g}{t_g}}\right) \\ V_g = V_d - I_g R_{int} \end{cases} \tag{4.27}$$

The model was fit to the experimental data in two steps. First, the measured NDR curve below the onset of oscillations was fit to the steady state of the model by varying parameters α, b, γ, t_g, r, R_{th} and R_p. The remaining parameters, C_p, C_{th}, and R_{int}, were then fit by a brute force method where a cartesian product of the three fit parameters within reasonable ranges were tested in $15\times15\times15$ separate numerical simulations. In each case, the system was driven by the same input voltage waveform as used in the experiment, and the oscillation frequencies and amplitudes were extracted. From the 3,375 total simulations, the best fit was chosen to minimize the squared residuals between the measured and simulated oscillation frequency and amplitude vs. voltage curves.

Using this fitting technique, we identified model parameters that give extraordinarily close agreement to the measurement data (see Fig. 4.32). The resulting set of parameters, shown in Table 4.3, are in a physically plausible range, but are effective values within the uniform temperature and thermal resistance approximation. For this reason, parameters like the heat capacity may differ somewhat from bulk values measured in comparable materials.

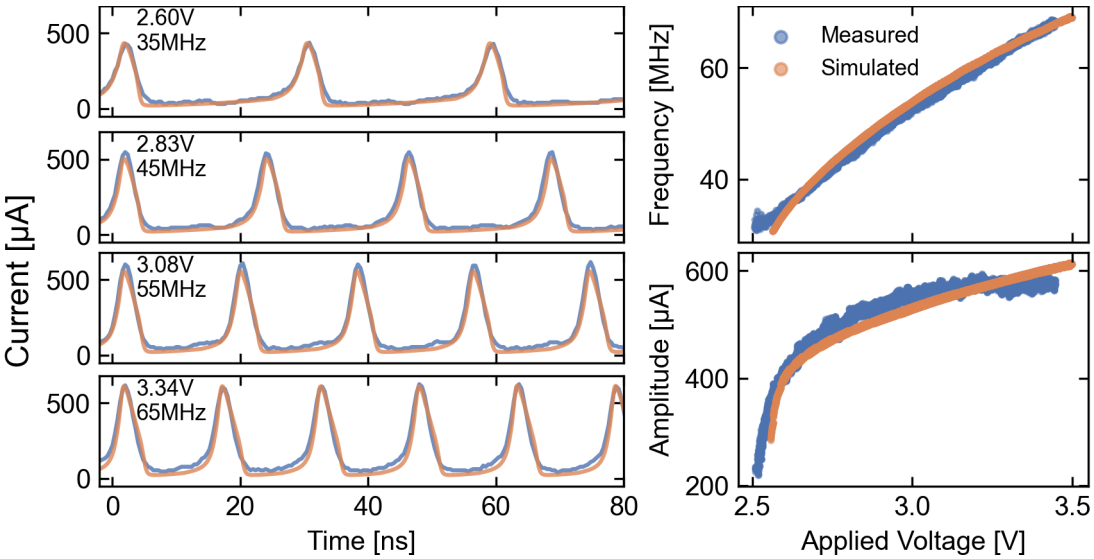

Figure 4.32: Close agreement between experimental and simulated relaxation oscillations in a circuit containing an electroformed $150\times150\times30$ nm^3 $(V_{0.85}Cr_{0.15})_2O_3$ nanodevice. The oscillation frequency and amplitude is controlled by the value of the sourced voltage.

Table 4.3: Parameters for the fitted oscillator model.

Symbol	Value	Unit
α	1.328×10^{4}	S m^{-1}
b	3.212×10^{3}	K
γ	2.750×10^{-4}	m$^{1/2}$ V$^{-1/2}$
t_g	6.000×10^{-9}	m
r	2.257×10^{-8}	m
R_{int}	1.200×10^{3}	Ω
R_p	3.000×10^{5}	Ω
R_{th}	4.100×10^{5}	K W^{-1}
C_{th}	5.278×10^{-16}	J K^{-1}
R_s	1.000×10^{4}	Ω
C_p	3.333×10^{-12}	F

4.5.4 Finite element analysis

Though many of the key concepts of electro-thermal systems can be understood and reasonably modeled using the simplifying assumption of a uniform temperature and current distribution, one aspect not revealed by the simple model is the appearance of temperature and current constriction. In a simple uniform model, the stability landscape can be visualized and understood via phase plane analysis. Spiking and oscillatory modes in circuits can also be predicted in a straightforward way. However, the model cannot explain why filaments sometimes arise in the electroformed states, nor any other possible consequences of strongly constricted temperature profiles.

A partial differential equation (PDE) formulation of an electro-thermal cell, while harder to analyze on a high-level, can take the variations in the temperature and current in one or more spatial dimensions into account. Numerical solutions to PDE boundary-value problems can be computed by the widely-used finite element method (FEM) [248]. One practical issue with this computationally heavy method is that, currently, the most sophisticated FEM solvers run on proprietary programs with poor programming interfaces. Nevertheless, provided enough mental fortitude, certain insights can still be gained from numerical simulation using FEM.

An FEM model was adapted from the existing model of Funck et al [37], where it was used to show that threshold switching in NbO_2-based devices can occur due to thermal runaway at temperatures far below the IMT temperature. The electrical conduction equation was substituted with the fitted conduction equation of Eq. 4.7, and the materials and cell structure were changed to match the $(V_{1-x}Cr_x)_2O_3$ devices measured in this work [249]. It is an axisymmetric 2D model running on the commercial FEM software COMSOL, covering the radial and axial dimensions of a cylindrical geometry with boundary conditions as shown in Fig. 4.33.

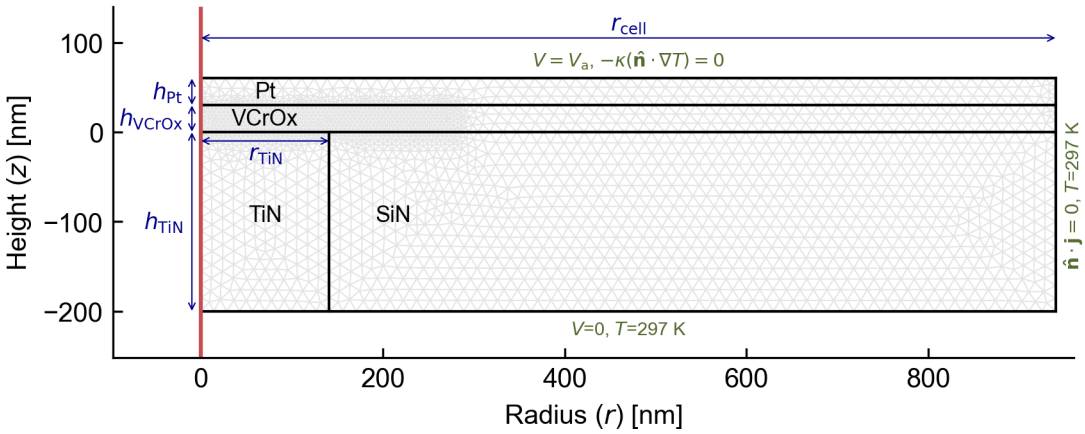

Figure 4.33: The cell geometry, boundary conditions, and meshing used for FEM simulations of a $(V_{1-x}Cr_x)_2O_3$-based electro-thermal cell. The red line at $r = 0$ nm shows the axis of rotational symmetry.

The equation system to solve in the cell volume starts with the point form of Ohm's law,

$$\boldsymbol{j} = \sigma \boldsymbol{E}, \tag{4.28}$$

where $\boldsymbol{j}$ is the current density vector, $\boldsymbol{E}$ is the electric field, and σ is the electric conductivity, which takes a constant value in each material except for the $(V_{1-x}Cr_x)_2O_3$ layer, where it has temperature and field dependence defined by

$$\sigma_{\mathrm{VCrOx}} = \alpha \exp\left(\frac{-b}{T}\right) \exp\left(\gamma\sqrt{|\boldsymbol{E}|}\right). \tag{4.29}$$

The rest of PDE system consists of static form of the equation of current continuity,

$$\nabla \cdot \boldsymbol{j} = \nabla \cdot (\sigma \nabla V) = 0, \tag{4.30}$$

and the heat equation,

$$\rho_{\mathrm{m}} C_{\mathrm{th,p}} \frac{\partial T}{\partial t} = Q + \nabla(\kappa \nabla T), \tag{4.31}$$

where ρ_{m} is the density of mass, $C_{\mathrm{th,p}}$ is the specific heat capacity, κ is the thermal conductivity, and the heat energy per unit volume is generated by Joule heating,

$$Q = \frac{|\boldsymbol{j}|^2}{\sigma}. \tag{4.32}$$

The material parameter values were largely sourced from thin film literature and are provided in Table 4.4.

When driving the cells with a static current source (implemented as a boundary condition), S-type NDR shape of the (I, V) steady states appear in analogy with the uniform temperature case. However, along with the maximum temperature T_{max}, each steady state is associated with a temperature profile that changes shape as the NDR curve current is increased (see Fig. 4.34). The full width at half maximum (FWHM), defined as the diameter of the temperature contour at $T = (T_0 + T_{\mathrm{max}})/2$ in the middle oxide thickness, gives a scalar indication of the constriction of the temperature distribution. Below the NDR onset current, the temperature distribution of the thermal steady state is nearly uniformly distributed within the via radius, and is elevated less than 40 °C from baseline. As the current increases inside the NDR range, the temperature distribution constricts slightly, with the FWHM gradually reducing below the device diameter but never reaching less than 75% of it. The temperature profile also constricts asymmetrically in the z dimension, with the maximum temperature being reached slightly below the middle of the oxide thickness.

Although the physical processes behind NV switching are not explicitly modeled in this work, we hypothesize a connection between the temperature profiles induced by Joule heating and the different types of electro-forming effects that were observed experimentally.

For the model parameters assigned here, thermal steady states along the

Table 4.4: Parameters for the FEM model.

Symbol	Value	Unit	Description
r_{cell}	940	nm	Radius of the cylindrical cell
r_{TiN}	141	nm	Radius of the TiN via
h_{TiN}	200	nm	Thickness of the TiN via
h_{VCrOx}	30	nm	Thickness of the VCrOx layer
h_{Pt}	30	nm	Thickness of the Pt layer
κ_{Pt}	30	$\text{W m}^{-1}\text{ K}^{-1}$	Thermal conductivity of Pt [250]
σ_{Pt}	9.50×10^{6}	S m^{-1}	Electrical conductivity of Pt [250]
$\rho_{\text{m,Pt}}$	21.5	g cm^{-3}	Mass density of Pt [250]
$C_{\text{th,p,Pt}}$	133	$\text{J kg}^{-1}\text{ K}^{-1}$	Heat capacity of Pt [250]
κ_{TiN}	2.7	$\text{W m}^{-1}\text{ K}^{-1}$	Thermal conductivity of TiN [251]
σ_{TiN}	6.25×10^{6}	S m^{-1}	Electrical conductivity of TiN [252]
$\rho_{\text{m,TiN}}$	5.22	g cm^{-3}	Mass density of TiN [253]
$C_{\text{th,p,TiN}}$	601	$\text{J kg}^{-1}\text{ K}^{-1}$	Heat capacity of TiN [254]
κ_{SiN}	2.0	$\text{W m}^{-1}\text{ K}^{-1}$	Thermal conductivity of SiN [255]
σ_{SiN}	9.35×10^{-14}	S m^{-1}	Electrical conductivity of SiN [256]
$\rho_{\text{m,SiN}}$	3.19	g cm^{-3}	Mass density of SiN [256]
$C_{\text{th,p,SiN}}$	900	$\text{J kg}^{-1}\text{ K}^{-1}$	Heat capacity of SiN [256]
κ_{VCrOx}	1.0	$\text{W m}^{-1}\text{ K}^{-1}$	Thermal conductivity of VCrOx
$\rho_{\text{m,VCrOx}}$	4.50	g cm^{-3}	Mass density of VCrOx [220]
$C_{\text{th,p,VCrOx}}$	3333	$\text{J kg}^{-1}\text{ K}^{-1}$	Heat capacity of VCrOx [257]
α	1.429×10^{4}	S m^{-1}	VCrOx conduction prefactor
b	3.017×10^{3}	K	VCrOx conduction energy barrier
γ	2.600×10^{-4}	$\text{m}^{1/2}\text{ V}^{-1/2}$	VCrOx conduction nonlinearity coefficient

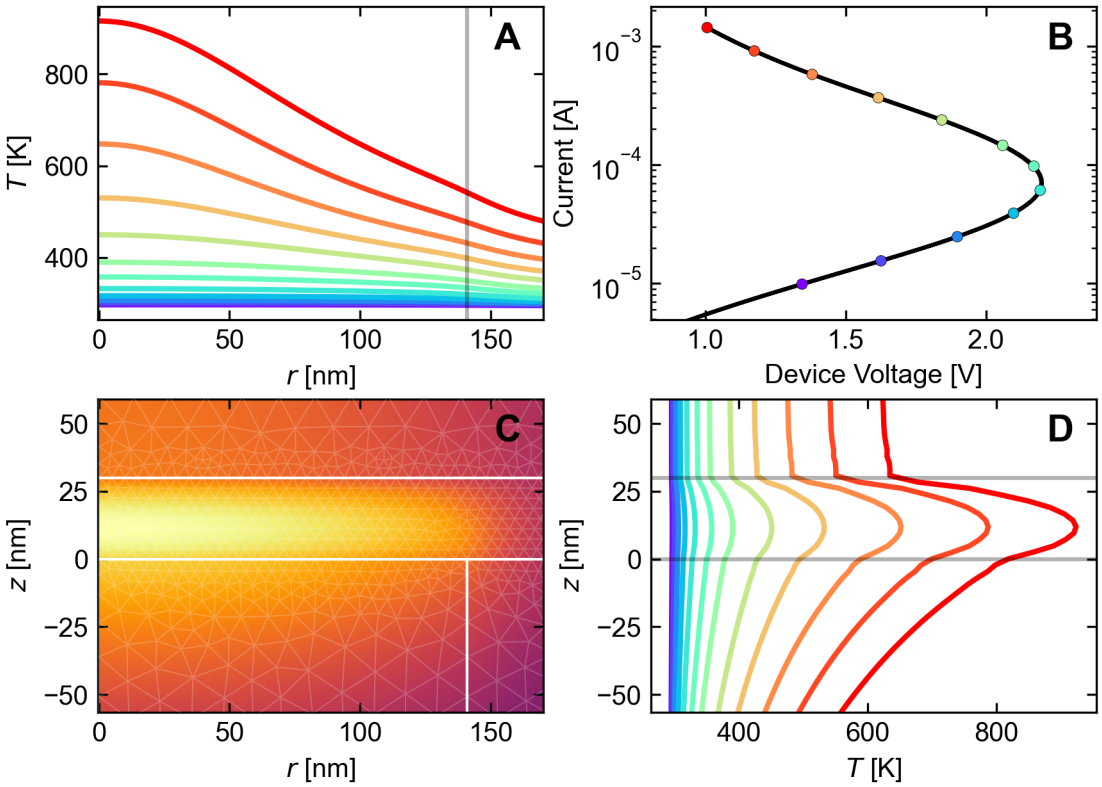

Figure 4.34: Simulated temperature (T) profiles of the thermal steady state for different source current levels along the NDR curve. **(C)** shows the cell geometry and the temperature distribution at 200 µA. **(A)** shows the radial profile at half oxide thickness and **(D)** the thickness profile at $r = 0$. **(B)** displays the location of the NDR bias for each of the temperature profiles.

NDR curve have a relatively broad temperature distribution. Assuming stable NDR biasing over a sufficiently long period of time, it can be expected that a temperature-dependent NV process (such as the production and drift of oxygen vacancies) will also follow the radial profile of the temperature distribution previously set up by Joule heating. This helps explain the relatively uniform NV layer growth observed for these conditions in our experiments. The main parameter influencing the FWHM of the temperature profile is the thermal conductivity of the oxide, which may vary with temperature according to the Wiedemann–Franz law in general but is assumed constant here. As shown in Fig. 4.35, with all else held constant, lower thermal conductivities increase the steady-state constriction significantly while also slightly modifying the NDR shape itself. For this model, the thermal conductivity value for the $(V_{1-x}Cr_x)_2O_3$ layer was determined by time-domain thermoreflectance to be 1.0 $W\,m^{-1}\,K^{-1}$. By producing a more pronounced constriction, we expect that lower κ values would make it more likely to locally induce chemical or structural changes at an earlier point along the NDR curve, potentially even below the knee current. Therefore, it is possible that for lower κ materials, a reversible NDR region would not be observable and the formed states could be intrinsically more constricted than we observed in our experiments.

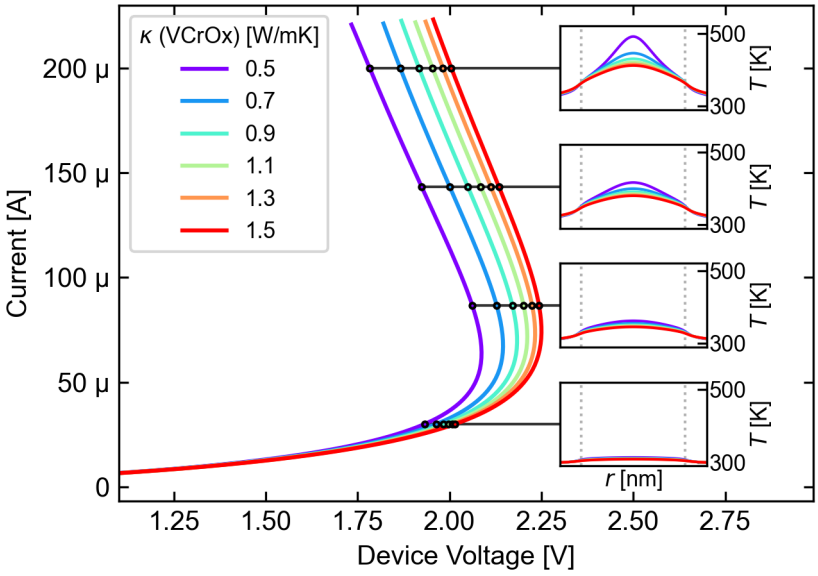

Figure 4.35: The effect of the assumed thermal conductivity of the oxide layer, κ, on the simulated temperature profile in the NDR steady states. With all else equal, a lower thermal conductivity gives more constricted solutions and higher peak temperatures at the same current values.

It is also illustrative to consider the temperature profiles induced outside of thermal equilibrium, as will occur when the cell is driven by fast voltage signals or when there is a driving signal-induced circuit bifurcation. To investigate this, we compare three simulated driving scenarios for the cell. The external series resistance and the applied voltage waveform are varied in order to

1. trace the device NDR curve,
2. produce a saddle-node bifurcation and a load-line jump to a new, distant equilibrium, and
3. induce an out-of-equilibrium switching trajectory by overdriving with a square pulse.

The results, summarized in Fig. 4.36, show that while the cells eventually settle into the same thermal profiles of the NDR steady state, these situations each result in very different transient levels of thermal filamentation. At every point of the switching trajectories, the FWHM was smaller than or equal to the corresponding NDR steady state at the same current level. The maximum degree of filamentation roughly followed the speed of the transition in question. The load-line jump in scenario 2 lasted ∼150 ns and had a minimum FWHM of 140 nm, while in scenario 3 the jump lasted ∼50 ns and had a lower minimum FWHM of 60 nm. The maximum temperatures reached at the center of the cells that experienced constriction were also much higher than seen in the steady state, even at very similar instantaneous (I, V) values. It is easy to

imagine that if these operations were carried out using any real-world material, the extreme constriction seen at some point during the trajectories could lead to decomposition in narrow filamentary region of the material, in contrast with the more uniform case seen with steady NDR biasing.

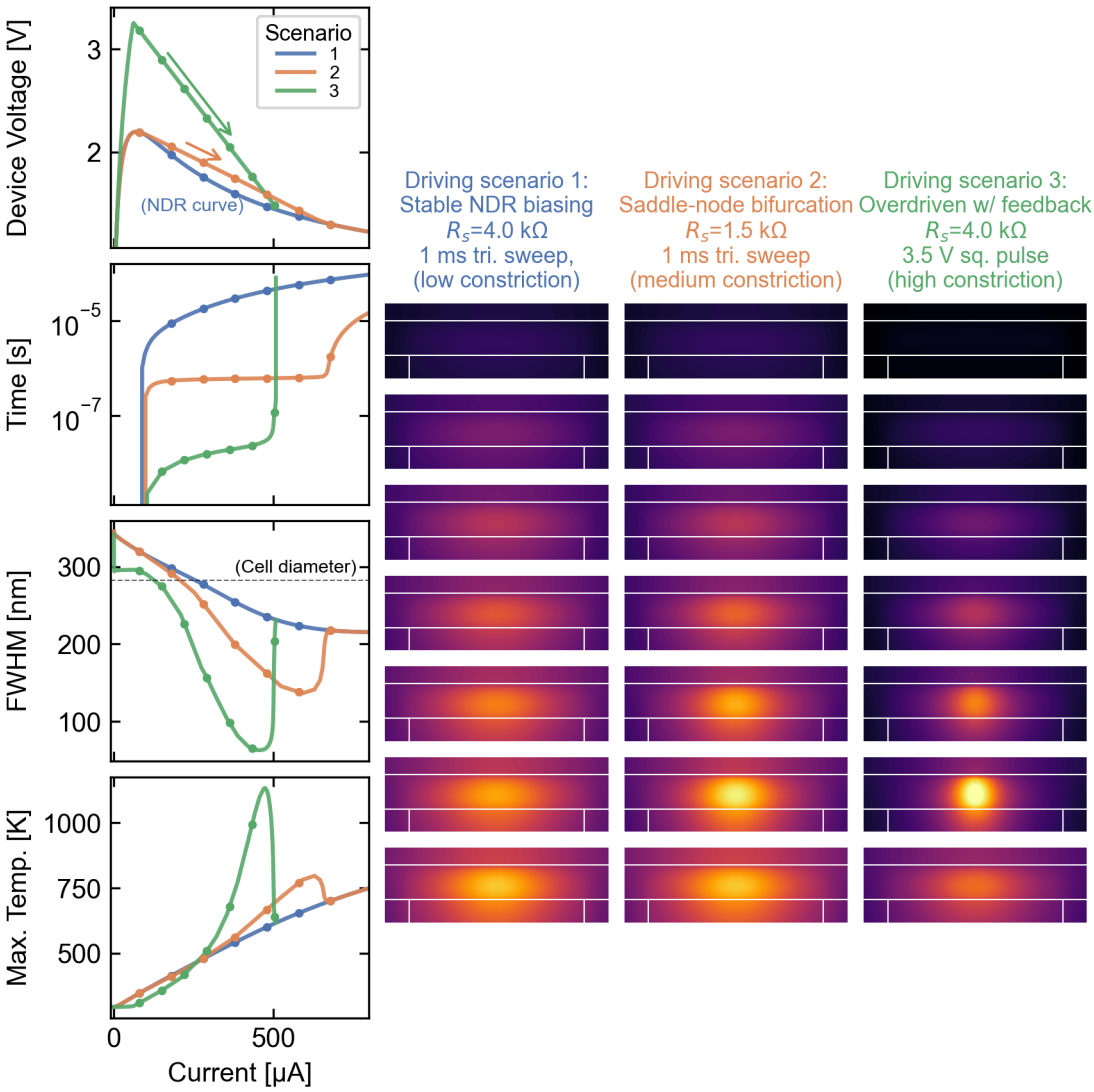

Figure 4.36: A comparison of simulated electro-thermal switching for three different driving scenarios that show strong differences in the amount of thermal constriction. On the left side of the figure, parameters at each current level are plotted. On the right side of the figure, corresponding 2D temperature profiles are shown in vertically stacked frames depicting increasing current levels.

Among the experimental results shown previously, the sudden forming mode was associated with large current spikes from the parasitic capacitance. However, we did not include parasitic capacitance in the FEM model and still encountered strong thermal constriction without the spikes. A similar finding was also recently pointed out by Goodwill et al. [258]. Fundamentally, current

constriction comes from sustained positive feedback between local conductivity enhancement and its driving force (here, heat generation), and there are therefore certain scenarios that predictably lead to filamentation. Whether and to what degree the positive feedback occurs is determined by the overall measurement circuit, of which the DUT is only one component. In the experimental case, the parasitic capacitance further degrades the ability to stabilize the cells, but we would also expect to see the filamentation effect even if it were possible to reduce the parasitic capacitance to zero. The distinguishing point is not the spiking itself, but that whenever the driving circuit pushes the (I, V) trajectory of the cell far away from its NDR curve, the heating rates become large relative to the equalizing factor of thermal diffusion. This can happen for variety of different reasons and has multiple contributing factors in an experiment. However induced, we expect that rapid changes in the device resistance state will tend to be associated with filamentary conduction paths.

On the materials side, an important quantity when considering the balance between filamentary and non-filamentary switching is the thermal diffusivity of the medium,

$$\alpha_{\text{th}} = \frac{\kappa}{\rho_{\text{m}} C_{\text{th,p}}}, \tag{4.33}$$

which is roughly the rate that the temperature, and thereafter the NV filament, can spread through the material. It evaluates to $\sim$60 nm^2/ns for the parameters used for the $(V_{1-x}Cr_x)_2O_3$ film. This establishes an informal connection between the timescale of the switching event and the area of the filament produced in the process. At the same time, we expect the thermal conductivity to also set a kind of upper bound on the filamentary radius, due to its effect on the steady-state temperature profiles, the FWHMs of which are not exceeded during driven transients.

4.6 Discussion

This chapter described the fabrication of nano-devices based on the Mott insulating material $(V_{1-x}Cr_x)_2O_3$. After verifying a high compositional and structural quality, we applied non-standard electrical measurement techniques to study both volatile and non-volatile switching phenomena in the devices. The cells have an adaptable NDR characteristic, leading to highly endurant excitable dynamics of interest for emulating spiking neurons in neuromorphic circuits.

The volatile switching component is comprehensively explained by Joule heating and conductivity enhancement at elevated temperatures due to carrier activation. The astute reader will notice a close resemblance of the switching behavior to that of the macroscopic thermistor measurement in Appendix B. While some of the phenomena strain human intuition, they have the same character as those observed in a simple and decidedly ordinary system. The remaining differences in speed, voltage, and current levels can be explained simply by scaling to nanometer dimensions.

Nano-electronic applications ask for smaller device dimensions than are easy to fabricate in a research setting. By measuring and analyzing the NDR curves of devices in a variety of different sizes (thicknesses and widths), we were able to determine the scaling behavior of electrical and thermal conduction in the cells. This information allowed us to build a device model that predicts good scaling behavior down to a 10 nm length scale.

Furthermore, we observed that NV resistance transitions can occur in different ways and with different consequences. A gradual process is induced by extended biasing at (I, V) points near the NDR curve, while a rapid process can be induced through various circuit instabilities. By a combination of evidence from experiment and simulation, we postulate that heating activates a NV process and results in the formation of a highly conductive pillar inside the cell. Depending on how the NV transition is induced, this pillar can be either highly localized or distributed across the entire cell area. The pillar could consist of an oxygen deficient or (re)crystallized oxide phase, or possibly a compressed metallic domain(s) having undergone a metal-insulator transition [259]. We judge ionic motion to be most likely given the prevalence of this mechanism in TMOs, the similarity of the bipolar (I, V) measurements to those produced in VCM cells, and the absence of a detectable change in cross-sectional TEM. However, we did not make a final determination of which ionic species are involved.

By using measurement circuits with limiting feedback during NV switching, we are able to access a range of states with NDR and TS modified from the pristine characteristic. Our model suggests that after these NV (forming) events, the modified volatile effects proceed in the same material and due to the same mechanism as in the forming-free operation, but occur in the decreased volume of the pillar gap. By supposing a functional equivalence between a formed device and a device with a smaller size (corresponding to the gap size), we showed evidence that scaling the device down will further reduce the leakage, significantly speed up the switching, and eliminate the necessity of a forming step; all of which are beneficial for applications.

Besides careful attention to the measurement circuitry, three experimental conveniences contributed to the ability to record these effects in this particular material system. First, the relatively high electrical conductivity of the pristine material reduces the series resistance and applied voltage needed to stabilize its NDR curve. Second, the high thermal conductivity makes the thermal steady-state temperature profile relatively uniform across the device width, making the material more likely to reversibly withstand NDR biasing on relatively long timescales. Finally, the apparent absence of interfacial Schottky barriers gives visibility of the conductance change in the bulk of the device as the pillar grows, which might not be the case for material stacks with interface-limited conduction.

Conclusion

Resistive switching devices are promising building blocks for future memory and unconventional computing architectures, but remaining non-idealities presently stand in the way of their production readiness. Two of the most salient challenges are that they are prone to electrical instabilities and have large cycle-to-cycle and device-to-device variability. Conventional lab measurements of these cells commonly represent very different conditions from integrated systems, and often have unclear implications for device applications. In particular, without integrated current-limiting devices, overshoots during runaway resistance transitions hinder the ability to control and characterize switching processes. With isolated devices vulnerable to runaway transitions, parasitic capacitance is an important factor, and both the type of external feedback and the amount of parasitic capacitance lead to different switching outcomes in general.

Due to the strong influence of the measurement setup on switching data, device evaluation requires accurate modeling of the driving circuitry, which rules out the use of proprietary "black box" feedback circuits. The two measurement circuit designs introduced in Chapter 2 provide a new, simple means for control and tractable analysis of switching processes, without the prohibitive requirement of CMOS integration. They provide synchronized measurement of switching trajectories at high speed while using adjustable feedback mechanisms to stabilize the system and to control the amount of resistive switching induced. The minimal designs with low part counts are relatively robust against load-induced instability and have the important advantage that their response is accurately predictable using only a few idealized component models.

We demonstrated collection of 10^5 to 10^7 switching cycles per second with the new system, which is highly useful for studying the stochastic nature of switching processes. The DUT current and voltage are simultaneously sampled for up to ten million switching cycles per second, and millions of such cycles are endured even by devices which do not survive switching using commercial SMUs. By combining the strengths of conventional approaches — the short time scales and fast sampling rates of pulsed measurements and the higher information content of quasistatic cycles — this new technique is capable generating rich, statistically significant datasets in a short amount of time. With this technique, the nature of physical processes that support neuromorphic functionalities can be conveniently investigated, yielding insights into how optimal control can eventually be achieved.

Using the statistical information collected by applying the new measurement circuitry to a fabricated HfO_2-based ReRAM device design, we constructed a new type of stochastic device model. The hierarchical statistical model is optimized for low memory footprint and fast, parallel simulation of large synaptic arrays on both CPUs and GPUs. This work is intended to seamlessly provide a physically validated drop-in model of a resistive RAM cell for large-scale neuromorphic simulations, which automatically adapts to new measurement data. We have open-sourced the model implementation and hope that it will be of use to the community as a plug-in model for RS devices, or as a framework to create further device models.

Finally, nanodevices based on the correlated-electron material $(V_{1-x}Cr_x)_2O_3$ were fabricated and characterized in this work by X-ray spectroscopy and by a variety of electrical measurements. High speed current excursions were demonstrated in simple excitable circuits containing the devices, making them interesting as compact spiking elements for use in neuromorphic computing. While the material characterization suggests high phase purity and correct stoichiometry, unambiguous evidence of an electrically induced Mott-type IMT in these devices is still absent, and may require further material engineering to realize in thin films. We interpret the various volatile switching results without reference to phase transitions, and instead find very close agreement with an electro-thermal model. Therefore, considerations not only of the material properties but also the thermal environment are critical for determining device behavior, and adjusting the thermal isolation provides an engineering path for modifying important variables such as the threshold voltage and leakage current. A major challenge in applying the devices will be reducing the quiescent power consumption. Our model suggests favorable scaling properties in this regard, but it remains to be seen whether fully scaled devices will follow the prediction and whether they are capable of self-oscillation without the addition of bulky capacitive components.

In addition to the electro-thermal volatile switching effects, a NV effect comparable to the VCM-type acts in parallel in the $(V_{1-x}Cr_x)_2O_3$ devices. Although a physical/dynamical model for this effect was not considered, the approximate shape of the NV conductance redistribution was inferred through its effect on the remaining volatile switching behavior. The results suggest that filaments are not a fundamental and unavoidable material property, but rather a controllable consequence of the overall stability of the measurement circuitry. We showed electrical measurements that are capable of differentiating these conditions, and we expect that a range of filament diameters could be realized in the same system in a predictable way by modifying the driving circumstances, and that there could be accompanying differences in switching performance. Furthermore, to explain the persistent volatile switching after NV perturbation, our model supposes that the volatile effect occurs in a remaining gap of pristine material due to the same electro-thermal effect as the

as-deposited device. Approached in these terms, the results contribute valuable insight about how both the volatile and non-volatile properties of intermediate resistance states can be harnessed for use in neuromorphic circuits.

Appendix A

Parameter assignment for the thermistor model

In a basic thermistor model, we assume a temperature dependent electrical conduction relation where

$$I = AV \exp\left(\frac{-E_\mathrm{b}}{k_\mathrm{B}T}\right), \tag{A.1}$$

and a lumped thermal model,

$$C_\mathrm{th}\frac{dT}{dt} = IV - \frac{T - T_0}{R_\mathrm{th}}. \tag{A.2}$$

Plausible parameter values are chosen for a clear visual representation on a linear axis scale. The material is assumed an intrinsic semiconductor with $E_\mathrm{b} = 0.5$ eV, and $A = 5 \times 10^3$ S such that the cell has 50 kΩ resistance at room temperature $T_0 = 300$ K.

For a rough idea of the expected order of magnitude of R_th, consider the steady-state 3D heat equation in an isotropic medium with uniform thermal conductivity κ,

$$-\kappa\nabla^2 T(r) = q(r) \tag{A.3}$$

With a point heat source q at the origin and with the boundary condition $T \to T_0$ as the radius $r \to \infty$, we have the solution

$$T(r) = \frac{q}{4\pi\kappa r} + T_0. \tag{A.4}$$

The average temperature inside of radius $r \leq R$ is

$$\overline{T} = \frac{3}{4\pi R^3}\int_0^R T(r)4\pi r^2 dr = \frac{3q}{8\pi\kappa R} + T_0, \tag{A.5}$$

and the effective thermal resistance is

$$R_\mathrm{th} = \frac{d\overline{T}}{dq} = \frac{3}{8\pi\kappa R}, \tag{A.6}$$

which is inversely proportional to the radius and the thermal conductivity. Assuming hypothetical values of $\kappa = 10\ \mathrm{W/m/K}$ and $R = 50\ \mathrm{nm}$, we assign $R_{\mathrm{th}} = 2.4 \times 10^5\ \mathrm{K/W}$. For the same cell volume, assuming a heat capacity of $2\ \mathrm{J/K/cm^3}$, we use a lumped value of $C_{\mathrm{th}} = 1.05 \times 10^{-15}\ \mathrm{J/K}$.

Appendix B

Zoo of thermistor behavior

An interesting set of dynamical effects can appear in electrical measurements, which simply follow from Joule heating in a device with strongly temperature dependent conductivity. To show this unambiguously, we measured a commercial negative temperature coefficient (NTC) thermistor intended for temperature sensing applications (TDK part number NTCG063JF103FT). The small surface mount device (300 μm × 300 μm × 500 μm) was presumably engineered for high voltage linearity, using a resistive material composed of unspecified semiconductor ceramics, and layered internal Pd electrodes. The conduction model given in the device datasheet specifies the static resistance as a function of temperature,

$$R = R_0 \exp\left(B\left(\frac{1}{T} - \frac{1}{T_0}\right)\right) \tag{B.1}$$

which equals the reference resistance $R_0 = 10$ kΩ at temperature $T_0 = 25$ °C, and $B = 3400$ K corresponds to a bandgap of 0.3 eV.

The device can also be operated outside the intended temperature sensing regime, where the self-heating effect produces NDR curves, which are also provided in the datasheet. A variety of measurements of the NTC thermistor using a Keithley 2636B SMU are shown in Fig. B.1. In every case, there was no permanent change in the NTC after the voltage was removed. These effects mirror those observed in scaled electro-thermal cells, but they occur on a much larger time and power scale, which is explained by the relatively large heat capacity, low thermal resistance, and absence of barrier lowering due to low electric field.

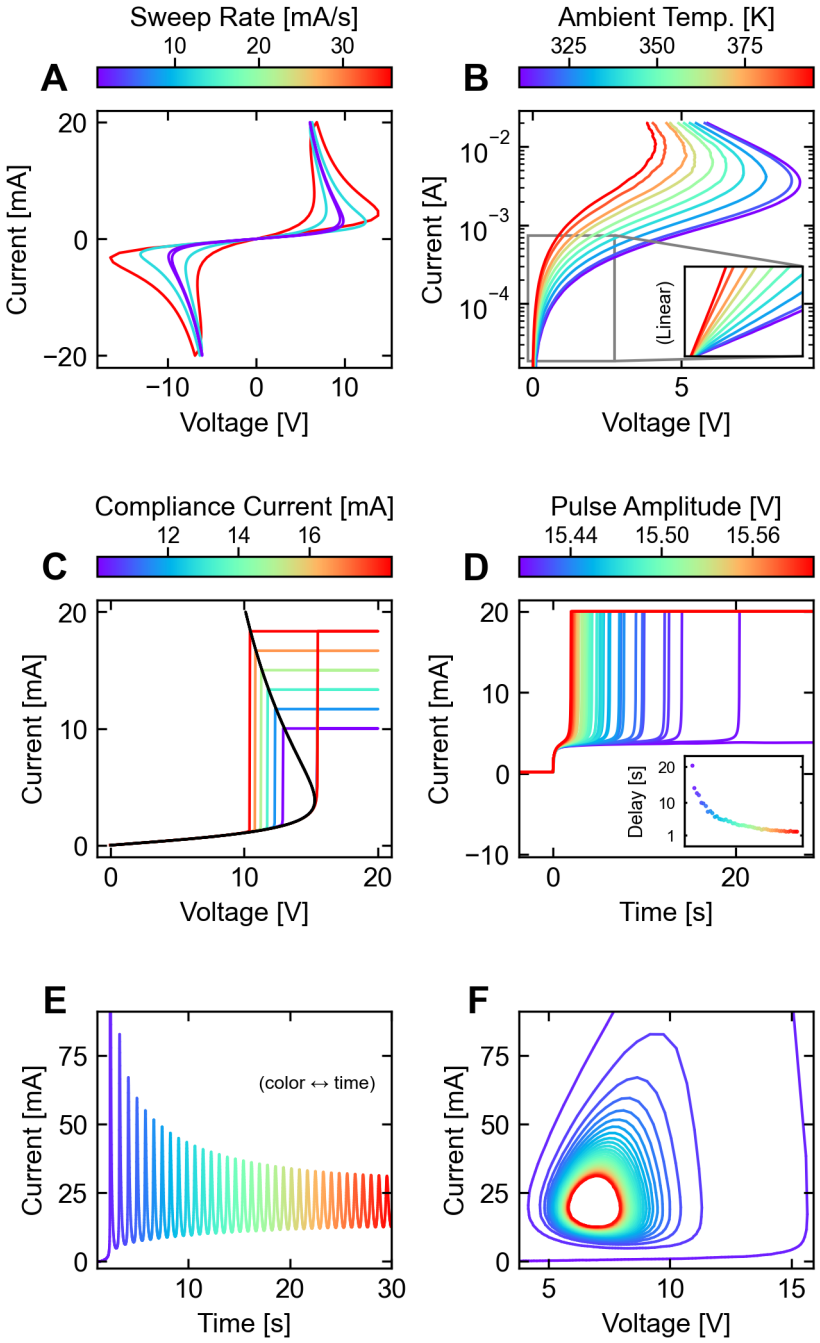

Figure B.1: Electrical measurements of a commercial surface-mount NTC thermistor. **(A)** NDR due to self-heating is measured with a sweeping current source. NDR curve retraces itself at low sweep rates, and a hysteresis opens at higher sweep rates. **(B)** The NDR curve depends strongly on the ambient temperature. **(C)** Sudden current jumps are measured when controlled with a sweeping voltage source. **(D)** Current jumps (duration ∼100 ms) occur after a delay which depends sensitively on the amplitude of an applied voltage pulse. **(E) (F)** Relaxation oscillations occur when driven by a 20 mA current source with a 470 µF capacitor in parallel.

Bibliography

[1] J. Backus, "Can programming be liberated from the von Neumann style?: A functional style and its algebra of programs," *Commun. ACM*, vol. 21, no. 8, pp. 613–641, Aug. 1978, ISSN: 0001-0782, 1557-7317. DOI: 10.1145/359576.359579.

[2] W. A. Wulf and S. A. McKee, "Hitting the memory wall: Implications of the obvious," *SIGARCH Comput. Archit. News*, vol. 23, no. 1, pp. 20–24, Mar. 1995, ISSN: 0163-5964. DOI: 10.1145/216585.216588.

[3] G. Indiveri and S.-C. Liu, "Memory and Information Processing in Neuromorphic Systems," *Proc. IEEE*, vol. 103, no. 8, pp. 1379–1397, Aug. 2015, ISSN: 0018-9219, 1558-2256. DOI: 10.1109/JPROC.2015.2444094.

[4] I. L. Markov, "Limits on fundamental limits to computation," *Nature*, vol. 512, no. 7513, pp. 147–154, Aug. 2014, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature13570.

[5] C. Mead, "Neuromorphic Electronic Systems," *Proc. IEEE*, vol. 78, p. 8, 1990.

[6] D. V. Christensen, R. Dittmann, *et al.*, "2022 roadmap on neuromorphic computing and engineering," *Neuromorph. Comput. Eng.*, vol. 2, no. 2, p. 022 501, Jun. 2022, ISSN: 2634-4386. DOI: 10.1088/2634-4386/ac4a83.

[7] W. Wan, R. Kubendran, *et al.*, "A compute-in-memory chip based on resistive random-access memory," *Nature*, vol. 608, no. 7923, pp. 504–512, Aug. 2022, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/s41586-022-04992-8.

[8] P. A. Merolla, J. V. Arthur, *et al.*, "A million spiking-neuron integrated circuit with a scalable communication network and interface," *Science*, vol. 345, no. 6197, pp. 668–673, Aug. 2014, ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1254642.

[9] M. Davies, N. Srinivasa, *et al.*, "Loihi: A Neuromorphic Manycore Processor with On-Chip Learning," *IEEE Micro*, vol. 38, no. 1, pp. 82–99, Jan. 2018, ISSN: 0272-1732, 1937-4143. DOI: 10.1109/MM.2018.112130359.

[10] D. J. Wouters, R. Waser, and M. Wuttig, "Phase-Change and Redox-Based Resistive Switching Memories," *Proc. IEEE*, vol. 103, no. 8, pp. 1274–1288, Aug. 2015, ISSN: 0018-9219. DOI: 10.1109/JPROC.2015.2433311.

[11] S. Yu and P.-Y. Chen, "Emerging Memory Technologies: Recent Trends and Prospects," *IEEE Solid-State Circuits Mag.*, vol. 8, no. 2, pp. 43–56, 2016, ISSN: 1943-0582. DOI: 10.1109/MSSC.2016.2546199.

[12] H. Li, T. F. Wu, S. Mitra, and H.-S. P. Wong, "Resistive RAM-Centric Computing: Design and Modeling Methodology," *IEEE Trans. Circuits*

Syst. I, vol. 64, no. 9, pp. 2263–2273, Sep. 2017, ISSN: 1549-8328, 1558-0806. DOI: 10.1109/TCSI.2017.2709812.

[13] G. W. Burr, R. M. Shelby, *et al.*, "Neuromorphic computing using non-volatile memory," *Adv. Phys. X*, vol. 2, no. 1, pp. 89–124, Jan. 2017, ISSN: 2374-6149. DOI: 10.1080/23746149.2016.1259585.

[14] V. K. Sangwan and M. C. Hersam, "Neuromorphic nanoelectronic materials," *Nat. Nanotechnol.*, vol. 15, no. 7, pp. 517–528, Jul. 2020, ISSN: 1748-3387, 1748-3395. DOI: 10.1038/s41565-020-0647-z.

[15] V. Milo, G. Malavena, C. Monzio Compagnoni, and D. Ielmini, "Memristive and CMOS Devices for Neuromorphic Computing," *Materials*, vol. 13, no. 1, p. 166, Jan. 2020, ISSN: 1996-1944. DOI: 10.3390/ma13010166.

[16] D. Ielmini and R. Waser, *Resistive Switching: From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications*. John Wiley & Sons, 2015.

[17] A. Chen, J. Hutchby, V. V. Zhirnov, and G. Bourianoff, Eds., *Emerging Nanoelectronic Devices*. Chichester, West Sussex, United Kingdom: John Wiley & Sons Inc, 2014, ISBN: 978-1-118-44774-1.

[18] You Zhou and S. Ramanathan, "Mott Memory and Neuromorphic Devices," *Proc. IEEE*, vol. 103, no. 8, pp. 1289–1310, Aug. 2015, ISSN: 0018-9219, 1558-2256. DOI: 10.1109/JPROC.2015.2431914.

[19] H. Liu, D. Bedau, J. Sun, S. Mangin, E. Fullerton, J. Katine, and A. Kent, "Dynamics of spin torque switching in all-perpendicular spin valve nanopillars," *Journal of Magnetism and Magnetic Materials*, vol. 358–359, pp. 233–258, May 2014, ISSN: 03048853. DOI: 10.1016/j.jmmm.2014.01.061.

[20] R. Waser, R. Dittmann, G. Staikov, and K. Szot, "Redox-Based Resistive Switching Memories - Nanoionic Mechanisms, Prospects, and Challenges," *Adv. Mater.*, vol. 21, no. 25-26, pp. 2632–2663, Jul. 2009, ISSN: 09359648, 15214095. DOI: 10.1002/adma.200900375.

[21] L. O. Chua, "Memristor-The missing circuit element," *IEEE Trans. Circuit Theory*, vol. 18, no. 5, pp. 507–519, 1971, ISSN: 0018-9324. DOI: 10.1109/TCT.1971.1083337.

[22] J. Kim, Y. V. Pershin, M. Yin, T. Datta, and M. Di Ventra, "An Experimental Proof that Resistance-Switching Memory Cells are not Memristors," *Adv. Electron. Mater.*, vol. 6, no. 7, p. 2 000 010, Jul. 2020, ISSN: 2199-160X, 2199-160X. DOI: 10.1002/aelm.202000010.

[23] M. Di Ventra and Y. V. Pershin, *Memristors and Memelements: Mathematics, Physics and Fiction* (SpringerBriefs in Physics). Cham: Springer International Publishing, 2023, ISBN: 978-3-031-25624-0 978-3-031-25625-7. DOI: 10.1007/978-3-031-25625-7.

[24] L. O. Chua and S. M. Kang, *Memristive Devices and Systems*, 1976.

[25] D. B. Strukov, G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, vol. 453, no. 7191, pp. 80–83, May 2008, ISSN: 0028-0836. DOI: 10.1038/nature06932.

[26] S. Vongehr and X. Meng, "The Missing Memristor has Not been Found," *Sci. Rep.*, vol. 5, no. 1, Dec. 2015, ISSN: 2045-2322. DOI: 10.1038/srep11657.

[27] S. Chandra, "On the Discovery of a Polarity-Dependent Memory Switch and/or Memristor (Memory Resistor)," *IETE Tech. Rev.*, vol. 27, no. 2, p. 179, 2010, ISSN: 0256-4602. DOI: 10.4103/0256-4602.60170.

[28] R. Williams, "Reply to - "On the Discovery of a Polarity-Dependent Memory Switch and/or Memristor (Memory Resistor)"," *IETE Tech. Rev.*, vol. 27, no. 2, p. 181, 2010, ISSN: 0256-4602. DOI: 10.4103/0256-4602.60171.

[29] B. Mouttet, "The Mythology of the Memristor," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2010.

[30] P. Meuffels and R. Soni, "Fundamental Issues and Problems in the Realization of Memristors," *ArXiv Prepr.*, 2012. DOI: 10.48550/arXiv.1207.7319.

[31] L. Chua, "Resistance switching memories are memristors," *Appl. Phys. A*, vol. 102, no. 4, pp. 765–783, Mar. 2011, ISSN: 0947-8396, 1432-0630. DOI: 10.1007/s00339-011-6264-9.

[32] S. Kim, M. Lim, Y. Kim, H.-D. Kim, and S.-J. Choi, "Impact of Synaptic Device Variations on Pattern Recognition Accuracy in a Hardware Neural Network," *Sci Rep*, vol. 8, no. 1, p. 2638, Dec. 2018, ISSN: 2045-2322. DOI: 10.1038/s41598-018-21057-x.

[33] C. Bengel, F. Cüppers, M. Payvand, R. Dittmann, R. Waser, S. Hoffmann-Eifert, and S. Menzel, "Utilizing the Switching Stochasticity of HfO_2/TiOx-Based ReRAM Devices and the Concept of Multiple Device Synapses for the Classification of Overlapping and Noisy Patterns," *Front. Neurosci.*, vol. 15, p. 661 856, Jun. 2021, ISSN: 1662-453X. DOI: 10.3389/fnins.2021.661856.

[34] E. Janod, J. Tranchant, *et al.*, "Resistive Switching in Mott Insulators and Correlated Systems," *Adv. Funct. Mater.*, vol. 25, no. 40, pp. 6287–6305, Oct. 2015, ISSN: 1616301X. DOI: 10.1002/adfm.201500823.

[35] M. Querré, J. Tranchant, *et al.*, "Non-volatile resistive switching in the Mott insulator $(V_{1-x}Cr_x)_2O_3$," *Phys. B Condens. Matter*, vol. 536, pp. 327–330, May 2018, ISSN: 09214526. DOI: 10.1016/j.physb.2017.10.060.

[36] J. S. Brockman, "Electric Field-Induced Conductivity Switching in Vanadium Sesquioxide Nanostructures," Ph.D. dissertation, 2012.

[37] C. Funck, S. Menzel, N. Aslam, H. Zhang, A. Hardtdegen, R. Waser, and S. Hoffmann-Eifert, "Multidimensional Simulation of Threshold Switching in NbO_2 Based on an Electric Field Triggered Thermal Runaway Model," *Adv. Electron. Mater.*, vol. 2, no. 7, p. 1 600 169, Jul. 2016, ISSN: 2199160X. DOI: 10.1002/aelm.201600169.

[38] M. Le Gallo, A. Athmanathan, D. Krebs, and A. Sebastian, "Evidence for thermally assisted threshold switching behavior in nanoscale phase-change memory cells," *J. Appl. Phys.*, vol. 119, no. 2, p. 025 704, Jan. 2016, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.4938532.

[39] J. M. Goodwill, D. K. Gala, J. A. Bain, and M. Skowronski, "Switching dynamics of TaO_x -based threshold switching devices," *J. Appl. Phys.*,

vol. 123, no. 11, p. 115 105, Mar. 2018, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.5020070.

[40] Z. Wang, S. Kumar, Y. Nishi, and H.-S. P. Wong, "Transient dynamics of NbO_x threshold switches explained by Poole-Frenkel based thermal feedback mechanism," *Appl. Phys. Lett.*, vol. 112, no. 19, p. 193 503, May 2018, ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.5027152.

[41] J. Lin, S. Guha, and S. Ramanathan, "Vanadium Dioxide Circuits Emulate Neurological Disorders," *Front. Neurosci.*, vol. 12, p. 856, Nov. 2018, ISSN: 1662-453X. DOI: 10.3389/fnins.2018.00856.

[42] M. D. Pickett, G. Medeiros-Ribeiro, and R. S. Williams, "A scalable neuristor built with Mott memristors," *Nat. Mater.*, vol. 12, no. 2, pp. 114–117, Dec. 2012, ISSN: 1476-1122, 1476-4660. DOI: 10.1038/nmat3510.

[43] T. Brown, B. Mann, *et al.*, "Language models are few-shot learners," in *Adv. Neural Inf. Process. Syst.*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, Eds., vol. 33, Curran Associates, Inc., 2020, pp. 1877–1901.

[44] T. Hennen, D. Bedau, *et al.*, "Forming-Free Mott-Oxide Threshold Selector Nanodevice Showing S-Type NDR with High Endurance ($> 10^{12}$ Cycles), Excellent V_{th} Stability ($< 5\%$), Fast (< 10 ns) Switching, and Promising Scaling Properties," in *IEDM*, San Francisco, CA, USA, 2018, pp. 37.5.1–37.5.4. DOI: 10.1109/IEDM.2018.8614618.

[45] T. Hennen, D. Bedau, *et al.*, "Switching Speed Analysis and Controlled Oscillatory Behavior of a Cr-Doped V_2O_3 Threshold Switching Device for Memory Selector and Neuromorphic Computing Application," in *2019 IEEE 11th Int. Mem. Workshop IMW*, Monterey, CA, USA: IEEE, May 2019, pp. 1–4, ISBN: 978-1-72810-981-7. DOI: 10.1109/IMW.2019.8739556.

[46] W. Ma, T. Hennen, *et al.*, "A Mott Insulator-Based Oscillator Circuit for Reservoir Computing," in *2020 IEEE Int. Symp. Circuits Syst. ISCAS*, Seville, Spain: IEEE, Oct. 2020, pp. 1–5, ISBN: 978-1-72813-320-1. DOI: 10.1109/ISCAS45731.2020.9181105.

[47] D. J. Dumin, Ed., *Oxide Reliability: A Summary of Silicon Oxide Wearout, Breakdown, and Reliability* (Selected Topics in Electronics and Systems 23). Singapore: World Scientific, 2002, ISBN: 978-981-02-4842-0.

[48] J. Kim, V. J. Dowling, T. Datta, and Y. V. Pershin, "Holy memristor," *arXiv:2111.11557*, Nov. 2021. arXiv: 2111.11557.

[49] A. A. Sivkov, Y. Xing, K. Y. Cheong, X. Zeng, and F. Zhao, "Investigation of honey thin film as a resistive switching material for nonvolatile memories," *Materials Letters*, vol. 271, p. 127 796, Jul. 2020, ISSN: 0167577X. DOI: 10.1016/j.matlet.2020.127796.

[50] R. Y. Adhikari, N. E. Harmon, and K. P. Williams, "Pristine leaf based electrochemical resistive switching device," *Applied Materials Today*, vol. 24, p. 101 077, Sep. 2021, ISSN: 23529407. DOI: 10.1016/j.apmt.2021.101077.

[51] F. Yasmin Rahman, S. Sarkar, H. Banik, M. Jashim Uddin, D. Bhattacharjee, and S. Arshad Hussain, "Investigation of non volatile resistive switching behaviour using rose petal," *Materials Today: Proceedings*, vol. 65, pp. 2693–2697, 2022, ISSN: 22147853. DOI: 10.1016/j.matpr.2022.05.341.

[52] Z. W. Dlamini, S. Vallabhapurapu, T. S. Mahule, and V. S. Vallabhapurapu, "Electrical conduction and resistive switching in cow milk-based devices prepared using the spin-coat method," *AIP Advances*, vol. 12, no. 9, p. 095 321, Sep. 2022, ISSN: 2158-3226. DOI: 10.1063/5.0098976.

[53] R. Muenstermann, T. Menke, R. Dittmann, and R. Waser, "Coexistence of Filamentary and Homogeneous Resistive Switching in Fe-Doped $SrTiO_3$ Thin-Film Memristive Devices," *Adv. Mater.*, vol. 22, no. 43, pp. 4819–4822, Nov. 2010, ISSN: 09359648. DOI: 10.1002/adma.201001872.

[54] M. D. Pickett, J. Borghetti, J. J. Yang, G. Medeiros-Ribeiro, and R. S. Williams, "Coexistence of Memristance and Negative Differential Resistance in a Nanoscale Metal-Oxide-Metal System," *Adv. Mater.*, vol. 23, no. 15, pp. 1730–1733, Apr. 2011, ISSN: 09359648. DOI: 10.1002/adma.201004497.

[55] J. Bae, I. Hwang, *et al.*, "Coexistence of bi-stable memory and monostable threshold resistance switching phenomena in amorphous NbO_x films," *Appl. Phys. Lett.*, vol. 100, no. 6, p. 062 902, Feb. 2012, ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.3685485.

[56] H. Abbas, Y. Abbas, *et al.*, "The coexistence of threshold and memory switching characteristics of ALD HfO_2 memristor synaptic arrays for energy-efficient neuromorphic computing," *Nanoscale*, vol. 12, no. 26, pp. 14 120–14 134, 2020, ISSN: 2040-3364, 2040-3372. DOI: 10.1039/D0NR02335C.

[57] R. Waser and M. Aono, "Nanoionics-based resistive switching memories," *Nat Mater*, vol. 6, no. 11, pp. 833–840, Nov. 2007, ISSN: 1476-1122. DOI: 10.1038/nmat2023.

[58] 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*, vol. 20, no. 21, p. 215 201, 2009, ISSN: 0957-4484. DOI: 10.1088/0957-4484/20/21/215201.

[59] K. M. Kim, D. S. Jeong, and C. S. Hwang, "Nanofilamentary resistive switching in binary oxide system; a review on the present status and outlook," *Nanotechnology*, vol. 22, no. 25, p. 254 002, Jun. 2011, ISSN: 0957-4484, 1361-6528. DOI: 10.1088/0957-4484/22/25/254002.

[60] G.-S. Park, Y. B. Kim, *et al.*, "In situ observation of filamentary conducting channels in an asymmetric Ta_2O_{5-x}/TaO_{2-x} bilayer structure," *Nat Commun*, vol. 4, no. 1, p. 2382, Dec. 2013, ISSN: 2041-1723. DOI: 10.1038/ncomms3382.

[61] J. J. Yang, M. D. Pickett, X. Li, D. A. A. Ohlberg, D. R. Stewart, and R. S. Williams, "Memristive switching mechanism for metal/oxide/metal nanodevices," *Nature Nanotech*, vol. 3, no. 7, pp. 429–433, Jul. 2008, ISSN: 1748-3387, 1748-3395. DOI: 10.1038/nnano.2008.160.

[62] H. Akinaga and H. Shima, "Resistive random access memory (ReRAM) based on metal oxides," *Proc. IEEE*, vol. 98, no. 12, pp. 2237–2251, 2010.

[63] Y. Chen, "ReRAM: History, Status, and Future," *IEEE Trans. Electron Devices*, vol. 67, no. 4, p. 14, 2020.

[64] C. Nail, G. Molas, *et al.*, "Understanding RRAM endurance, retention and window margin trade-off using experimental results and simulations," in *2016 IEEE Int. Electron Devices Meet. IEDM*, San Francisco, CA, USA: IEEE, Dec. 2016, pp. 4.5.1–4.5.4, ISBN: 978-1-5090-3902-9. DOI: 10.1109/IEDM.2016.7838346.

[65] 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*, vol. 22, no. 50, p. 505 402, Dec. 2011, ISSN: 0957-4484, 1361-6528. DOI: 10.1088/0957-4484/22/50/505402.

[66] A. C. Torrezan, J. P. Strachan, G. Medeiros-Ribeiro, and R. S. Williams, "Sub-nanosecond switching of a tantalum oxide memristor," *Nanotechnology*, vol. 22, no. 48, p. 485 203, Dec. 2011, ISSN: 1361-6528. DOI: 10.1088/0957-4484/22/48/485203.

[67] S. Menzel, M. von Witzleben, V. Havel, and U. Böttger, "The ultimate switching speed limit of redox-based resistive switching devices," *Faraday Discuss.*, vol. 213, pp. 197–213, 2019, ISSN: 1359-6640, 1364-5498. DOI: 10.1039/C8FD00117K.

[68] M. von Witzleben, T. Hennen, A. Kindsmüller, S. Menzel, R. Waser, and U. Böttger, "Study of the SET switching event of VCM-based memories on a picosecond timescale," *J. Appl. Phys.*, vol. 127, no. 20, p. 204 501, May 2020, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/5.0003840.

[69] H.-Y. Chen, S. Brivio, *et al.*, "Resistive random access memory (RRAM) technology: From material, device, selector, 3D integration to bottom-up fabrication," *J Electroceram*, vol. 39, no. 1-4, pp. 21–38, Dec. 2017, ISSN: 1385-3449, 1573-8663. DOI: 10.1007/s10832-017-0095-9.

[70] B. Govoreanu, G. Kar, *et al.*, "10x10nm^2 Hf/HfOx crossbar resistive RAM with excellent performance, reliability and low-energy operation," in *2011 Int. Electron Devices Meet.*, Washington, DC, USA: IEEE, Dec. 2011, pp. 31.6.1–31.6.4, ISBN: 978-1-4577-0504-5. DOI: 10.1109/IEDM.2011.6131652.

[71] N. Raghavan, R. Degraeve, A. Fantini, L. Goux, D. J. Wouters, G. Groeseneken, and M. Jurczak, "Stochastic variability of vacancy filament configuration in ultra-thin dielectric RRAM and its impact on OFF-state reliability," in *2013 IEEE Int. Electron Devices Meet.*, Washington, DC, USA: IEEE, Dec. 2013, pp. 21.1.1–21.1.4, ISBN: 978-1-4799-2306-9. DOI: 10.1109/IEDM.2013.6724674.

[72] R. Degraeve, A. Fantini, *et al.*, "Hourglass concept for RRAM: A dynamic and statistical device model," in *Phys. Fail. Anal. Integr. Circuits IPFA 2014 IEEE 21st Int. Symp. On*, IEEE, 2014, pp. 245–249.

[73] A. Chen and M.-R. Lin, "Variability of resistive switching memories and its impact on crossbar array performance," in *2011 Int. Reliab. Phys. Symp.*, Monterey, CA, USA: IEEE, Apr. 2011, MY.7.1–MY.7.4, ISBN: 978-1-4244-9113-1. DOI: 10.1109/IRPS.2011.5784590.

[74] A. Fantini, L. Goux, *et al.*, "Intrinsic switching variability in HfO_2 RRAM," in *2013 5th IEEE Int. Mem. Workshop*, Monterey, CA, USA: IEEE, May 2013, pp. 30–33, ISBN: 978-1-4673-6169-9. DOI: 10.1109/IMW.2013.6582090.

[75] J. Hubbard, "Electron Correlations in Narrow Energy Bands," *Proc. R. Soc. Math. Phys. Eng. Sci.*, vol. 276, no. 1365, pp. 238–257, Nov. 1963, ISSN: 1364-5021, 1471-2946. DOI: 10.1098/rspa.1963.0204.

[76] N. F. Mott, *Metal-Insulator Transitions*, First edition. London: Taylor & Francis, 1990, ISBN: 978-0-203-21059-8.

[77] V. Dobrosavljevic, "Introduction to metal-insulator transitions," in *Conductor-Insulator Quantum Phase Transitions*, Oxford University Press, 2012, pp. 3–58, ISBN: 978-0-19-959259-3.

[78] L. Cario, C. Vaju, B. Corraze, V. Guiot, and E. Janod, "Electric-Field-Induced Resistive Switching in a Family of Mott Insulators: Towards a New Class of RRAM Memories," *Adv. Mater.*, vol. 22, no. 45, pp. 5193–5197, Dec. 2010, ISSN: 09359648. DOI: 10.1002/adma.201002521.

[79] D. S. Jeong, R. Thomas, R. S. Katiyar, J. F. Scott, H. Kohlstedt, A. Petraru, and C. S. Hwang, "Emerging memories: Resistive switching mechanisms and current status," *Rep. Prog. Phys.*, vol. 75, no. 7, p. 076 502, Jul. 2012, ISSN: 0034-4885, 1361-6633. DOI: 10.1088/0034-4885/75/7/076502.

[80] H. A. Wriedt, "The O-V (Oxygen-Vanadium) system," *Bull. Alloy Phase Diagr.*, vol. 10, no. 3, pp. 271–277, Jun. 1989, ISSN: 0197-0216. DOI: 10.1007/BF02877512.

[81] A. L. Pergament, G. B. Stefanovich, N. A. Kuldin, and A. A. Velichko, "On the Problem of Metal-Insulator Transitions in Vanadium Oxides," *ISRN Condens. Matter Phys.*, vol. 2013, pp. 1–6, 2013, ISSN: 2090-7400. DOI: 10.1155/2013/960627.

[82] H. Schuler, S. Klimm, G. Weissmann, C. Renner, and S. Horn, "Influence of strain on the electronic properties of epitaxial V_2O_3 thin films," *Thin Solid Films*, vol. 299, no. 1, pp. 119–124, 1997. DOI: 10.1016/S0040-6090(96)09399-6.

[83] P. Homm, L. Dillemans, *et al.*, "Collapse of the low temperature insulating state in Cr-doped V_2O_3 thin films," *Appl. Phys. Lett.*, vol. 107, no. 11, p. 111 904, Sep. 2015, ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.4931372.

[84] M. Querré, E. Janod, *et al.*, "Metal–insulator transitions in $(V_{1-x}Cr_x)_2O_3$ thin films deposited by reactive direct current magnetron co-sputtering," *Thin Solid Films*, vol. 617, pp. 56–62, Oct. 2016, ISSN: 00406090. DOI: 10.1016/j.tsf.2015.12.043.

[85] E. B. Thorsteinsson, S. Shayestehaminzadeh, and U. B. Arnalds, "Tuning metal-insulator transitions in epitaxial V_2O_3 thin films," *Appl. Phys. Lett.*, vol. 112, no. 16, p. 161 902, Apr. 2018, ISSN: 0003-6951. DOI: 10.1063/1.5023180.

[86] J. Trastoy, Y. Kalcheim, J. del Valle, I. Valmianski, and I. K. Schuller, "Enhanced metal–insulator transition in V_2O_3 by thermal quenching

after growth," *J. Mater. Sci.*, vol. 53, no. 12, pp. 9131–9137, Jun. 2018, ISSN: 0022-2461, 1573-4803. DOI: 10.1007/s10853-018-2214-7.

[87] C. Adda, L. Cario, *et al.*, "First demonstration of "Leaky Integrate and Fire" artificial neuron behavior on $(V_{0.95}Cr_{0.05})_2O_3$ thin film," *MRS Communications*, vol. 8, no. 3, pp. 835–841, Sep. 2018, ISSN: 2159-6859, 2159-6867. DOI: 10.1557/mrc.2018.90.

[88] F. Rodolakis, P. Hansmann, *et al.*, "Inequivalent Routes across the Mott Transition in V_2O_3 Explored by X-Ray Absorption," *Phys. Rev. Lett.*, vol. 104, no. 4, Jan. 2010, ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.104.047401.

[89] D. B. McWhan, A. Menth, J. P. Remeika, W. F. Brinkman, and T. M. Rice, "Metal-insulator transitions in pure and doped V_2O_3," *Phys. Rev. B*, vol. 7, no. 5, p. 1920, 1973.

[90] J. Meng, B. Zhao, Q. Xu, J. M. Goodwill, J. A. Bain, and M. Skowronski, "Temperature overshoot as the cause of physical changes in resistive switching devices during electro-formation," *J. Appl. Phys.*, vol. 127, no. 23, p. 235 107, Jun. 2020, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/5.0010882.

[91] H. Wan, P. Zhou, L. Ye, Y. Lin, T. Tang, H. Wu, and M. Chi, "In Situ Observation of Compliance-Current Overshoot and Its Effect on Resistive Switching," *IEEE Electron Device Lett.*, vol. 31, no. 3, pp. 246–248, Mar. 2010, ISSN: 0741-3106, 1558-0563. DOI: 10.1109/LED.2009.2039694.

[92] A. Kalantarian, G. Bersuker, *et al.*, "Controlling uniformity of RRAM characteristics through the forming process," in *Reliab. Phys. Symp. IRPS 2012 IEEE Int.*, IEEE, 2012, pp. 6C–4. DOI: 10.1109/IRPS.2012.6241874.

[93] J. Song, D. Lee, *et al.*, "Effects of RESET Current Overshoot and Resistance State on Reliability of RRAM," *IEEE Electron Device Lett.*, vol. 35, no. 6, pp. 636–638, Jun. 2014, ISSN: 0741-3106, 1558-0563. DOI: 10.1109/LED.2014.2316544.

[94] V. Havel, "Transient Processes in Resistive Switching Memory Devices at Ultimate Time Scale Down to Sub-Nanosecond Range," Ph.D. dissertation, RWTH Aachen, 2016.

[95] R. E. Kalman, "Mathematical Description of Linear Dynamical Systems," *Journal of the Society for Industrial and Applied Mathematics Series A Control*, vol. 1, no. 2, pp. 152–192, Jan. 1963, ISSN: 0887-4603. DOI: 10.1137/0301010.

[96] D. K. Schroder, *Semiconductor Material and Device Characterization*, 3. ed. Hoboken, NJ: Wiley-Interscience [u.a.], 2006, ISBN: 978-0-471-73906-7.

[97] "Pulsed I-V Testing for Components and Semiconductor Devices," Keithley, Applications Guide.

[98] M. Lanza, H.-S. P. Wong, *et al.*, "Recommended Methods to Study Resistive Switching Devices," *Adv. Electron. Mater.*, vol. 5, no. 1, p. 1 800 143, Jan. 2019, ISSN: 2199160X. DOI: 10.1002/aelm.201800143.

[99] T. D. Brown, S. Kumar, and R. S. Williams, "Physics-based compact modeling of electro-thermal memristors: Negative differential resistance, local activity, and non-local dynamical bifurcations," *Applied*

Physics Reviews, vol. 9, no. 1, p. 011 308, Mar. 2022, ISSN: 1931-9401. DOI: 10.1063/5.0070558.

[100] A. Fantini, D. J. Wouters, *et al.*, "Intrinsic switching behavior in HfO_2 RRAM by fast electrical measurements on novel 2R test structures," in *Mem. Workshop IMW 2012 4th IEEE Int.*, IEEE, 2012, pp. 1–4.

[101] S. Balatti, S. Ambrogio, Z. Wang, S. Sills, A. Calderoni, N. Ramaswamy, and D. Ielmini, "Voltage-Controlled Cycling Endurance of HfO_x-Based Resistive-Switching Memory," *IEEE Trans. Electron Devices*, vol. 62, no. 10, pp. 3365–3372, Oct. 2015, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2015.2463104.

[102] 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," *Adv. Funct. Mater.*, vol. 21, no. 23, pp. 4487–4492, Dec. 2011, ISSN: 1616301X. DOI: 10.1002/adfm.201101117.

[103] Y. M. Lu, M. Noman, W. Chen, P. A. Salvador, J. A. Bain, and M. Skowronski, "Elimination of high transient currents and electrode damage during electroformation of TiO_2-based resistive switching devices," *J. Phys. Appl. Phys.*, vol. 45, no. 39, p. 395 101, Oct. 2012, ISSN: 0022-3727, 1361-6463. DOI: 10.1088/0022-3727/45/39/395101.

[104] S. Tirano, L. Perniola, *et al.*, "Accurate analysis of parasitic current overshoot during forming operation in RRAMs," *Microelectron. Eng.*, vol. 88, no. 7, pp. 1129–1132, Jul. 2011, ISSN: 01679317. DOI: 10.1016/j.mee.2011.03.062.

[105] K. Kinoshita, K. Tsunoda, Y. Sato, H. Noshiro, S. Yagaki, M. Aoki, and Y. Sugiyama, "Reduction in the reset current in a resistive random access memory consisting of NiOx brought about by reducing a parasitic capacitance," *Appl. Phys. Lett.*, vol. 93, no. 3, p. 033 506, Jul. 2008, ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.2959065.

[106] S. Ambrogio, V. Milo, Z. Wang, S. Balatti, and D. Ielmini, "Analytical Modeling of Current Overshoot in Oxide-Based Resistive Switching Memory (RRAM)," *IEEE Electron Device Lett.*, vol. 37, no. 10, pp. 1268–1271, Oct. 2016, ISSN: 0741-3106, 1558-0563. DOI: 10.1109/LED.2016.2600574.

[107] F. Nardi, D. Ielmini, C. Cagli, S. Spiga, M. Fanciulli, L. Goux, and D. Wouters, "Control of filament size and reduction of reset current below 10μA in NiO resistance switching memories," *Solid-State Electron.*, vol. 58, no. 1, pp. 42–47, Apr. 2011, ISSN: 00381101. DOI: 10.1016/j.sse.2010.11.031.

[108] C. Nguyen, C. Cagli, *et al.*, "Advanced 1T1R test vehicle for RRAM nanosecond-range switching-time resolution and reliability assessment," in *2015 IEEE Int. Integr. Reliab. Workshop IIRW*, South Lake Tahoe, CA: IEEE, Oct. 2015, pp. 17–20, ISBN: 978-1-4673-7395-1. DOI: 10.1109/IIRW.2015.7437059.

[109] J. P. Strachan, A. C. Torrezan, *et al.*, "State Dynamics and Modeling of Tantalum Oxide Memristors," *IEEE Trans. Electron Devices*, vol. 60, no. 7, pp. 2194–2202, Jul. 2013, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2013.2264476.

[110] Y.-S. Fan, L. Zhang, D. Crotti, T. Witters, M. Jurczak, and B. Govoreanu, "Direct Evidence of the Overshoot Suppression in Ta_2O_5-Based Resistive Switching Memory With an Integrated Access Resistor," *IEEE Electron Device Lett.*, vol. 36, no. 10, pp. 1027–1029, Oct. 2015, ISSN: 0741-3106, 1558-0563. DOI: 10.1109/LED.2015.2470081.

[111] A. Hardtdegen, C. La Torre, F. Cuppers, 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 Trans. Electron Devices*, vol. 65, no. 8, pp. 3229–3236, Aug. 2018, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2018.2849872.

[112] M. B. Gonzalez, M. Maestro-Izquierdo, F. Jiménez-Molinos, J. B. Roldán, and F. Campabadal, "Current transient response and role of the internal resistance in HfOx-based memristors," *Appl. Phys. Lett.*, vol. 117, p. 262 902, 2020. DOI: 10.1063/5.0031575.

[113] D. Ielmini, D. Mantegazza, A. Lacaita, A. Pirovano, and F. Pellizzer, "Parasitic reset in the programming transient of PCMs," *IEEE Electron Device Lett.*, vol. 26, no. 11, pp. 799–801, Nov. 2005, ISSN: 0741-3106. DOI: 10.1109/LED.2005.857719.

[114] K. M. Kim, J. J. Yang, *et al.*, "Voltage divider effect for the improvement of variability and endurance of TaOx memristor," *Sci Rep*, vol. 6, no. 1, p. 20 085, Apr. 2016, ISSN: 2045-2322. DOI: 10.1038/srep20085.

[115] Y. Ma, D. Li, *et al.*, "Formation of the Conducting Filament in TaO_x-Resistive Switching Devices by Thermal-Gradient-Induced Cation Accumulation," *ACS Appl. Mater. Interfaces*, vol. 10, no. 27, pp. 23 187–23 197, Jul. 2018, ISSN: 1944-8244, 1944-8252. DOI: 10.1021/acsami.8b03726.

[116] P. R. Shrestha, D. M. Nminibapiel, J. P. Campbell, J. T. Ryan, D. Veksler, H. Baumgart, and K. P. Cheung, "Analysis and Control of RRAM Overshoot Current," *IEEE Trans. Electron Devices*, vol. 65, no. 1, pp. 108–114, Jan. 2018, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2017.2776860.

[117] V. Ostrovskii, P. Fedoseev, Y. Bobrova, and D. Butusov, "Structural and Parametric Identification of Knowm Memristors," *Nanomaterials*, vol. 12, no. 1, p. 63, Dec. 2021, ISSN: 2079-4991. DOI: 10.3390/nano12010063.

[118] D. Ielmini, C. Cagli, and F. Nardi, "Resistance transition in metal oxides induced by electronic threshold switching," *Appl. Phys. Lett.*, vol. 94, no. 6, p. 063 511, Feb. 2009, ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.3081401.

[119] P. Shrestha, D. Nminibapiel, J. P. Campbell, K. P. Cheung, H. Baumgart, S. Deora, and G. Bersuker, "Dependence of the filament resistance on the duration of current overshoot," in *Integr. Reliab. Workshop Final Rep. IRW 2013 IEEE Int.*, IEEE, 2013, pp. 55–58.

[120] A. A. Sharma, M. Noman, M. Abdelmoula, M. Skowronski, and J. A. Bain, "Electronic Instabilities Leading to Electroformation of Binary Metal Oxide-based Resistive Switches," *Adv. Funct. Mater.*, vol. 24, no. 35, pp. 5522–5529, Sep. 2014, ISSN: 1616301X. DOI: 10.1002/adfm.201400461.

[121] K. M. Kim, J. J. Yang, *et al.*, "Low Variability Resistor-Memristor Circuit Masking the Actual Memristor States," *Adv. Electron. Mater.*, vol. 1, no. 6, p. 1 500 095, Jun. 2015, ISSN: 2199160X. DOI: 10.1002/aelm.201500095.

[122] D. C. Gilmer, G. Bersuker, *et al.*, "Effects of RRAM stack configuration on forming voltage and current overshoot," in *Mem. Workshop IMW 2011 3rd IEEE Int.*, IEEE, 2011, pp. 1–4.

[123] W. Chen, W. Lu, *et al.*, "Switching characteristics of W/Zr/HfO_2/TiN ReRAM devices for multi-level cell non-volatile memory applications," *Semicond. Sci. Technol.*, vol. 30, no. 7, p. 075 002, Jul. 2015, ISSN: 0268-1242, 1361-6641. DOI: 10.1088/0268-1242/30/7/075002.

[124] J. Diaz-Fortuny, M. Maestro, J. Martin-Martinez, A. Crespo-Yepes, R. Rodriguez, M. Nafria, and X. Aymerich, "Current-limiting and ultrafast system for the characterization of resistive random access memories," *Review of Scientific Instruments*, vol. 87, no. 6, p. 064 705, Jun. 2016, ISSN: 0034-6748, 1089-7623. DOI: 10.1063/1.4954973.

[125] A. Mikhaylov, A. Belov, *et al.*, "Multilayer Metal-Oxide Memristive Device with Stabilized Resistive Switching," *Adv. Mater. Technol.*, vol. 5, no. 1, p. 1 900 607, Jan. 2020, ISSN: 2365-709X, 2365-709X. DOI: 10.1002/admt.201900607.

[126] D. A. Neamen, *Semiconductor Physics and Devices: Basic Principles*, 3rd ed. Boston: McGraw-Hill, 2003, ISBN: 978-0-07-232107-4.

[127] F.-C. Chiu, "A Review on Conduction Mechanisms in Dielectric Films," *Adv. Mater. Sci. Eng. Adv. Mater. Sci. Eng.*, vol. 2014, 2014, e578168, 2014, ISSN: 1687-8434, 1687-8434. DOI: 10.1155/2014/578168,10.1155/2014/578168.

[128] M. P. Shaw, H. L. Grubin, and I. J. Gastman, "Analysis of an Inhomogeneous Bulk "S-shaped" Negative Differential Conductivity Element in a Circuit Containing Reactive Elements," *IEEE Trans. Electron Devices*, vol. ED-20, pp. 169–177, 1973.

[129] M. P. Shaw, V. V. Mitin, E. Schöll, and H. L. Grubin, Eds., *The Physics of Instabilities in Solid State Electron Devices*. Boston, MA: Springer US, 1992, ISBN: 978-1-4899-2344-8. DOI: 10.1007/978-1-4899-2344-8.

[130] R. E. Burgess, "The A.C. Admittance of Temperature-Dependent Circuit Elements," *Proc. Phys. Soc. B*, vol. 68, no. 10, pp. 766–774, Oct. 1955, ISSN: 0370-1301. DOI: 10.1088/0370-1301/68/10/309.

[131] C. Karakotsou, A. N. Anagnostopoulos, K. Kambas, and J. Spyridelis, "Chaotic voltage oscillations in the negative-differential-resistance region of the *I* - *U* curves of V 2 O 5 crystals," *Phys. Rev. B*, vol. 46, no. 24, pp. 16 144–16 147, Dec. 1992, ISSN: 0163-1829, 1095-3795. DOI: 10.1103/PhysRevB.46.16144.

[132] S. Menzel, U. Böttger, M. Wimmer, and M. Salinga, "Physics of the Switching Kinetics in Resistive Memories," *Adv. Funct. Mater.*, vol. 25, no. 40, pp. 6306–6325, Oct. 2015, ISSN: 1616-3028. DOI: 10.1002/adfm.201500825.

[133] J. B. Roldán, G. González-Cordero, *et al.*, "On the Thermal Models for Resistive Random Access Memory Circuit Simulation," p. 46, 2021.

[134] E. C. Zeeman, "Catastrophe Theory," in *Structural Stability in Physics*, W. Güttinger and H. Eikemeier, Eds., vol. 4, Berlin, Heidelberg: Springer Berlin Heidelberg, 1979, pp. 12–22, ISBN: 978-3-642-67363-4. DOI: 10.1007/978-3-642-67363-4_3.

[135] A. Fuchs, "Dynamical Systems in One and Two Dimensions: A Geometrical Approach," in *Nonlinear Dynamics in Human Behavior*, J. Kacprzyk, R. Huys, and V. K. Jirsa, Eds., vol. 328, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 1–33, ISBN: 978-3-642-16262-6. DOI: 10.1007/978-3-642-16262-6_1.

[136] M. A. Zidan, H. A. H. Fahmy, M. M. Hussain, and K. N. Salama, "Memristor-based memory: The sneak paths problem and solutions," *Microelectron. J.*, vol. 44, no. 2, pp. 176–183, Feb. 2013, ISSN: 00262692. DOI: 10.1016/j.mejo.2012.10.001.

[137] A. Chen, "Nonlinearity and Asymmetry for Device Selection in Cross-Bar Memory Arrays," *IEEE Trans. Electron Devices*, vol. 62, no. 9, pp. 2857–2864, Sep. 2015, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2015.2450712.

[138] L. Zhang, S. Cosemans, D. J. Wouters, G. Groeseneken, M. Jurczak, and B. Govoreanu, "Selector design considerations and requirements for 1 SIR RRAM crossbar array," in *2014 IEEE 6th Int. Mem. Workshop IMW*, May 2014, pp. 1–4. DOI: 10.1109/IMW.2014.6849358.

[139] W. Banerjee, "Challenges and Applications of Emerging Nonvolatile Memory Devices," *Electronics*, vol. 9, no. 6, p. 1029, Jun. 2020, ISSN: 2079-9292. DOI: 10.3390/electronics9061029.

[140] D. Li, J. M. Goodwill, J. A. Bain, and M. Skowronski, "Scaling behavior of oxide-based electrothermal threshold switching devices," *Nanoscale*, vol. 9, no. 37, pp. 14 139–14 148, 2017, ISSN: 2040-3364, 2040-3372. DOI: 10.1039/C7NR03865H.

[141] S. A. Chekol, F. Cuppers, R. Waser, and S. Hoffmann-Eifert, "An $Ag/HfO_2/Pt$ Threshold Switching Device with an Ultra-Low Leakage (< 10 fA), High On/Off Ratio ($> 10^{11}$), and Low Threshold Voltage ($<$ 0.2 V) for Energy-Efficient Neuromorphic Computing," in *2021 IEEE Int. Mem. Workshop IMW*, Dresden, Germany: IEEE, May 2021, pp. 1–4, ISBN: 978-1-72818-517-0. DOI: 10.1109/IMW51353.2021.9439601.

[142] D. Lee, M. Kwak, *et al.*, "Various Threshold Switching Devices for Integrate and Fire Neuron Applications," *Adv. Electron. Mater.*, vol. 5, no. 9, p. 1 800 866, Sep. 2019, ISSN: 2199-160X, 2199-160X. DOI: 10.1002/aelm.201800866.

[143] A. J. Hughes, P. A. Holland, and A. H. Lettington, "Control of holding currents in amorphous threshold switches," *Journal of Non-Crystalline Solids*, vol. 17, no. 1, pp. 89–99, Jan. 1975, ISSN: 00223093. DOI: 10.1016/0022-3093(75)90116-7.

[144] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier, 2013, ISBN: 978-0-12-382010-5. DOI: 10.1016/C2009-0-61160-0.

[145] S. O. Pearson and H. S. G. Anson, "The Neon Tube as a Means of Producing Intermittent Currents," *Proc. Phys. Soc. London*, vol. 34, no. 1,

pp. 204–212, Dec. 1921, ISSN: 1478-7814. DOI: 10.1088/1478-7814/34/1/341.

[146] S. Lavizzari, D. Ielmini, and A. L. Lacaita, "A New Transient Model for Recovery and Relaxation Oscillations in Phase-Change Memories," *IEEE Trans. Electron Devices*, vol. 57, no. 8, pp. 1838–1845, Aug. 2010, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2010.2050963.

[147] A. A. Sharma, Y. Li, M. Skowronski, J. A. Bain, and J. A. Weldon, "High-Frequency TaO_x-Based Compact Oscillators," *IEEE Trans. Electron Devices*, vol. 62, no. 11, pp. 3857–3862, Nov. 2015, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2015.2475623.

[148] A. L. Hodgkin and A. F. Huxley, "Currents carried by sodium and potassium ions through the membrane of the giant axon of *Loligo*," *The Journal of Physiology*, vol. 116, no. 4, pp. 449–472, Apr. 1952, ISSN: 0022-3751, 1469-7793. DOI: 10.1113/jphysiol.1952.sp004717.

[149] A. S. Dmitrichev, V. I. Nekorkin, D. V. Kasatkin, V. V. Klinshov, S. Y. Kirillov, O. V. Maslennikov, and D. S. Shchapin, "Nonlinear dynamical models of neurons: Review," *AND*, vol. 26, no. 4, pp. 5–58, 2018, ISSN: 08696632. DOI: 10.18500/0869-6632-2018-26-4-5-58.

[150] J. Rinzel and G. B. Ermentrout, "Analysis of neural excitability and oscillations," in *Methods in Neuronal Modeling: From Synapses to Networks*, Cambridge, MA, USA: MIT Press, 1989, pp. 135–169, ISBN: 0-262-11133-0.

[151] E. M. Izhikevich, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting* (Computational Neuroscience). Cambridge, Mass: MIT Press, 2007, ISBN: 978-0-262-09043-8.

[152] M. Desroches, J. Rinzel, and S. Rodrigues, "Classification of bursting patterns: A tale of two ducks," *PLoS Comput Biol*, vol. 18, no. 2, H. Berry, Ed., e1009752, Feb. 2022, ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1009752.

[153] R. FitzHugh, "Mathematical models of threshold phenomena in the nerve membrane," *Bulletin of Mathematical Biophysics*, vol. 17, no. 4, pp. 257–278, Dec. 1955, ISSN: 0007-4985, 1522-9602. DOI: 10.1007/BF02477753.

[154] W. E. Sherwood, "FitzHugh–Nagumo Model," in *Encyclopedia of Computational Neuroscience*, D. Jaeger and R. Jung, Eds., New York, NY: Springer New York, 2014, pp. 1–11, ISBN: 978-1-4614-7320-6. DOI: 10.1007/978-1-4614-7320-6_147-1.

[155] R. E. Burgess, "Negative Resistance in Semiconductor Devices," *Can. J. Phys.*, vol. 38, no. 3, pp. 369–375, Mar. 1960, ISSN: 0008-4204, 1208-6045. DOI: 10.1139/p60-038.

[156] H. K. Henisch, "Amorphous Semiconductor Switching," vol. 236, p. 3, 1972.

[157] D. M. Kroll, "Theory of electrical instabilities of mixed electronic and thermal origin," *Phys. Rev. B*, vol. 9, no. 4, pp. 1669–1706, Feb. 1974, ISSN: 0556-2805. DOI: 10.1103/PhysRevB.9.1669.

[158] M. Shaw, "Thermal instability—The precursor to switching in inhomogeneous thin films," *IEEE Trans. Electron Devices*, vol. 26,

no. 11, pp. 1766–1771, Nov. 1979, ISSN: 0018-9383. DOI: 10.1109/T-ED.1979.19683.

[159] Montani, "An electrothermal model for high-field conduction and switching phenomena in TeO_2-V_2O_5 glasses," *J. Non-Cryst. Solids*, vol. 149, no. 3, pp. 249–256, 1992. DOI: 10.1016/0022-3093(92)90073-S.

[160] S. Slesazeck, H. Mähne, *et al.*, "Physical model of threshold switching in NbO_2 based memristors," *RSC Adv.*, vol. 5, no. 124, pp. 102 318–102 322, 2015, ISSN: 2046-2069. DOI: 10.1039/C5RA19300A.

[161] G. A. Gibson, "Designing Negative Differential Resistance Devices Based on Self-Heating," *Adv. Funct. Mater.*, vol. 28, no. 22, p. 1 704 175, May 2018, ISSN: 1616301X. DOI: 10.1002/adfm.201704175.

[162] T. Hennen, E. Wichmann, R. Waser, D. J. Wouters, and D. Bedau, "Stabilizing amplifier with a programmable load line for characterization of nanodevices with negative differential resistance," *Review of Scientific Instruments*, vol. 93, no. 2, p. 024 705, Feb. 2022, ISSN: 0034-6748, 1089-7623. DOI: 10.1063/5.0080532.

[163] T. Hennen, E. Wichmann, *et al.*, "Current-limiting amplifier for high speed measurement of resistive switching data," *Rev. Sci. Instrum.*, vol. 92, p. 054 701, 2021. DOI: 10.1063/5.0047571.

[164] A. Hardtdegen, C. La Torre, H. Zhang, C. Funck, S. Menzel, R. Waser, and S. Hoffmann-Eifert, "Internal Cell Resistance as the Origin of Abrupt Reset Behavior in HfO_2-Based Devices Determined from Current Compliance Series," in *2016 IEEE 8th Int. Mem. Workshop IMW*, Paris, France: IEEE, May 2016, pp. 1–4, ISBN: 978-1-4673-8833-7. DOI: 10.1109/IMW.2016.7495280.

[165] M. J. Ibáñez, D. Barrera, D. Maldonado, R. Yáñez, and J. B. Roldán, "Non-Uniform Spline Quasi-Interpolation to Extract the Series Resistance in Resistive Switching Memristors for Compact Modeling Purposes," *Mathematics*, vol. 9, no. 17, p. 2159, Sep. 2021, ISSN: 2227-7390. DOI: 10.3390/math9172159.

[166] D. Maldonado, F. Aguirre, *et al.*, "Experimental study of the series resistance effect and its impact on the compact modeling of the conduction characteristics of HfO_2-based resistive switching memories," *Journal of Applied Physics*, vol. 130, no. 5, p. 054 503, Aug. 2021, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/5.0055982.

[167] "DS1808 dual log digital potentiometer," Maxim Integrated, Datasheet DS1808Z, 2001, Archived at https://perma.cc/VK7V–8MSG.

[168] "THS309x high-voltage, low-distortion, current-feedback operational amplifiers," Texas Instruments, Datasheet THS3091, 2015, Archived at https://perma.cc/PQX3–RUZD.

[169] E. Wichmann, "Computer Controlled Variable Impedance Circuit for Electrical Characterization of Resistive Switching Cells," Bachelor Thesis, 2019.

[170] P. R. Mickel, A. J. Lohn, and M. J. Marinella, "Detection and characterization of multi-filament evolution during resistive switching," *Appl. Phys. Lett.*, vol. 105, no. 5, p. 053 503, Aug. 2014, ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.4892471.

[171] F. Cüppers, S. Menzel, *et al.*, "Exploiting the switching dynamics of HfO_2-based ReRAM devices for reliable analog memristive behavior," *APL Materials*, vol. 7, no. 9, p. 091 105, Sep. 2019, ISSN: 2166-532X. DOI: 10.1063/1.5108654.

[172] A. García, P. Stasner, and T. Hennen, "Impact of Series Resistance on ReRAM Switching Stability," M.S. thesis, RWTH Aachen, 2022.

[173] T. Hennen, A. Elias, J.-F. Nodin, G. Molas, R. Waser, D. J. Wouters, and D. Bedau, "A high throughput generative vector autoregression model for stochastic synapses," *Front. Neurosci.*, vol. 16, p. 941 753, Aug. 2022, ISSN: 1662-453X. DOI: 10.3389/fnins.2022.941753.

[174] H. Jiang and D. A. Stewart, "Using Dopants to Tune Oxygen Vacancy Formation in Transition Metal Oxide Resistive Memory," *ACS Appl. Mater. Interfaces*, vol. 9, no. 19, pp. 16 296–16 304, May 2017, ISSN: 1944-8244, 1944-8252. DOI: 10.1021/acsami.7b00139.

[175] D. A. Stewart, "Diffusion of oxygen in amorphous tantalum oxide," *Phys. Rev. Materials*, vol. 3, no. 5, p. 055 605, May 2019, ISSN: 2475-9953. DOI: 10.1103/PhysRevMaterials.3.055605.

[176] 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 Appl. Mater. Interfaces*, vol. 13, no. 48, pp. 58 066–58 075, Dec. 2021, ISSN: 1944-8244, 1944-8252. DOI: 10.1021/acsami.1c14667.

[177] A. Ascoli, R. Tetzlaff, Z. Biolek, Z. Kolka, V. Biolkova, and D. Biolek, "The Art of Finding Accurate Memristor Model Solutions," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 5, no. 2, pp. 133–142, Jun. 2015, ISSN: 2156-3357, 2156-3365. DOI: 10.1109/JETCAS.2015.2426493.

[178] I. Messaris, A. Serb, S. Stathopoulos, A. Khiat, S. Nikolaidis, and T. Prodromakis, "A Data-Driven Verilog-A ReRAM Model," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 37, no. 12, pp. 3151–3162, Dec. 2018, ISSN: 0278-0070, 1937-4151. DOI: 10.1109/TCAD.2018.2791468.

[179] D. Ielmini and V. Milo, "Physics-based modeling approaches of resistive switching devices for memory and in-memory computing applications," *J. Comput. Electron.*, vol. 16, no. 4, pp. 1121–1143, Dec. 2017, ISSN: 1569-8025, 1572-8137. DOI: 10.1007/s10825-017-1101-9.

[180] E. Abbaspour, S. Menzel, and C. Jungemann, "Studying the switching variability in redox-based resistive switching devices," *J. Comput. Electron.*, vol. 19, no. 4, pp. 1426–1432, Dec. 2020, ISSN: 1569-8025, 1572-8137. DOI: 10.1007/s10825-020-01537-y.

[181] J. Reuben, D. Fey, and C. Wenger, "A Modeling Methodology for Resistive RAM Based on Stanford-PKU Model With Extended Multilevel Capability," *IEEE Trans. Nanotechnology*, vol. 18, pp. 647–656, 2019, ISSN: 1536-125X, 1941-0085. DOI: 10.1109/TNANO.2019.2922838.

[182] J. Mayer, K. Khairy, and J. Howard, "Drawing an elephant with four complex parameters," *American Journal of Physics*, vol. 78, no. 6, pp. 648–649, Jun. 2010, ISSN: 0002-9505, 1943-2909. DOI: 10.1119/1.3254017.

[183] H. Li, P. Huang, B. Gao, X. Liu, J. Kang, and H.-S. Philip Wong, "Device and Circuit Interaction Analysis of Stochastic Behaviors in Cross-Point RRAM Arrays," *IEEE Trans. Electron Devices*, vol. 64, no. 12, pp. 4928–4936, Dec. 2017, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2017.2766046.

[184] F. Maria Puglisi, L. Larcher, A. Padovani, and P. Pavan, "Bipolar Resistive RAM Based on HfO_2: Physics, Compact Modeling, and Variability Control," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, 2015, ISSN: 2156-3357, 2156-3365.

[185] C. Bengel, A. Siemon, *et al.*, "Variability-Aware Modeling of Filamentary Oxide-Based Bipolar Resistive Switching Cells Using SPICE Level Compact Models," *IEEE Trans. Circuits Syst. Regul. Pap.*, pp. 1–13, 2020, ISSN: 1549-8328, 1558-0806. DOI: 10.1109/TCSI.2020.3018502.

[186] Z. Jiang, Y. Wu, S. Yu, L. Yang, K. Song, Z. Karim, and H.-S. P. Wong, "A Compact Model for Metal–Oxide Resistive Random Access Memory With Experiment Verification," *IEEE Trans. Electron Devices*, vol. 63, no. 5, pp. 1884–1892, May 2016, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2016.2545412.

[187] P.-Y. Chen and S. Yu, "Compact Modeling of RRAM Devices and Its Applications in 1T1R and 1S1R Array Design," *IEEE Trans. Electron Devices*, vol. 62, no. 12, pp. 4022–4028, Dec. 2015, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2015.2492421.

[188] A. Siemon, D. Wouters, S. Hamdioui, and S. Menzel, "Memristive Device Modeling and Circuit Design Exploration for Computation-in-Memory," in *2019 IEEE Int. Symp. Circuits Syst. ISCAS*, Sapporo, Japan: IEEE, May 2019, pp. 1–5, ISBN: 978-1-72810-397-6. DOI: 10.1109/ISCAS.2019.8702600.

[189] M. Bocquet, H. Aziza, *et al.*, "Compact Modeling Solutions for Oxide-Based Resistive Switching Memories (OxRAM)," *JLPEA*, vol. 4, no. 1, pp. 1–14, Jan. 2014, ISSN: 2079-9268. DOI: 10.3390/jlpea4010001.

[190] P. Huang, D. Zhu, *et al.*, "Compact Model of HfO_x-Based Electronic Synaptic Devices for Neuromorphic Computing," *IEEE Trans. Electron Devices*, vol. 64, no. 2, pp. 614–621, Feb. 2017, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2016.2643162.

[191] N. Wald and S. Kvatinsky, "Understanding the influence of device, circuit and environmental variations on real processing in memristive memory using Memristor Aided Logic," *Microelectronics Journal*, vol. 86, pp. 22–33, Apr. 2019, ISSN: 00262692. DOI: 10.1016/j.mejo.2019.02.013.

[192] M. C. Cario and B. L. Nelson, "Autoregressive to anything: Time-series input processes for simulation," *Operations Research Letters*, vol. 19, no. 2, pp. 51–58, Aug. 1996, ISSN: 01676377. DOI: 10.1016/0167-6377(96)00017-X.

[193] D. J. Rezende and S. Mohamed, "Variational Inference with Normalizing Flows," in *Proceedings of the 32nd International Conference on Machine Learning (PMLR)*, vol. 37, Lille, France, 2015, pp. 1530–1538. DOI: 10.5555/3045118.3045281.

[194] A. Chen, "Utilizing the Variability of Resistive Random Access Memory to Implement Reconfigurable Physical Unclonable Functions," *IEEE Electron Device Lett.*, vol. 36, no. 2, pp. 138–140, Feb. 2015, ISSN: 0741-3106, 1558-0563. DOI: 10.1109/LED.2014.2385870.

[195] J. D. Hamilton, *Time Series Analysis*. Princeton, N.J: Princeton University Press, 1994, ISBN: 978-0-691-04289-3.

[196] H. Lütkepohl, *New Introduction to Multiple Time Series Analysis*. Berlin: New York : Springer, 2005, ISBN: 978-3-540-40172-8.

[197] A. Fantini, G. Gorine, *et al.*, "Intrinsic program instability in HfO_2 RRAM and consequences on program algorithms," in *2015 IEEE Int. Electron Devices Meet. IEDM*, Washington, DC, USA: IEEE, Dec. 2015, pp. 7.5.1–7.5.4, ISBN: 978-1-4673-9894-7. DOI: 10.1109/IEDM.2015.7409648.

[198] J. B. Roldán, F. J. Alonso, A. M. Aguilera, D. Maldonado, and M. Lanza, "Time series statistical analysis: A powerful tool to evaluate the variability of resistive switching memories," *Journal of Applied Physics*, vol. 125, no. 17, p. 174 504, May 2019, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.5079409.

[199] S. Seabold and J. Perktold, "Statsmodels: Econometric and Statistical Modeling with Python," in *Python in Science Conference*, Austin, Texas, 2010, pp. 92–96. DOI: 10.25080/Majora-92bf1922-011.

[200] T. Dalgaty, N. Castellani, C. Turck, K.-E. Harabi, D. Querlioz, and E. Vianello, "In situ learning using intrinsic memristor variability via Markov chain Monte Carlo sampling," *Nat Electron*, vol. 4, no. 2, pp. 151–161, Feb. 2021, ISSN: 2520-1131. DOI: 10.1038/s41928-020-00523-3.

[201] B. Butcher, G. Bersuker, *et al.*, "Hot Forming to Improve Memory Window and Uniformity of Low-Power HfOx-Based RRAMs," in *2012 4th IEEE Int. Mem. Workshop*, Milan: IEEE, May 2012, pp. 1–4, ISBN: 978-1-4673-1081-9. DOI: 10.1109/IMW.2012.6213647.

[202] L. Zhao, H.-Y. Chen, *et al.*, "Multi-level control of conductive nano-filament evolution in HfO_2 ReRAM by pulse-train operations," *Nanoscale*, vol. 6, no. 11, pp. 5698–5702, 2014, ISSN: 2040-3364, 2040-3372. DOI: 10.1039/C4NR00500G.

[203] J. Moon, W. Ma, J. H. Shin, F. Cai, C. Du, S. H. Lee, and W. D. Lu, "Temporal data classification and forecasting using a memristor-based reservoir computing system," *Nat. Electron.*, vol. 2, no. 10, pp. 480–487, Oct. 2019, ISSN: 2520-1131. DOI: 10.1038/s41928-019-0313-3.

[204] W. Ma, P.-F. Chiu, W. H. Choi, M. Qin, D. Bedau, and M. Lueker-Boden, "Non-Volatile Memory Array Based Quantization- and Noise-Resilient LSTM Neural Networks," in *2019 IEEE Int. Conf. Rebooting Comput. ICRC*, San Mateo, CA, USA: IEEE, Nov. 2019, pp. 1–9, ISBN: 978-1-72815-221-9. DOI: 10.1109/ICRC.2019.8914713.

[205] T. Besard, C. Foket, and B. De Sutter, "Effective Extensible Programming: Unleashing Julia on GPUs," *IEEE Trans. Parallel Distrib. Syst.*, vol. 30, no. 4, pp. 827–841, Apr. 2019, ISSN: 1045-9219, 1558-2183, 2161-9883. DOI: 10.1109/TPDS.2018.2872064.

[206] T. Hennen, "StochasticSynapses.jl," *Zenodo*, May 2022. DOI: 10.5281/zenodo.6535411.

[207] B. U. Pedroni, S. Joshi, *et al.*, "Memory-Efficient Synaptic Connectivity for Spike-Timing- Dependent Plasticity," *Front. Neurosci.*, vol. 13, p. 357, Apr. 2019, ISSN: 1662-453X. DOI: 10.3389/fnins.2019.00357.

[208] B. U. Pedroni, S. R. Deiss, N. Mysore, and G. Cauwenberghs, "Design Principles of Large-Scale Neuromorphic Systems Centered on High Bandwidth Memory," in *2020 Int. Conf. Rebooting Comput. ICRC*, Atlanta, GA, USA: IEEE, Dec. 2020, pp. 90–94, ISBN: 978-1-66541-975-8. DOI: 10.1109/ICRC2020.2020.00013.

[209] L. V. Kantorovich, "Mathematical Methods of Organizing and Planning Production," *Management Science*, vol. 6, no. 4, pp. 366–422, Jul. 1960, ISSN: 0025-1909, 1526-5501. DOI: 10.1287/mnsc.6.4.366.

[210] S. Park, J. Noh, *et al.*, "Nanoscale RRAM-based synaptic electronics: Toward a neuromorphic computing device," *Nanotechnology*, vol. 24, no. 38, p. 384 009, Sep. 2013, ISSN: 0957-4484, 1361-6528. DOI: 10.1088/0957-4484/24/38/384009.

[211] S. Ambrogio, S. Balatti, *et al.*, "Neuromorphic Learning and Recognition With One-Transistor-One-Resistor Synapses and Bistable Metal Oxide RRAM," *IEEE Trans. Electron Devices*, vol. 63, no. 4, pp. 1508–1515, Apr. 2016, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2016.2526647.

[212] D. Ielmini, "Modeling the Universal Set/Reset Characteristics of Bipolar RRAM by Field- and Temperature-Driven Filament Growth," *IEEE Trans. Electron Devices*, vol. 58, no. 12, pp. 4309–4317, Dec. 2011, ISSN: 0018-9383. DOI: 10.1109/TED.2011.2167513.

[213] F. Nardi, S. Larentis, S. Balatti, D. C. Gilmer, and D. Ielmini, "Resistive Switching by Voltage-Driven Ion Migration in Bipolar RRAM—Part I: Experimental Study," *IEEE Trans. Electron Devices*, vol. 59, no. 9, pp. 2461–2467, Sep. 2012, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2012.2202319.

[214] 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 Trans. Electron Devices*, vol. 62, no. 5, pp. 1561–1567, May 2015, ISSN: 0018-9383. DOI: 10.1109/TED.2015.2411748.

[215] W. Kim, S. Menzel, *et al.*, "Impact of oxygen exchange reaction at the ohmic interface in Ta_2O_5-based ReRAM devices," *Nanoscale*, vol. 8, no. 41, pp. 17 774–17 781, 2016, ISSN: 2040-3364, 2040-3372. DOI: 10.1039/C6NR03810G.

[216] 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," *Phys. Status Solidi A*, vol. 213, no. 2, pp. 316–319, Feb. 2016, ISSN: 18626300. DOI: 10.1002/pssa.201532375.

[217] G. Piccolboni, G. Molas, *et al.*, "Investigation of the potentialities of Vertical Resistive RAM (VRRAM) for neuromorphic applications," in *2015 IEEE Int. Electron Devices Meet. IEDM*, Washington, DC, USA: IEEE,

Dec. 2015, pp. 17.2.1–17.2.4, ISBN: 978-1-4673-9894-7. DOI: 10.1109/IEDM.2015.7409717.

[218] D. Ielmini, F. Nardi, and C. Cagli, "Universal Reset Characteristics of Unipolar and Bipolar Metal-Oxide RRAM," *IEEE Trans. Electron Devices*, vol. 58, no. 10, pp. 3246–3253, Oct. 2011, ISSN: 0018-9383, 1557-9646. DOI: 10.1109/TED.2011.2161088.

[219] J. A. J. Rupp, "Synthesis and Resistive Switching Mechanisms of Mott Insulators based on Undoped and Cr-doped Vanadium Oxide Thin Films," Ph.D. dissertation, RWTH Aachen University, 2020.

[220] J. A. J. Rupp, E. Janod, *et al.*, "Competition between V_2O_3 phases deposited by one-step reactive sputtering process on polycrystalline conducting electrode," *Thin Solid Films*, vol. 705, p. 138 063, Jul. 2020, ISSN: 00406090. DOI: 10.1016/j.tsf.2020.138063.

[221] J. A. J. Rupp, M. Querré, *et al.*, "Different threshold and bipolar resistive switching mechanisms in reactively sputtered amorphous undoped and Cr-doped vanadium oxide thin films," *J. Appl. Phys.*, vol. 123, no. 4, p. 044 502, Jan. 2018, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.5006145.

[222] J. A. J. Rupp, R. Waser, and D. J. Wouters, "Threshold Switching in Amorphous Cr-Doped Vanadium Oxide for New Crossbar Selector," IEEE, May 2016, pp. 1–4, ISBN: 978-1-4673-8833-7. DOI: 10.1109/IMW.2016.7495293.

[223] S. A. Shivashankar and J. M. Honig, "Metal-antiferromagnetic-insulator transition in V_2O_3 alloys," *Phys. Rev. B*, vol. 28, no. 10, pp. 5695–5701, Nov. 1983, ISSN: 0163-1829. DOI: 10.1103/PhysRevB.28.5695.

[224] H. Kuwamoto, J. M. Honig, and J. Appel, "Electrical properties of the $(V_{1-x}Cr_x)_2O_3$ system," *Phys. Rev. B*, vol. 22, no. 6, p. 2626, 1980. DOI: 10.1103/PhysRevB.22.2626.

[225] J. Wong, F. W. Lytle, R. P. Messmer, and D. H. Maylotte, "K-edge absorption spectra of selected vanadium compounds," *Phys. Rev. B*, vol. 30, no. 10, pp. 5596–5610, Nov. 1984, ISSN: 0163-1829. DOI: 10.1103/PhysRevB.30.5596.

[226] D. Haskel, Z. Islam, *et al.*, "Local structural order in the disordered vanadium tetracyanoethylene room-temperature molecule-based magnet," *Phys. Rev. B*, vol. 70, no. 5, Aug. 2004, ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.70.054422.

[227] W. Wisawapipat and R. Kretzschmar, "Solid Phase Speciation and Solubility of Vanadium in Highly Weathered Soils," *Environ. Sci. Technol.*, vol. 51, no. 15, pp. 8254–8262, Aug. 2017, ISSN: 0013-936X, 1520-5851. DOI: 10.1021/acs.est.7b01005.

[228] P. Hansmann, A. Toschi, G. Sangiovanni, T. Saha-Dasgupta, S. Lupi, M. Marsi, and K. Held, "Mott-Hubbard transition in V_2O_3 revisited," *Phys. Status Solidi B*, vol. 250, no. 7, pp. 1251–1264, Jul. 2013, ISSN: 03701972. DOI: 10.1002/pssb.201248476.

[229] J. H. Hur, M.-J. Lee, C. B. Lee, Y.-B. Kim, and C.-J. Kim, "Modeling for bipolar resistive memory switching in transition-metal oxides," *Phys.*

Rev. B, vol. 82, no. 15, p. 155 321, Oct. 2010, ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.82.155321.

[230] 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," in *2014 IEEE Int. Symp. Circuits Syst. ISCAS*, Melbourne VIC, Australia: IEEE, Jun. 2014, pp. 1420–1423, ISBN: 978-1-4799-3432-4. DOI: 10.1109/ISCAS.2014.6865411.

[231] C. Baeumer, C. Funck, *et al.*, "In-Gap States and Band-Like Transport in Memristive Devices," *Nano Lett.*, vol. 19, no. 1, pp. 54–60, Jan. 2019, ISSN: 1530-6984, 1530-6992. DOI: 10.1021/acs.nanolett.8b03023.

[232] I. Baek, M. Lee, *et al.*, "Highly scalable non-volatile resistive memory using simple binary oxide driven by asymmetric unipolar voltage pulses," in *Electron Devices Meet. 2004 IEDM Tech. Dig. IEEE Int.*, Dec. 2004, pp. 587–590. DOI: 10.1109/IEDM.2004.1419228.

[233] H. Mähne, L. Berger, *et al.*, "Filamentary resistive switching in amorphous and polycrystalline Nb_2O_5 thin films," *Solid-State Electronics*, vol. 72, pp. 73–77, Jun. 2012, ISSN: 00381101. DOI: 10.1016/j.sse.2012.01.005.

[234] C. Baeumer, R. Valenta, *et al.*, "Subfilamentary Networks Cause Cycle-to-Cycle Variability in Memristive Devices," *ACS Nano*, vol. 11, no. 7, pp. 6921–6929, Jul. 2017, ISSN: 1936-0851, 1936-086X. DOI: 10.1021/acsnano.7b02113.

[235] In Kyeong Yoo, Bo Soo Kang, Seung Eon Ahn, Chang Bum Lee, Myoung Jae Lee, Gyeong-Su Park, and Xiang-Shu Li, "Fractal Dimension of Conducting Paths in Nickel Oxide (NiO) Thin Films During Resistance Switching," *IEEE Trans. Nanotechnol.*, vol. 9, no. 2, pp. 131–133, Mar. 2010, ISSN: 1536-125X, 1941-0085. DOI: 10.1109/TNANO.2010.2041670.

[236] S. Menzel, S. Tappertzhofen, R. Waser, and I. Valov, "Switching kinetics of electrochemical metallization memory cells," *Phys. Chem. Chem. Phys.*, vol. 15, no. 18, p. 6945, 2013, ISSN: 1463-9076, 1463-9084. DOI: 10.1039/c3cp50738f.

[237] P. Stoliar, M. Rozenberg, E. Janod, B. Corraze, J. Tranchant, and L. Cario, "Nonthermal and purely electronic resistive switching in a Mott memory," *Phys. Rev. B*, vol. 90, no. 4, Jul. 2014, ISSN: 1098-0121, 1550-235X. DOI: 10.1103/PhysRevB.90.045146.

[238] D. P. Kennedy, "Spreading Resistance in Cylindrical Semiconductor Devices," *Journal of Applied Physics*, vol. 31, no. 8, pp. 1490–1497, Aug. 1960, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.1735869.

[239] S. Song and K. P. Moran, "Constriction/spreading resistance model for electronics packaging," in *ASME/JSME Thermal Engineering Conference*, vol. 4, 1995, p. 8.

[240] G. Ellison, "Maximum thermal spreading resistance for rectangular sources and plates with nonunity aspect ratios," *IEEE Trans. Comp. Packag. Technol.*, vol. 26, no. 2, pp. 439–454, Jun. 2003, ISSN: 1521-3331. DOI: 10.1109/TCAPT.2003.815088.

[241] R. Holm and E. Holm, *Electric Contacts: Theory and Application*. Berlin; London: Springer, 2011, ISBN: 978-3-642-05708-3.

[242] T. Chen, T.-Y. Lee, J. Allum, and M. McPartlin, "The thermal scaling: From transistor to array," in *2014 IEEE Radio Freq. Integr. Circuits Symp.*, Tampa, FL, USA: IEEE, Jun. 2014, pp. 123–126, ISBN: 978-1-4799-3864-3. DOI: 10.1109/RFIC.2014.6851675.

[243] S. Long, C. Cagli, D. Ielmini, M. Liu, and J. Suñé, "Analysis and modeling of resistive switching statistics," *J. Appl. Phys.*, p. 20, 2012. DOI: 10.1063/1.3699369.

[244] S. Long, L. Perniola, *et al.*, "Voltage and Power-Controlled Regimes in the Progressive Unipolar RESET Transition of HfO_2-Based RRAM," *Sci. Rep.*, vol. 3, no. 1, Dec. 2013, ISSN: 2045-2322. DOI: 10.1038/srep02929.

[245] M. D. Pickett and R. Stanley Williams, "Sub-100 fJ and sub-nanosecond thermally driven threshold switching in niobium oxide crosspoint nanodevices," *Nanotechnology*, vol. 23, no. 21, p. 215 202, Jun. 2012, ISSN: 0957-4484, 1361-6528. DOI: 10.1088/0957-4484/23/21/215202.

[246] X. Liu, S. Li, S. K. Nandi, D. K. Venkatachalam, and R. G. Elliman, "Threshold switching and electrical self-oscillation in niobium oxide films," *J. Appl. Phys.*, vol. 120, no. 12, p. 124 102, Sep. 2016, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.4963288.

[247] J.-G. Zhang and P. Eklund, "Filament formation in switching devices based on V_2O_5 gel films," *J. Mater. Res.*, vol. 8, no. 03, pp. 558–564, Mar. 1993, ISSN: 0884-2914, 2044-5326. DOI: 10.1557/JMR.1993.0558.

[248] M. Mori, *The finite element method and its applications*. New York: Macmillan, 1986, ISBN: 978-0-02-948621-4.

[249] M. Beschow, "Theoretical investigation of volatile resistive switching in VOx based on an electronic induced thermal instability," Bachelor Thesis, RWTH Aachen, 2019.

[250] X. Zhang, H. Xie, *et al.*, "Thermal and electrical conductivity of a suspended platinum nanofilm," *Appl. Phys. Lett.*, vol. 86, no. 17, p. 171 912, Apr. 2005, ISSN: 0003-6951, 1077-3118. DOI: 10.1063/1.1921350.

[251] A. Albert Irudayaraj, R. Srinivasan, P. Kuppusami, E. Mohandas, S. Kalainathan, and K. Ramachandran, "Photoacoustic measurement of thermal properties of TiN thin films," *J. Mater. Sci.*, vol. 43, no. 3, pp. 1114–1120, Feb. 2008, ISSN: 0022-2461, 1573-4803. DOI: 10.1007/s10853-007-2248-8.

[252] Y. Igasaki, H. Mitsuhashi, K. Azuma, and T. Muto, "Structure and Electrical Properties of Titanium Nitride Films," *Jpn. J. Appl. Phys.*, vol. 17, no. 1, pp. 85–96, Jan. 1978, ISSN: 0021-4922, 1347-4065. DOI: 10.1143/JJAP.17.85.

[253] F. Kauffmann, "Microstructure and properties of titanium nitride/silicon nitride coatings," 2003. DOI: 10.18419/OPUS-6523.

[254] M. Chase, *NIST-JANAF Thermochemical Tables, 4th Edition*, 1998.

[255] M. von Arx, O. Paul, and H. Baltes, "Process-dependent thin-film thermal conductivities for thermal CMOS MEMS," *J. Microelectromech. Syst.*, vol. 9, no. 1, pp. 136–145, Mar. 2000, ISSN: 1057-7157, 1941-0158. DOI: 10.1109/84.825788.

[256] H. S. Dow, W. S. Kim, and J. W. Lee, "Thermal and electrical properties of silicon nitride substrates," *AIP Adv.*, vol. 7, no. 9, p. 095 022, Sep. 2017, ISSN: 2158-3226. DOI: 10.1063/1.4996314.

[257] H. V. Keer, D. L. Dickerson, H. Kuwamoto, H. L. C. Barros, and J. M. Honig, "Heat capacity of pure and doped V_2O_3 single crystals," *J. Solid State Chem.*, vol. 19, no. 1, pp. 95–102, 1976.

[258] J. M. Goodwill and M. Skowronski, "Intrinsic current overshoot during thermal-runaway threshold switching events in TaO_x devices," *J. Appl. Phys.*, vol. 126, no. 3, p. 035 108, Jul. 2019, ISSN: 0021-8979, 1089-7550. DOI: 10.1063/1.5087560.

[259] V. Dubost, T. Cren, *et al.*, "Resistive Switching at the Nanoscale in the Mott Insulator Compound $GaTa_4Se_8$," *Nano Lett.*, vol. 13, no. 8, pp. 3648–3653, Aug. 2013, ISSN: 1530-6984, 1530-6992. DOI: 10.1021/nl401510p.

[260] Y. Ma, J. Goodwill, and M. Skowronski, "Quantification of Compositional Runaway during Electroformation in TaO_x Resistive Switching Devices," in *2019 IEEE 11th Int. Mem. Workshop IMW*, Monterey, CA, USA: IEEE, May 2019, pp. 1–4, ISBN: 978-1-72810-981-7. DOI: 10.1109/IMW.2019.8739727.

Tyler A. Hennen

Education

JUNE 2016	Master of Science, **The University of California**, San Diego Electrical and Computer Engineering: Applied Physics
MAY 2010	Bachelor of Science, **The University of Minnesota**, Minneapolis Physics (*Magna Cum Laude*)
MAY 2010	Bachelor of Science, **The University of Minnesota**, Minneapolis Mathematics

Experience

SEPT 2016 *Current*	**Wissenschaftlicher Mitarbeiter** at RWTH AACHEN UNIVERSITY *Institut für Werkstoffe der Elektrotechnik 2*
SEPT 2014 SEPT 2016	**Graduate Student Researcher** at UC SAN DIEGO *Center for Memory and Recording Research*
JAN 2011 AUG 2014	**Research Engineer** at HGST, A WESTERN DIGITAL COMPANY, San Jose *Magnetic Recording Research (Formerly IBM Storage Division)*
SEPT 2010 JAN 2011	**Teaching Assistant** at THE UNIVERSITY OF MINNESOTA, Minneapolis *Introductory College Physics*
MAY 2007 SEPT 2010	**Research Assistant** at THE UNIVERSITY OF MINNESOTA, Minneapolis *Magnetic Microscopy Center*

Awards

SEPT 2014	UCSD Electrical and Computer Engineering Departmental Fellowship
SEPT 2009	Edmond G. Franklin Scholarship in Physics
JUNE 2008	Undergraduate Research Opportunity Grant

List of Publications

T. Hennen, A. Elias, J. F. Nodin, G. Molas, R. Waser, D. J. Wouters, and D. Bedau, "A high throughput generative vector autoregression model for stochastic synapses," *Frontiers in Neuroscience*, vol. 16, p. 941753, 2022

T. Hennen, E. Wichmann, R. Waser, D. J. Wouters, and D. Bedau, "Stabilizing amplifier with a programmable load line for characterization of nanodevices with negative differential resistance," *Review of Scientific Instruments* 93, 024705, 2022.

T. Hennen, E. Wichmann, A. Elias, J. Lille, O. Mosendz, R. Waser, D. J. Wouters, and D. Bedau, "Current-limiting amplifier for high speed measurement of resistive switching data," *Review of Scientific Instruments* 92, 054701, 2021.

T. Hennen, D. Bedau, J. A. J. Rupp, C. Funck, S. Menzel, M. Grobis, R. Waser, and D. J. Wouters, "Switching speed analysis and controlled oscillatory behavior of a Cr-doped V_2O_3 threshold switching device for memory selector and neuromorphic computing application," *IEEE 11th International Memory Workshop (IMW)*, Monterey, CA, USA, 2019

T. Hennen, D. Bedau, J. A. J. Rupp, C. Funck, S. Menzel, M. Grobis, R. Waser, and D. J. Wouters, "Forming-free Mott-oxide threshold selector nanodevice showing S-type NDR with high endurance (> 10^{12} cycles), excellent V_{th} stability (< 5%), fast (< 10 ns) switching, and promising scaling properties," *2018 IEEE International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, pp. 37.5.1-37.5.4, 2018

W. Ma, **T. Hennen**, M. Lueker-Boden, R. Galbraith, J. Goode, W. H. Choi, P Chiu, J. A. J. Rupp, D. J. Wouters, R. Waser, D. Bedau, "A Mott insulator-based oscillator circuit for reservoir computing," *IEEE International Symposium on Circuits and Systems (ISCAS)*, Seville, Spain, 2020

J. Mohr, **T. Hennen**, D. Bedau, J. Nag, R. Waser, and D. J. Wouters, "Fabrication of highly resistive NiO thin films for nanoelectronic applications," *Advanced Physics Research*, 2200008, 2022

M. von Witzleben, **T. Hennen**, A. Kindsmüller, S. Menzel, R. Waser, and U. Böttger, "Study of the SET Switching Event of VCM-based Memories on a Picosecond Timescale," *Journal of Applied Physics*, vol. 127, no. 20, p. 204501, 2020

C. Funck, C. Bäumer, S. Wiefels, **T. Hennen**, R. Waser, S. Hoffmann-Eifert, R. Dittmann, and S. Menzel, "Comprehensive model for the electronic transport in $Pt/SrTiO_3$ analog memristive devices," *Physical Review B*, vol. 102, no. 3, 2020

D. J. Wouters, S. Menzel, J. A. J. Rupp, **T. Hennen**, and R. Waser, "On the universality of the *I–V* switching characteristics in non-volatile and volatile resistive switching oxides," *Faraday Discussions* Vol. 213, pp. 183-196, 2019

J. Hellerstedt, A. Cahlík, M. Švec, O. Stetsovych, and **T. Hennen**, "Counting Molecules: Python based scheme for automated enumeration and categorization of molecules in scanning tunneling microscopy images," *Software Impacts*, Vol. 12, 100301, 2022