Complex Oxides for Computing Beyond von Neumann
Goossens, Anouk

DOI:
10.33612/diss.766596376

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2023

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 20-12-2023
Chapter 4

Memristive Memory Enhancement by Device Miniaturization for Neuromorphic Computing

Abstract

The areal footprint of memristive devices is a key consideration in material-based computing and for their applications in large-scale architectures. In ionic material relying on filamentary switching mechanisms, nanoscale control of the filament crucial to resistive switching in miniaturised devices, is a roadblock to their successful implementation in such architectures. Areal switching mechanisms have the potential to be more readily scalable while reducing the power. This chapter explores the effect of areal downsizing of Schottky devices, with radial dimension varying between 100 μm - 800 nm. A reliable device fabrication method is developed that allows for minimal device-to-device variability so that the effect of area can be isolated. Using device designs that exploit additional electric field effects at the device edges, an unexpected enhancement in the resistance window when reducing device area is found. The microscopic nature of the switching is studied using scanning transmission electron microscopy on interfaces in the virgin state of the device and after switching to high and low resistance states. An interfacial layer is observed whose physical extent is influenced by the applied electric field, highlighting the role of oxygen vacancy trap migration to memristive switching and augmented by a generalised device modelling that captures the kinetics of the switching process.

This chapter is based on work published in:
4.1 Introduction

The ability of memristors to co-locate memory and computation, and exhibit characteristics absent in digital computing makes them important for novel computing approaches. Given the robust way in which the human brain is able to process large amounts of data with remarkably low power, it is unsurprising that it serves as a source of inspiration to the development of computing beyond using CMOS. As the brain utilises a vast network, downscaling memristive devices is a crucial area of research to develop large scale neuromorphic systems.

Resistive switching has been reported in many ionic materials, where the mechanism can be characterised as either filamentary and occurring through the material bulk between two electrodes or interface-type where switching takes place in a region underneath the area of the electrodes \cite{2}. Considerable research has been devoted to this in the realm of non-volatile conventional filamentary devices. The challenges in their implementation in such novel architectures, besides the requirement for unfavourable electroforming processes, lie in their switching endurance \cite{3}, and their efficacy to exhibit discernible analogue resistance states. Memristive devices that exhibit more than two stable states also greatly enhance integration density because each device can store multiple data bits in an analogue manner.

In valence change memristors, where switching originates from filaments, such behaviour is observed in large areal dimensions but is lost when devices are downcaled and conduction is mediated by a single nanoscale filament causing an abrupt transition between the two resistance states \cite{4}. The current is also not sensitive to the electrode area as only a small area is responsible for the conduction. Further, the effects of Joule heating on filaments are an important consideration as devices shrink; Joule heating can cause a wide distribution of switching voltages and endurance deterioration. These limitations in device stability, endurance and associated enhanced power of operation are major roadblocks in the successful implementation of filamentary devices in large scale architectures.

Herein the areal switching mechanism has some potential advantages. A model system in which this mechanism is dominant is Schottky contacts on Nb-doped SrTiO$_3$ (Nb:STO), formed at the interface with a high work function metal. Distinguishing Nb:STO from conventional semiconductors such as Si, widely used in conventional architectures, is its dielectric permittivity which is comparatively large (300) and is strongly dependent on electric field. This property extends the parameter space for designing functionality: electric fields can be used to tune the barrier height and width relevant for memristive device design. In chapter 3 we have shown that such Schottky contacts form robust memristors, exhibiting non-linear transport and continuous conductance modulation. However, for the applicability of Nb:STO-based memristors as hardware elements for non-von Neumann comput-
ing architecture beyond CMOS, the focus should be on establishing their memristive performance with device miniaturisation, which has not been shown on such semiconducting platforms. In this chapter, we demonstrate that memristive devices of Co Schottky contacts on Nb:STO exhibit an increase in the analogue memristive memory window in devices down to 1 μm, contrary to expectations. Ionic defects are at the heart of memristive behaviour, hence one of the following two scenarios is expected. For a homogeneous areal mechanism, the current density will scale with device area so that the device resistance in both the high resistance state (HRS) and the low resistance state (LRS) scales with the electrode size, but the ratio between them is area independent. Alternatively, the resistance window can be severely reduced or even vanish with downscaling due to insufficient ionic defects. However, we observe an enhancement in the memory window as the device area is reduced, with minimal device-to-device variation, an unforeseen finding.

To understand the microscopic nature of the switching, we conducted scanning transmission electron microscopy (STEM) on virgin samples and on samples subjected to either a positive (set) or negative (reset) voltage. Using integrated differential phase contrast (iDPC) we image oxygen atomic columns next to the heavy metal atomic columns. Virgin samples show the existence of a layer near the interface with neither the perovskite structure of the substrate nor that of the Co electrode. Applying a bias across the interface results in oxygen vacancy movement, which is a key factor controlling the resistance states. These new revelations are consolidated with a mathematical model describing the kinetics of trapping and de-trapping in dielectric materials and relate experimental results to the effective trapping density. Surprisingly, this is found to be larger for smaller junctions, suggesting that an increase in the density of traps is responsible for the increased resistance ratio and attributed to inhomogeneous distribution of the electric field due to device edges.

### 4.2 Device Fabrication

The devices were fabricated using Nb-doped SrTiO$_3$ (001) substrates with a doping concentration of 0.1 wt% from Crystec. SrTiO$_3$ consists of alternating SrO and TiO$_2$ planes along the [001] direction. The as-received substrates have a slight miscut from the exact crystallographic direction and as a result, a mixture of both terminations exists at the surface. It has been shown that the local properties of Schottky barriers grown on the different terminations may differ, hence to minimise the variation of different areas on the substrate a single termination is desired. To ensure that the terminating layer is TiO$_2$, a chemical treatment was carried out with buffered hydrofluoric acid (BHF). A further annealing treatment at 960°C in an O$_2$ flow of 300 ccmin$^{-1}$ to facilitate the reorientation of surface atoms to form an atomically flat and straight
4.2. Device Fabrication

Figure 4.1: Device fabrication. Center: schematic of the fabricated devices on Nb-doped SrTiO$_3$, electrical connections. Black lines represent the varying overall electric fields acting over each area. The field strengths at the interface are also indicated by a colour gradient, showing the fields are weakest in the central area (blue) and strongest around the perimeter (red).

terraced surface. Atomic force microscopy images were taken at different parts of the substrate and confirmed the existence of uniform terraces. The substrate was then coated with a negative resist (AZ nLOF 2020) and using electron beam lithography circles of different areas were patterned. A thick insulation layer of AlO$_x$ was deposited using electron beam evaporation and lift-off was carried out to define a set of direct contacts to the substrate. By means of a second lithography step with a positive resist (950K PMMA), square contact pads were defined, each covering a hole and part of the surrounding AlO$_x$: the dimensions of these pads were identical for
each device to minimise spurious effects arising from significantly different contact resistances. Co (20 nm) and a capping layer of Au (100 nm) were then deposited using electron beam evaporation in high vacuum ($\sim 10^{-6}$ Torr). These steps are shown in Fig. 4.1 and can be summarised:

1. First a surface treatment protocol using buffered hydrofluoric acid (BHF) is carried out to ensure a termination layer of TiO$_2$.
2. The surface preparation is completed by annealing in oxygen with a flow rate of 300 sccm.
3. A layer of AZ nLOF 2020, a negative resist, is spincoated on the substrate.
4. Using electron beam lithography, the resist is patterned into circular pillars.
5. A layer of Al$_2$O$_3$ is deposited using electron beam evaporation.
6. After lift-off, the substrate is covered in a layer of Al$_2$O$_3$ with circular holes that form a direct connection to the underlying substrate.
7. Next, a layer of PMMA 950K, a positive resist, is spincoated on the substrate.
8. Electron beam lithography is used to define square contact pads directly on top of the circular openings.
9. A layer of Co, capped with Au is deposited using electron beam evaporation.
10. The device is completed by lift-off of the excess Co/Au.

### 4.3 Electrical Characterisation

We investigated a series of Co/Nb-doped SrTiO$_3$ devices, where the device area was varied across the series over a range spanning six orders of magnitude ranging from $10^{-13}$ to $10^{-8}$ m$^2$, with radii between 800 nm and 100 μm. The centre of Fig. 4.1 shows a schematic of the device structure used for the electrical measurements. Ag paste is used to attach the chip to a carrier. Electrical measurements are carried out using a probe station; a voltage is applied to a probe placed on an electrode while the bottom of the substrate serves as a back contact for the devices.

#### 4.3.1 Virgin State

After fabrication, we performed small-range voltage sweeps to characterise the virgin states of each device on a chip. The results for devices with radial dimension
from 100 μm to 800 nm are shown in Fig. 4.3, where each sweep followed a voltage sequence from 0 to +1 V to -1 V and back to 0 V. We show four devices of each area, which are plotted in Fig. 4.3(a)-(f).

The current magnitudes for different devices of the same area show no significant differences down to 1 μm, indicating device-to-device variations are minimal. Establishing this is important as this signifies the sole influence of device area in determining the resistance ratio and rules out contributions from device-to-device variation. The 800 nm devices show a greater degree of variation; this is likely due to small differences in their areas and edges arising from the fabrication process and not inherent to the material or due to device fallibility. No significant differences in the

Figure 4.2: State stability and multilevel memristive operation. (a) Current densities of virgin devices. Results are shown for devices of radial dimensions of 100 μm (black), 10 μm (blue) and 1 μm (red) (b) and (c) show the current densities and currents read at +0.3 V during the switching cycles shown in Fig. 4.4, respectively. (d) Current read at 0.3 V after switching between a set voltage of +1 V (black, red and blue) or +2 V (green, purple and orange) and a reset voltage of -2 V (black and green), -2.5 V (red and purple) or -3 V (blue and orange). Each combination was repeated over 100 cycles.
current densities at low bias values are found in the virgin state, confirming that the entire device area contributes to the charge transport (Fig. 4.2(a)). For all the devices, the current gradually increases and exhibits a small hysteretic effect from the virgin state, indicating that no forming step is required.

### 4.3.2 Resistive Switching

Figure 4.4(a)-(f) shows 1000 consecutive current-voltage (I-V) sweeps of these devices. Starting from a set voltage of +2 V, each device is in an LRS, represented by the upper branches. After reaching the reset voltage of -3 V and sweeping back, the devices are switched to an HRS (represented by the lower branches). In all device areas both the set and reset operations remain continuous, indicating the resistive switching retains its analogue nature when downscaling. The cycling endurance was measured for over $10^5$ switching cycles without device failure, illustrating an endurance of $>10^5$. The current in the HRSs scales approximately with area at low bias values, while the low resistance current, is less closely correlated to the area. As a result, the resistance window increases with decreasing device area in both forward and reverse bias. Figure 4.2(b) and (c) show the current density and current at a low read voltage of 0.3 V, respectively. Minimal cycle-to-cycle variations at
4.3. Electrical Characterisation

Figure 4.4: Resistance ratio, cycling endurance and state stability. (a)-(f) show 1000 consecutive current-voltage sweeps from +2 V to -3 V to +2 V at a rate of 1.52 Vs$^{-1}$ for devices of 100 μm down to 800 nm. Starting from a set voltage of +2 V, each device is in an LRS, represented by the upper branch reaching the reset voltage of -3 V and sweeping back, the devices are switched to an HRS (represented by the lower branch).

low reading voltages are found with reproducible switching between clearly distinguishable states without degradation in device performance. This also establishes the low-power operation of these devices after downscaling, which is important for memristor operation. As shown in Appendix C the device-to-device variation remains low down to 1 μm. The variation in the resistance ratio in the 800 nm devices is larger (Appendix C), and will be discussed later.

The set and reset transitions are gradual and highly tunable. To demonstrate this, a 1 μm device was subjected to voltage sweeps varying between different positive (set) and negative (reset) voltages. Figure 4.2(c) shows that a wide range of stable states is available at a low read voltage of +0.3 V. The wide dynamic range combined with the large number of distinct addressable states ensures device reliability and increased memory storage capabilities. Each state maintains a narrow distribution of current values over the 100 cycles shown, reiterating the stability of the switching process.
4. Memristive Memory Enhancement by Device Miniaturization for Neuromorphic Computing

Figure 4.5: Visualisation of oxygen vacancy migration using STEM. iDPC-STEM images of Co/Nb:STO samples in (a) the virgin (unbiased) state, (b) the LRS and (c) the HRS, highlighting the structure close and far from the interface. The perovskite unit cell of STO, showing Sr in green, O in dark red and Ti in light red, viewed along the $<110>$ in (d) the pristine state and (e) with oxygen vacancies. The deficiency of O causes Ti atoms to move away from the vacancies as shown by the arrows. (f) shows a schematic representation of how the interfacial layer is affected by biasing.

4.4 Scanning Transmission Electron Microscopy*

4.4.1 Experimental Details

A microscopy study of the Schottky interface was carried out using STEM. Three types of STEM lamellae were prepared: virgin (unbiased) samples, low resistance state (LRS) samples and high resistance state (HRS) samples. Using a probe station, samples were subjected to bias values of +2 V and -3 V to prepare samples in the LRS and HRS respectively. STEM lamellae were extracted from samples along the $<110>$ direction using a Helios G4 CX dual beam system with a Ga focused ion beam. The lamellae were thinned to make them transparent to electron using the focused ion beam. Imaging was carried out using a Thermo Fisher Scientific Themis Z S/TEM system operating at 300 kV. STEM-High-angle annular dark-field (HAADF)

*Work in collaboration with M. Ahmadi and B. J. Kooi
4.4. Scanning Transmission Electron Microscopy

Images are readily interpretable with atomic columns being bright spots in a dark surrounding, where the brightness of the spots scale with the average atomic number $Z (\sim Z^{1.7})$. This technique is well suited to image heavy elements, but lighter elements, such as oxygen, are harder to detect, and cannot be detected properly when integrated into a matrix with significantly heavier elements (like Sr). Therefore, to gain more insight into the important role played here by the oxygen ions, we utilised here STEM-iDPC instead of STEM-HAADF imaging. This technique uses a four-quadrant annular bright field detector and can be used to acquire the projected local electrostatic potential of the sample (when thin) and has clear advantages over traditional annular bright field (ABF) imaging [5, 6].

4.4.2 Results

Figure 4.5 shows atomic resolution cross-section STEM-integrated Differential Phase Contrast (iDPC) images of the Co/Nb:STO interface for samples in the unbiased virgin condition (Fig. 4.5(a)), the LRS state (Fig. 4.5(b)) and the HRS state (Fig. 4.5(c)). The STEM images in Fig. 4.5(a) show that, apart from a thin interfacial region, the bulk STO consists of a cubic perovskite lattice and no defects are observable. All images taken within the bulk did not show any dislocation and possessed the expected perovskite structure as shown in Fig. 4.5(d). However, the structure close to the interface deviates from this perovskite structure and is deficient in oxygen. The migration of oxygen ions near the interface towards Co causes positively charged Ti ions to be displaced so that they no longer sit equidistantly from the Sr ions along $<001>$. Figure 4.5(e) illustrates how the loss of O ions gives rise to Ti displacements along the $<001>$ direction away from the interface as well as along $<1-10>$ (see Appendix B, Fig. B.3) and is similar to what was reported in ref. [7] in $\text{La}_{0.67}\text{Sr}_{0.33}\text{MnO}_3/\text{Hf}_{0.5}\text{Zr}_{0.5}\text{O}_2$. We believe the creation of this thin layer to be related to the formation of a Schottky barrier. The analysis for a non-memristive interface with Ti contacts can be found in Appendix B, Fig. B.1.

Figure 4.5(b) shows analogous results to Fig. 4.5(a), but now for the sample switched to the LRS, representing the upper branch in Fig. 4.4 after the application of a positive bias voltage of 2 V. Comparing the two figures shows that in the LRS state the extent of the interfacial layer has decreased. This suggests that under the influence of a positive voltage, the labile bonds between O and interfacial Co atoms are broken and oxygen moves back into the STO substrate. A negative bias voltage of -3 V (corresponding to the lower branch in Fig. 4.4), on the other hand, causes oxygen to move from STO to cobalt causing the formation of CoO and more oxygen vacancies in the STO, highlighted by a larger region over which Ti ions are displaced (see Fig. 4.5(c)). This indicates that the formation of the CoO switches the sample to the HRS state. It has been shown [8, 9] that the oxygen vacancy distri-
4. Memristive Memory Enhancement by Device Miniaturization for Neuromorphic Computing

The formation of an oxygen-deficient interfacial layer confirms that in these samples the oxygen vacancies are concentrated near the interface. In this case, it is expected that the application of a positive voltage will cause oxygen vacancies to be repelled from the interface while a negative voltage will cause oxygen vacancies to be attracted to the interface, consistent with our findings. After removing the voltage, the interfacial layer did not reform over time, suggesting the presence of an oxygen-migration blocking layer. These results are summarised in Fig. 4.5(f).

Our results directly confirm the existence of a homogeneous oxygen-deficient layer at the interface. The homogeneous nature of the defect state layer ensures ionic defects are retained with downscaling. We furthermore show that the physical extent of the layer is reduced or extended when a positive or negative voltage is applied respectively. Although the uniform nature of the ionic contribution to switching is now verified, this does not explain the origin of the unexpected enhancement of the resistance window with downscaling. This we discuss next by considering the trapping of electronic charges at oxygen vacancy sites.

4.5 Model

In order to understand how the electrical properties of the devices are influenced by these oxygen vacancies, we consider the interaction between electrons and defect states. This interaction is most strongly evidenced by the retention characteristics, which have a slow decaying component. This behaviour is caused by the detrapping of charges. It has been shown that this occurs over long timescales and the different states will remain clearly distinguishable for long time periods of hours and that the retention time is tunable by the applied stimuli [10]. We utilised short voltage pulses to measure the retention characteristics of each device in both an HRS and LRS. This was done by applying alternating set and reset pulses of +2 V and -3 V respectively, and reading the small-signal current at either +0.3 V or -0.5 V after each writing event. The state retention characteristics of the different devices are shown in Fig. 4.6 for the LRS (red) and HRS (black). Over time, the current in both states tends to an intermediate value. For the LRS, the rate of change follows a power law that is commonly observed for charge trapping under bias in high-κ dielectrics, referred to as the Curie-von Schweidler law.

This law describes a non-Debye type relaxation in dielectrics. Empirical evidence of this behaviour is seen in a wide variety of materials, but the precise physical origin remains unclear. Here we consider the effect of injected electrons becoming trapped in defects states within the dielectric. The space charge generated by these trapped...
4.5. Model

Figure 4.6: Trapping dynamics and Schottky interface energy landscapes. Retention characteristics of differently sized devices read at +0.3 V (a)-(c) and -0.5 V ((d)-(f)) after a set voltage of +2 V (red) or -3 V (black). (g) shows the energy landscape of a Schottky interface in equilibrium when the dielectric constant does not depend on electric field (solid line) and when the dielectric constant is field-dependent (dashed line). $E_F$ and $E_C$ are the Fermi level and conduction band respectively. The energy landscapes at the centre and edge of a device are compared in (h) in equilibrium and (i) in reverse bias. Red circles represent oxygen vacancy states and the green arrow indicates electron tunnelling.

Electrons lowers the electric field, in turn reducing the flow of current through the dielectric.

If we assume that the rate of trapping has no dependence on the location of traps, the electric field, $E$, can be expressed as:

$$E = E_{ap} - \frac{qnx}{\eta} \quad (4.1)$$

where $E_{ap}$ is the applied electric field, $q$ the electric charge, $n$ is the number density of trapped charges, $x$ is the centroid of the trapped charge with respect to the
4. Memristive Memory Enhancement by Device Miniaturization for Neuromorphic Computing

Figure 4.7: Schematic of parameters in section 4.5. $E_{ap}$ and $x$ represent the applied electric field and centroid of trapped charge, defined with respect to the interface, respectively. The number density of trapped charges, $n$, is depicted by the black curve as a function of position in the dielectric.

interface and $\eta$ is the dielectric permittivity. In [11], charge trapping was analysed on the basis of three mechanisms, namely first-order trapping, first-order trapping with Coulombic interactions, and trapping which increases during injection due to the generation of states. The expressions for current they derive are qualitatively similar for each mechanism. Hence, for simplicity, we consider the rate of trapping density to be a decay in first order with the addition of electron-electron interactions. Coulombic repulsion may inactivate trapping sites surrounding a trapped electron. This is included in the rate equation by multiplying a probability factor. If the volume of dielectric rendered inactive by a trap is $h$, then the trapping is reduced by a factor of $\left(1 - \frac{h}{V}\right)$, where $V$ is the volume of the dielectric. For $n$ trapped charges, the factor is $\left(1 - \frac{h}{V}\right)^n$. The trapping rate can be expressed as:

$$\frac{dn}{dt} = (n_0 - n)\sigma \frac{J_{v_{th}}}{q v_d} \left(1 - \frac{h}{V}\right)^n \quad (4.2)$$

where $n_0$ is the maximum number of traps available, $J/q$ is the net flux density, $v_{th}$ and $v_d$ are the thermal and drift velocities respectively, and $\sigma$ is capture cross-section. Assuming the total volume of the dielectric to be substantially larger than the volume deactivated by trapping events so that, $1 \gg h/V$ and $n_0 \gg n$, this expression can be simplified to:

$$\frac{dn}{dt} = n_0 \sigma \frac{J_{v_{th}}}{q v_d} (e)^{-\frac{n}{n}} \quad (4.3)$$
Solving this equation yields the following expression for $n$:

$$ n = \frac{V}{\hbar} \ln \left( \frac{Q}{Q^*} + 1 \right) $$

(4.4)

where $Q = \int J \, dt$ is the total injected charge and

$$ Q^* = \frac{V v_d q}{n_0 h v_{th} \sigma} $$

(4.5)

We express the current in terms of the electric field as:

$$ \ln \left( \frac{J}{J_0} \right) = \frac{E}{E_0} = \frac{1}{E_0} \left( E_{ap} - \frac{V}{\hbar} \frac{q x}{\eta} \ln \left( \frac{Q}{Q^*} + 1 \right) \right) $$

(4.6)

The current follows a decaying power law with time, $J = J_s t^{-\alpha}$, and the injected charge as a function of time is given by:

$$ Q(t) = \int J \, dt = \frac{J_s t^{1-\alpha}}{1 - \alpha} $$

(4.7)

Substituting 4.7 into 4.6 when $Q/Q^* \gg 1$ yields and noting $\beta = \frac{V q x}{\hbar E_0 \sigma}$:

$$ \ln \left( \frac{J}{J_0} \right) = \frac{1}{E_0} \left( E_{ap} - \beta \ln \left( \frac{J_s t^{1-\alpha}}{Q^*(1-\alpha)} \right) \right) $$

$$ = \frac{1}{E_0} \left( E_{ap} - \beta \ln \left( \frac{J_s t}{Q^*(1-\alpha)} \right) - \beta (1 - \alpha) \ln t \right) $$

(4.8)

Comparing 4.8 with $J = J_s t^{-\alpha}$ implies

$$ \alpha = \beta (1 - \alpha) $$

$$ = \frac{\beta}{1 + \beta} $$

(4.9)

and

$$ J_s \approx m E_{ap} + n - \beta \ln(J_s) $$

(4.10)

Where $m$ encompasses several material parameters. Writing $\beta$ in terms of $\alpha$, and since measured currents are less than $10^{-4}$ A, $J_s$ can be neglected in comparison to $\ln(J_s)$, leading to:

$$ \frac{\alpha}{1 - \alpha} \ln(J_s) \approx m E_{ap} + n $$

(4.11)

$\beta$ is positive, we know from Eq. 4.9 that $\alpha$ lies between 0 and 1, and is a monotonically increasing function of $\beta$. Considering that $\beta = \frac{V q x}{\hbar E_0}$, an increase in either the effective density, $V/\hbar$ or in $x$ gives rise to an increase in $\alpha$, with the former being physically more likely.
4. Memristive Memory Enhancement by Device Miniaturization for Neuromorphic Computing

Instead of deriving an explicit expression for the number density of trapped charge, we can also directly relate the trapping rate to the current, as was done in for example [12]. We use $Q_T$ to denote the charge that is trapped when charge $Q$ is injected into the dielectric. The ratio $\frac{dQ_T}{dQ}$ is assumed to be a function of current, i.e.

$$\frac{dQ_T}{dQ} = f\left(\frac{J}{J_0}\right)$$  \hspace{1cm} (4.12)

Substituting Eq. 4.12 into Eq. 4.1 gives:

$$\frac{dE}{dt} = J x \frac{dQ_T}{dQ}$$  \hspace{1cm} (4.13)

where $l$ is the length of the dielectric and $J = \frac{dQ}{dt}$. To relate this to the power law, we assume a solution of the form

$$\frac{dQ_T}{dQ} = \left(\frac{J}{J_0}\right)^\alpha,$$  \hspace{1cm} (4.14)

with $\alpha \geq 0$. A general expression for the current assumes the form:

$$J(E) = J_0(E)^{g(E_0)},$$  \hspace{1cm} (4.15)

where the specific functions are determined by the relevant conduction mechanisms. Specifically here, using Eq. 4.14 we can express the current as

$$J = J_s t^{1/(\alpha+1)}$$  \hspace{1cm} (4.16)

For conduction given by an exponential relation as in Eq. 4.6

$$J_s \propto e^{\frac{(1 - 1/(\alpha+1))V}{V_0}},$$  \hspace{1cm} (4.17)

while for conduction determined by the Poole-Frenkel equation‡, we arrive at:

$$J_s \propto V e^{\frac{(1 - 1/(\alpha+1))V^{1/2}}{V_0}},$$  \hspace{1cm} (4.18)

and for Fowler-Nordheim conduction we get:

$$J_s \propto V^2 e^{\frac{(1 - 1/(\alpha+1))V^{-1}}{V_0}}$$  \hspace{1cm} (4.19)

Here, $V_0$ is a constant.

In Table 4.1, we show the LRS exponents, $\alpha$ for each device. Larger values are observed for smaller devices indicating that the trap density is higher in the smallest device compared to the larger device.

‡Poole-Frenkel emission is a trap assisted thermionic emission mechanism.
Table 4.1: Magnitude of exponents, $\alpha$, extracted by fitting a power-law to the low resistance states in the graphs in Fig. 4.6

<table>
<thead>
<tr>
<th>Radius ($\mu$m)</th>
<th>Read at +0.3 V</th>
<th>Read at -0.5 V</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.85$\pm$0.03</td>
<td>2.17$\pm$0.02</td>
</tr>
<tr>
<td>10</td>
<td>0.47$\pm$0.02</td>
<td>0.987$\pm$0.002</td>
</tr>
<tr>
<td>100</td>
<td>0.041$\pm$0.004</td>
<td>0.626$\pm$0.007</td>
</tr>
</tbody>
</table>

4.6 Discussion

While this model provides a clear correlation between trapping density and device area, it does not give information about the traps; we implicitly take all traps to be of the same kind, while in reality, the nature of traps can vary greatly. The trapping rate can depend on the spatial location of the traps and new traps can be generated via defect migration. For a more precise picture of the mechanism, we need to consider a distribution of traps with respect to their location within the dielectric. Evidenced by the STEM study, oxygen vacancies are the most important class of trapping defects to consider. They are abundantly present in SrTiO$_3$ due to their low formation (0.51 eV$^{[13]}$) and migration (0.62 eV$^{[14]}$) enthalpies and their locations within the energy landscape are well documented$^{[15]}$.

From the discussion above, it is clear that the energy landscape of these Schottky junctions is far more complex than is captured by the most commonly used models that are based solely on parameters of the individual materials forming the contact$^{[16, 17]}$. Transport through these junctions is usually described by the thermionic emission equation, which includes an ideality factor accounting for the deviating transport from this ideal diode equation. This model furthermore does not consider that the interfacial area is not spatially homogeneous and that in devices of finite areas, the boundary of the device will be relevant. In particular, it is known that near the edges crowding of the field lines leads to an enhancement in the field strength which can decrease the barrier width$^{[18, 19]}$. This is supported by the results of the finite element simulations in section 4.7 showing a significant enhancement in the electric field around the edge and when downscaling. From the simulations it is evident that there is still a clear field gradient in the 1 $\mu$m devices, indicating that a further increase in ratio with downscaling can be expected, and the areal field shows no apparent saturation till around 10 nm (Fig. 4.10).

The observed enhancement is especially important in Nb:STO-based memristive devices as the dielectric constant of the substrate strongly depends on electric fields$^{[20, 21]}$. This will further alter the potential landscape of the Schottky interface in
such memristive devices. In particular, the dielectric permittivity of Nb:STO rapidly decreases in the presence of large electric fields which results in a decrease in the effective Schottky barrier width as illustrated in Fig. 4.6(g). Consequently, a large reduction in the barrier width is expected to occur near the device edges (Fig. 4.6(h)). It has also been shown that an electric field can modify the defect states and significantly affect trapping parameters[22].

Given that the charge transport is governed by the potential landscape, this will hugely impact the measured current, pictured in Fig. 4.6(i). Tunnelling through the barrier will be enhanced near the device edges leading to a larger current near the device perimeter. This will be especially important in the LRS where the interface is depleted of trapped charges and the Schottky barrier is narrower, leading to more tunnelling [10 23].

Transport across the interface is comprised of thermionic emission and tunnelling. The thermionic current density is expected to be independent of area and is the dominant mechanism in the HRS at low bias voltages, giving rise to the decreasing current in the HRS around zero with downscaling observed in Fig. 4.4. At higher voltage values, however, tunnelling will also contribute to the current; the tunnelling current density will increase with decreasing area. In Fig. 4.2(b), the current is read at +0.3 V where we expect both thermionic emission and tunnelling to contribute to transport, giving rise to similar currents measured for the 10 and 1 μm devices in the HRS. The tunnelling contribution increases in the LRS, especially in smaller devices due to the larger electric fields, resulting in the observed increase in current density with reducing area.

By applying a potential over the Schottky barrier, the Fermi level is shifted such that tunnelling electrons sample different oxygen vacancy energy levels. As the reverse bias voltage is increased, electrons are gradually exposed to larger ranges of states in which they can become trapped. In addition, in reverse bias, the electric field at the interface becomes larger leading to a reduction in the dielectric constant and a corresponding decrease of the Schottky barrier widths. This decrease in width will be more pronounced in regions closer to the edge due to the local field enhancement. As a result of the narrower barrier, electron-electron scattering will be reduced and the trap states will act as the main barrier for transport. The stronger edge field may additionally facilitate the migration of oxygen vacancies resulting in a higher number of vacancies accumulating around the perimeter. Consequently, the trapping efficiency will be greater near the edge than in the centre. This is a unique effect enabled by the electric field control of the dielectric permittivity, does not occur in conventional semiconductors and is relevant for Nb:STO memristive device design.

We can express the area and perimeter of a device with radius $r$ as $A = \pi r^2$ and
4.7 Simulating the Edge Fields

To visualise the field profiles in our devices we used finite element analysis (COMSOL). In each simulation, the Nb:STO substrate was modelled as a cube with a di-

\[
p = 2\pi r \quad \text{respectively. The ratio of the perimeter to area:}
\]

\[
\frac{p}{A} = \frac{2\pi r}{\pi r^2} = \frac{2}{r},
\]

indicates that the edge effects become more dominant as the device area is reduced. As a result, current flow at the perimeter will constitute a larger percentage to the overall transport behaviour in smaller devices. This explains the enhanced current densities observed when downscaling after applying large bias voltages as well as the larger effective trapping densities for smaller devices. Specifically, this field enhancement around the device edges gives rise to an increase in the dynamic range in smaller devices, and explains the unexpected resistance window scaling.

4.7 Simulating the Edge Fields

Figure 4.8: Electric field at -3 V: along the surface normal (z-direction) for (a)+(d) 1 μm, (b)+(e) 10 μm and (c)+(f) 100 μm devices. The plots on the top row ((a)-(c)) have the same scale bar, with a maximum field value of 3.5 \times 10^6 \text{ Vm}^{-1}. For the bottom row, the scale bar has a maximum value of (d) 2 \times 10^7 \text{ Vm}^{-1}, (e) 2 \times 10^6 \text{ Vm}^{-1} and (f) 3 \times 10^5 \text{ Vm}^{-1}.
4. Memristive Memory Enhancement by Device Miniaturization for Neuromorphic Computing

Figure 4.9: Electric field at +2 V: along the surface normal (z-direction) for (a)+(d) 1 μm, (b)+(e) 10 μm and (c)+(f) 100 μm devices. The plots on the top row ((a)-(c)) have the same scale bar, with a maximum field strength of $-3 \times 10^6$ Vm$^{-1}$. For the bottom row, the scale bar has a maximum value of (d) $-1 \times 10^7$ Vm$^{-1}$, (e) $-1.5 \times 10^6$ Vm$^{-1}$ and (f) $-2 \times 10^5$ Vm$^{-1}$.

We observe a clear increase in the electric field around the device edge compared to the centre; this is most apparent in Fig. 4.8 and 4.9(d)-(f) in which the scale is adjusted for each size. From Fig. 4.8 and 4.9(a)-(c), where the scales are the same for all devices, it is clear that the internal field strength also increases upon downscaling.

For the simulations in Fig. 4.10, the size of the substrate was reduced to improve the resolution of the mesh. This was required to retain the circular nature of the electrodes for the 10 nm devices and was determined not to influence the electric field strength. From Fig. 4.10, it is evident that there is still an observable field gradient in the 50 nm device, indicating that a further increase in ratio with downscaling can
be expected. The areal field shows no apparent saturation till around 10 nm; hence we expect that around 10 nm either (i) the resistance ratio becomes constant or (ii) it continues to increase due to an increasing electric field.

4.8 Conclusions

As a first demonstration of exploiting edge effect-related additional electric fields, our work successfully demonstrates the ability to increase the resistance window by device miniaturisation of interface memristors from 100 μm down to 1 μm, contrary to expectations, with exceptional robustness to device-to-device and cycle variability. Scanning transmission electron microscopy images taken in the virgin, high and low resistance states prove the existence of a homogeneous interfacial layer, deficient in oxygen, whose physical extent is influenced by applying an electric field. This, however, does not explain the enhancement in the resistance window with device downscaling. A model describing the interaction of electrons with oxygen vacancy trap states shows an increase in the effective trapping density with downscaling. The advantage of direct integration of devices on a semiconducting platform of Nb-doped SrTiO$_3$ allows for the locally enhanced fields to controllably tune the interfacial energy landscape at the interface, leading to a greater contribution of edge effects in smaller devices as confirmed by finite element simulations. With rapid advances made in the palette of materials and devices available for neuromorphic hardware, the thrust now should be in their efficient integration on semiconducting platforms for on-chip applications with substantial reduction in areal footprint. In this, our work provides an encouraging direction.
Bibliography


