Advanced TCAD Modeling of HfO_2-based ReRAM: Coupling Redox Reactions and Thermal Effects

Y. Hirchaou
Silvaco-France, Univ. Grenoble Alpes, CEA, Leti
F-38000, Grenoble, France
youssef.hirchaou@silvaco.com

W. Goes
Silvaco Europe Ltd.
Compass Point, St. Ives, UK
wolfgang.goes@silvaco.com

C. Hylin
Silvaco Europe Ltd.
Compass Point, St. Ives, UK
carl.hylin@silvaco.com

P. Blaise
Silvaco-France
38330 Montbonnot Saint-Martin, France
philippe.blaise@silvaco.com

J. Li
Univ. Grenoble Alpes, CEA, Leti
F-38000 Grenoble, France
jing.li@cea.fr

F. Triozon
Univ. Grenoble Alpes, CEA, Leti
F-38000 Grenoble, France
francois.triozon@cea.fr

***Abstract*—This paper presents a TCAD modeling approach for HfO_2-based ReRAM (Resistive Random Access Memory). For describing the switching and retention behaviors of a ReRAM cell, the proposed model includes the essential redox reactions coupled to an electron transport model and to heat generation. The effects of various parameters such as sweep time, and device geometry on the switching behavior are investigated. Simulation results demonstrate that thermal management is crucial both for the reliable operation of ReRAM cells and retention. The proposed TCAD model provides insight into the design and optimization of HfO_2-based ReRAM devices.**

***Index Terms*—OxRAM, TCAD, HfO_2, Redox Reactions, Thermal Effects**

I. Introduction

Oxide-based resistive random access memory (OxRAM) is a promising technology for embedded non-volatile memory devices. Due to the widespread use of HfO_2 in complementary metal-oxide-semiconductor (CMOS) fabrication, OxRAMs can be easily integrated into CMOS technology. In addition, OxRAMs offer a low energy consumption and the potential to be used as multilevel memories for neuromorphic neural networks. While kinetic Monte Carlo (KMC) simulations [1] have advanced the understanding of OxRAMs, faster tools, such as TCAD simulations, appear necessary for device design engineering. For optimizing the OxRAM technology, predictive and physics-based TCAD simulations are vital as they can dramatically speed up the design, fabrication, and commercial use of new microelectronic technologies through the elimination of expensive and time-consuming experimental test wafers during technology adoption. In this article, we present a complete TCAD model that incorporates the relevant electrochemistry coupled with thermal effects to explain the switching and retention behaviors of OxRAMs.

II. Model Description

The model is developed within the Victory Device simulator [2]. We describe a TiN/HfO_2/Ti stack schematized in Fig. 1, where the top electrode (TE) made of Ti can react with oxygen ions moving in the oxide. A cylindrical symmetry around a nanometric conductive filament (CF) is assumed, allowing a 2D description. The oxide layer has a typical thickness of a few nm and a radius of a few tens of nm. The OxRAM is connected in series with a compliance transistor described by a compact model. The species concentrations within the oxide and their reactions with the TE interface are described by continuous functions evolving through drift-diffusion and chemical kinetics equations. A non-stoichiometric oxide includes a large number

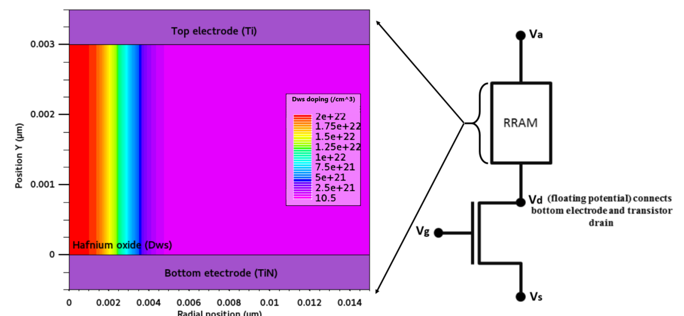

Fig. 1: Schematic of the simulated OxRAM device. (Left) Initial concentration of weak spots $D_{\rm ws}$ in cylindrical symmetry leading to the OxRAM switching. (Right) Applied voltages in a 1T1R structure with a current compliance.

of lattice defects that preferentially occur in clusters and can carry small currents. In our model, they are represented by the so-called weak spots $D_{\rm ws}$ and are assumed to decrease laterally following a Gaussian distribution (see Fig. 1). Upon the application of large electric fields, the $D_{\rm ws}$ can undergo a field-dependent forming reaction $D_{\rm ws} \rightleftharpoons V_{\rm O}^{2+} + I_{\rm O}^{2-}$ where they dissociate into oxygen vacancies $V_{\rm O}^{2+}$ and oxygen interstitials $I_{\rm O}^{2-}$. The $V_{\rm O}$ are immediately neutralized by electron capture as in [3] and become immobile in their neutral charge state due to a high diffusion barrier. Due to the capture processes, the reaction equation effectively modifies to $D_{\rm ws} + 2{\rm e}^- \rightleftharpoons V_{\rm O} + I_{\rm O}^{2-}$. The negatively charged $I_{\rm O}^{2-}$ are moving due to diffusion or drift under an electric field. There, they can also react with the interface of the active electrode, represented by the redox equation $T_{\rm If} + I_{\rm O}^{2-} \rightleftharpoons T_{\rm IfO} + 2{\rm e}^-$, where $T_{\rm IfO}$ and $T_{\rm If}$ denote the oxidized and reduced interface sites, respectively. Due to the reaction kinetics, notable amounts of $V_{\rm O}$ can appear within the oxide. At low concentrations, those defects can also carry an inelastic trap-to-trap tunneling current [4], giving rise to higher conductivity, while they can possibly form some kind of metallic state at higher concentrations.

The mathematical description of chemistry is based on the continuity equation for the species $X_i = D_{\rm ws}, V_{\rm O}, I_{\rm O}^{2-}$:

$$\partial_t X_i = D_{\rm X_i} \nabla \left\{ \frac{z_{X_i} q_0}{k_{\rm B} T} [X_i] \nabla \varphi + \nabla [X_i] \right\} + (G_{\rm X_i} - R_{\rm X_i}) \quad (1)$$

$D_{\rm X_i}$ is the corresponding species diffusivity using the Mott-Gurney model, which considers the field-dependence of the diffusion barriers. The charge of each ion species is given by the product of the charge number $z_{\rm X_i}$ and the elementary charge q_0.

k_B and T are the Boltzmann's constant and the temperature, respectively. The expression in square brackets $[X_i]$ denotes the concentration of the species X_i and φ stands for the electrostatic potential. The first term on the right-hand side of (1) represents the species transport due to drift and diffusion and is omitted for immobile species, ie. $D_{\rm ws}$, $V_{\rm O}$, $T_{\rm If}$, $T_{\rm IfO}$. The second term accounts for the species generation/recombination due to chemical reactions and is considered for all species in our model. Their corresponding reaction rates can be expressed as:

$$R_{D_{\rm ws}} = G_{V_{\rm O}} = G_{I_{\rm O}} = k_{\rm bf}\ [D_{\rm ws}]\ {\rm e}^{-(E_{\rm bf} \mp d_{\rm bf} \cdot E)/(k_{\rm B}T)} \quad (2)$$

$$G_{D_{\rm ws}} = R_{V_{\rm O}} = R_{I_{\rm O}} = k_{\rm br}\ [V_{\rm O}]\ [I_{\rm O}]\ {\rm e}^{-(E_{\rm br} \pm d_{\rm br} \cdot E)/(k_{\rm B}T)} \quad (3)$$

$$R_{T_{\rm IfO}} = G_{T_{\rm If}} = G_{I_{\rm O}} = k_{\rm if}\ [T_{\rm IfO}]\ {\rm e}^{-(E_{\rm if} \pm d_{\rm if} \cdot E)/(k_{\rm B}T)} \quad (4)$$

$$G_{T_{\rm IfO}} = R_{T_{\rm If}} = R_{I_{\rm O}} = k_{\rm ir}\ [T_{\rm If}]\ [I_{\rm O}]\ {\rm e}^{-(E_{\rm ir} \mp d_{\rm ir} \cdot E)/(k_{\rm B}T)} \quad (5)$$

with $k_{\rm bf}$, $k_{\rm br}$, $k_{\rm if}$, and $k_{\rm ir}$ being the reaction rate coefficients ((b,i) for bulk and interface, (f,r) for forward and reverse respectively), and $E_{\rm bf}$, $E_{\rm br}$, $E_{\rm if}$, and $E_{\rm ir}$ being the reaction barriers at zero electric field. These barriers are increased or decreased due to the presence of the electric field $E = -\nabla\varphi$ within the oxide. According to [5], the corresponding dipole moments for the forward and the reverse bulk reaction are of the form:

$$d_{\rm bf} = \frac{2+\varepsilon_{\rm r}}{3} p_{\rm b} \cdot x_{\rm b} \qquad d_{\rm br} = \frac{2+\varepsilon_{\rm r}}{3} p_{\rm b} \cdot (1 - x_{\rm b}) \quad (6)$$

where $\varepsilon_{\rm r}$ is the relative dielectric constant, $p_{\rm b}$ is the change in effective dipole moment, and $x_{\rm b}$ specifies the location of the saddle point of the barrier as a fraction of the reaction coordinate. The same relations are used for the interface reaction dipoles $d_{\rm if}$, $d_{\rm ir}$ with the parameters $p_{\rm i}$ and $x_{\rm i}$.

The species charges are included in the Poisson's equation

$$\nabla(\varepsilon_{\rm r}\varepsilon_0 \nabla\varphi) = -q_0(-n + \sum_i z_{\rm X_i}[X_i]) \quad (7)$$

where n is the electron density and ε_0 stands for the vacuum permittivity. The heat flux q inside a material is described by the Fourier equation

$$C\ \partial_t T = -\nabla \cdot q + (J_{\rm n} \cdot E) \quad (8)$$

$$q = -\kappa \nabla T \quad (9)$$

where C is the heat capacity, κ the thermal conductivity, and $J_{\rm n}$ the current density. The last term in equation (8) accounts for the heat that is released due to electron transport through the oxide. In line with [6], this process is modeled using a Joule heating term. The heat flux q across interfaces is not assumed to be perfect in agreement with [7] and is modeled using the following boundary condition for the interfacial thermal resistance.

$$q \cdot n_{ij} = G(T_i - T_j) \quad (10)$$

i and j denote the materials adjacent to the interface and n_{ij} is the corresponding interface normal vector. G is referred to as the thermal boundary conductance. The electron transport mechanism is a quite complex process as it includes trap-assisted tunneling, trap-to-trap tunneling, maybe even some kind of metallic conduction, and their transition regimes. For this reason, the transport mechanism is phenomenologically described by the effective bridging mobility model, which is included in the electron drift-diffusion equation and has already been successfully used for CBRAM applications [8].

$$\sigma_{\rm eff} = \sigma_{\rm oxide} + w(\delta, f_0, f)\big(\sigma_{\rm filament} - \sigma_{\rm oxide}\big) \quad (11)$$

Initial concentrations	$[D_{\rm ws}]_{max} = 2 \times 10^{22}$ cm^{-3} (with radial Gaussian decay) $[T_{\rm If}] = 10^{14}$ cm^{-2}
Bulk reaction	$E_{\rm bf} = 2$ eV, $E_{\rm br} = 0.5$ eV [9] $k_{\rm bf} = 10^{14}$ s^{-1}, $k_{\rm br} = 10^{-11}$ s^{-1}.cm^3 $p_{\rm b} = 0.8\ q_0$·nm, $\varepsilon_{\rm r} = 21$, $x_{\rm b} = 0.5$
Interface reaction	$E_{\rm if} = 2$ eV, $E_{\rm ir} = 1.15$ eV [9] $k_{\rm if} = 2.63 \times 10^{14}$ s^{-1}, $k_{\rm ir} = 3 \times 10^{-11}$ s^{-1}.cm^3 $p_{\rm i} = 0.7\ q_0$·nm, $\varepsilon_{\rm r} = 21$, $x_{\rm i} = 0.5$
Drift-diffusion of $I_{\rm O}^{2-}$	$D_{I_{\rm O}^{2-}} = D_0 e^{-E_{\rm a}/k_{\rm B}T}$ at zero field $D_0 = 0.4$ cm^2.s^{-1}, $E_{\rm a} = 0.2$ eV
Bridging conductivity	$\sigma_{\rm oxide} = 1.6$ S.cm^{-1} $\sigma_{\rm filament} = 3 \times 10^5$ S.cm^{-1} $f_0 = 0.4$, $\delta = 0.11$
Thermal parameters	$\kappa = 0.03$ W.cm^{-1}.K^{-1} $G = 10^4$ W.cm^{-2}.K^{-1} $C = 2.4$ J.cm^{-3}.K^{-1}

Tab. I: Physical parameters used in the simulations.

$\sigma_{\rm oxide}$ is the conductivity of the host material in the presence of weak spots but without oxygen vacancies. $\sigma_{\rm filament}$ corresponds to the enhanced conductivity of the filament due to the increased concentration of $V_{\rm O}$. The weighting factor w depends on the $V_{\rm O}$ fraction f with respect to its maximum concentration (equal to $[D_{\rm ws}]_{\rm max}$, see Tab. I). w starts to rise linearly at $f = f_0$, with a smoothing parameter δ. All aforementioned equations are solved self-consistently within a mixed-mode simulation, which considers the compliance transistor used for limiting the current and protecting the OxRAM device. This combined circuit and device simulation environment has been implemented into Silvaco's commercial TCAD tool Victory Device [2] in order to enable comprehensive studies of OxRAM devices.

It is noted that the forming of the conductive filament is not simulated throughout this work, but is assumed by imposing a high initial concentration of weak spots $D_{\rm ws}$, and by equilibrating the system of chemical species using a transient simulation at elevated temperatures and zero voltage.

III. Simulation Results

The simulation parameters are listed in Tab. I. For the model evaluation/validation, the most important commonly accepted features (resistive switching during the SET and the RESET phase, current run-away, repeatable cycling) as well as some experimentally observed trends (variation in the oxide thickness and the sweep times, temperature dependence of retention times) will be reproduced in the following.

A. SET Phase

At the beginning of this phase, the defective region within the oxide carries a small current, associated with the high-resistance state (HRS). At a higher biases, this current gives rise to a significant heat generation which raises the temperature within the oxide (see in Fig. 2). Together with the increasing electric field across the oxide, the field-dependent forming reaction is initiated, generating substantial concentrations of $V_{\rm O}$ and $I_{\rm O}^{2-}$. The newly created $V_{\rm O}$ vacancies cause an increase in the trap-to-trap tunneling current through the oxide and thereby trigger a positive feedback loop between the rising current and the increasing heat generation. This feedback loop allows the CF to grow until this process is stopped by the compliance transistor. At this point, the CF has reached its maximum extent (shown in Fig. 3) and the corresponding OxRAM current has arrived at its highest level, associated with the low-resistance state (LRS). The $I_{\rm O}^{2-}$ from the forming reaction drift towards the interface of TE, where they bond to interface sites $T_{\rm If}$. Even though this interface reaction has a relatively high energy

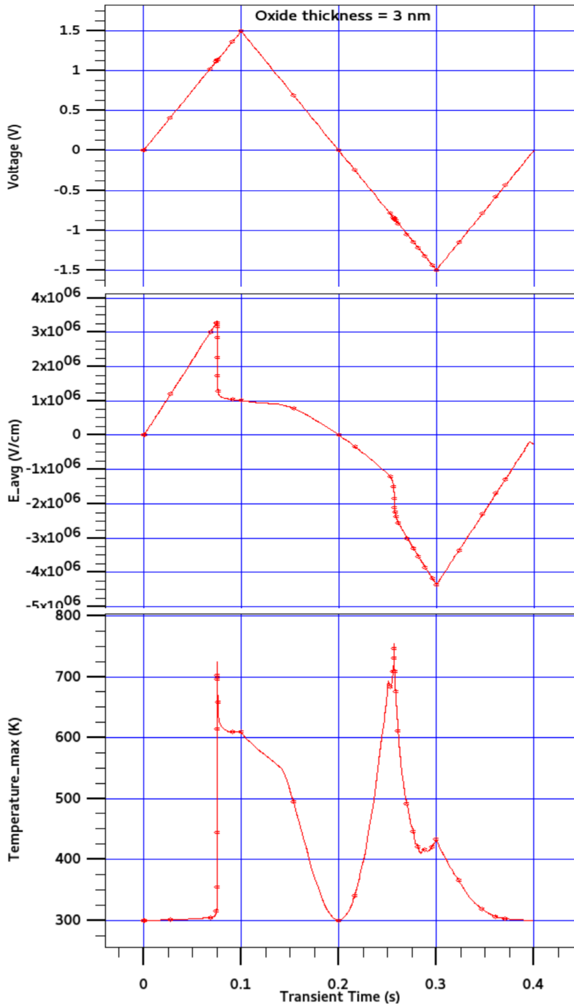

Fig. 2: The voltage wave (top), average electric field (middle), and maximum local temperature (bottom) profiles during one SET and RESET cycle with oxide thickness of 3 nm. The applied bias follows a triangle wave, where the voltage reaches a value of 1.5V during the SET phase (0-200 ms) and a value of -1.5 V during the RESET phase (200-400 ms). One can clearly observe the temperature spikes during SET forward (<100 ms) and RESET forward (<300 ms), which allow for the resistive switching.

barrier, it can be overcome due to the high electric field and high temperatures. During the subsequent SET backward phase, the CF remains stable since the I_O^{2-} are effectively locked at the interface and are unavailable for the reverse mode of the forming reaction.

B. RESET Phase

During this phase, the negative bias applied to the TE is gradually increased, causing an increase in current and heat generation, similar to the SET forward phase. The reverse mode of the interface reaction is activated due to the opposite polarity of the electric field and the higher device temperatures. The electric field pushes the I_O^{2-} towards the interface of the inert bottom electrode (BE), where they recombine with the V_O to form weak spots again. This causes a rupture of the CF and switches the OxRAM device back to its HRS. During the RESET backward phase, the increasing amount of released I_O^{2-} continues to dissolve the disrupted CF (see Fig. 3).

C. Cycling for Different Oxide Thicknesses and Sweep Times

The device characteristics for a SET-RESET cycle is depicted in Fig. 4 for oxide thicknesses in the range between 3 and 4

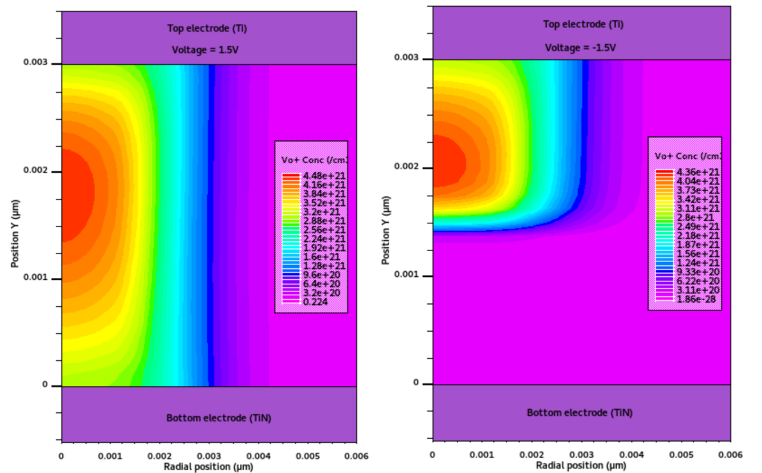

Fig. 3: Distribution of the V_O vacancies in the filamentary region during SET (left) and RESET (right). During SET, the complete CF connects the TE and the BE. During RESET, the conductive filament is broken close to the BE.

nm. The SET and RESET voltage is observed to increase with larger oxide thicknesses (t_{ox}). Due to a larger oxide thickness the electric field inside the oxide is reduced and larger applied voltages are required to initiate the forming and the interface reactions during the SET and the RESET phase. Fig. 5 shows the device characteristics for sweep times ranging from 100 μs to 1 s. Our model predicts an increase in the SET and RESET voltage with shorter sweep times. These trends are in qualitative agreement with recent experimental data obtained at CEA-Leti. Work is in progress both on the simulation and experimental sides to push further this comparison. This behavior can be explained by the fact that the time constants for the forming reaction during SET and the interface reaction during RESET decrease exponentially with the applied electric field. As a direct consequence, those reactions are initiated at lower fields for longer sweep times. In addition, it is demonstrated in Fig. 6 that the hysteresis behavior is nearly unchanged for 5 cycles and is therefore repeatable. As such, the current hysteresis is proven not to be the result of any spurious transient effect - with time constants slightly larger than the sweep time. Fig. 6 also shows the simulation of the current run-away in the absence of the compliance transistor. This transistor prevents this by limiting the electric field across the oxide.

D. Retention

The retention from the SET and RESET states (Fig. 3) is assessed using long transient simulations at different temperatures, during which the current is probed at different time points using a bias of 0.1 V. Fig. 7 shows that the HRS state remains stable at the considered temperatures and times, while the LRS evolves to the HRS, with a retention time which decreases with temperature. The degradation of the LRS state can be explained by the release of I_O^{2-} by the interface and by their recombination with V_O. The stability of the HRS state is explained by the same phenomenon: what remains of the initially broken conducting filament (Fig. 3, right panel) is dissolved, which does not yield a significant change of the resistance.

IV. Discussion

One major challenge in modeling of OxRAM devices is that the time constants for resistive switching span a wide range from μs during resistive switching, up to 10^6 s for retention. Due to the computational costs, KMC studies cannot cover

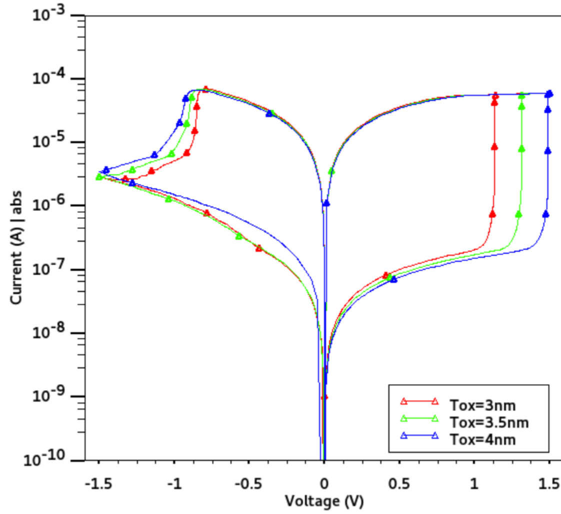

Fig. 4: Effect of different oxide thicknesses on the hysteresis curve for the voltage wave in Fig. 2 (sweep time 100 ms). The SET and RESET voltage clearly increases with larger oxide thicknesses.

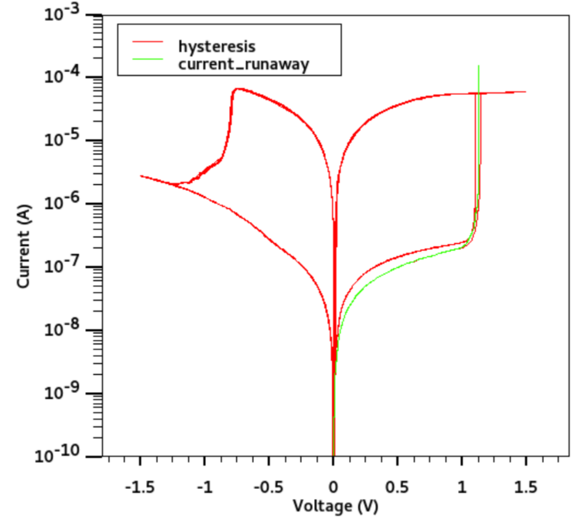

Fig. 6: Repeatable current hysteresis. The device is subject to a number of triangular bias waves. The resulting cycles almost overlap, demonstrating that the resistive switching is a repeatable behavior in our OxRAM model. Green line: current run-away.

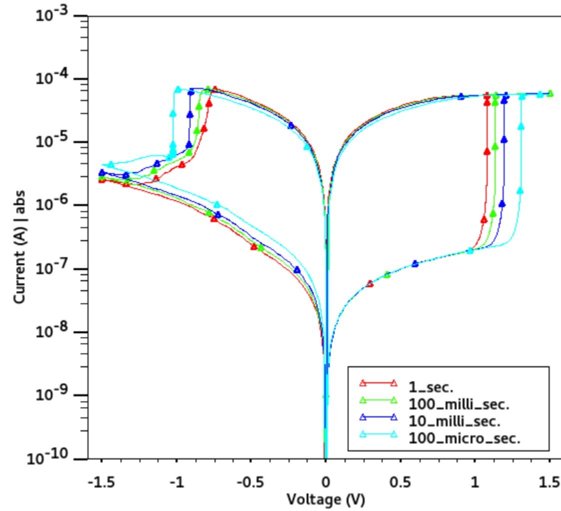

Fig. 5: Effect of different sweep times for a device with oxide thickness of 3 nm. The simulations demonstrate that the SET and RESET voltages are shifted to larger values with shorter sweep times.

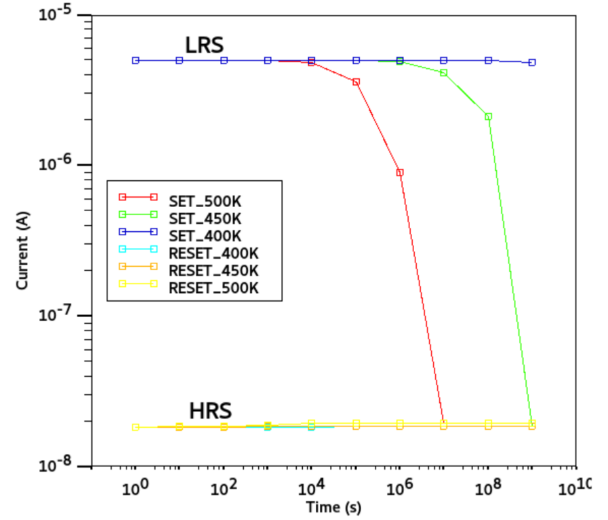

Fig. 7: Retention from SET and RESET states: current probe at voltage V = 0.1 V and different times for T = 400, 450, and 500 K.

the long retention times required to assess the stability of the resistance state and, therefore, do not reveal this critical aspect. In a previous TCAD model [10], the wide range of time constants has been achieved by introducing an empirical model for the interface reaction, which accelerates it at high fields. In this study, however, we focus on correctly incorporating the thermal effects, including the thermal interface conductance [7]. In agreement with the findings of [7], the OxRAM devices can heat up to 900 K during the SET and the RESET phase. As shown in Fig. 2, the highly defective oxide region can carry a substantial current and releases significant amounts of heat. The increased temperatures effectively activate the interface reaction and thereby allow resistive switching down to the microsecond regime. During retention, however, the oxide remains at constant temperature due to the lack of Joule heating and the chemical state of the device is nearly frozen. As a result, the LRS after the SET phase remains stable up to 10^4 s at 500 K, as typically observed in retention experiments.

V. Conclusion

We have developed a quantitative TCAD model, including the key physical and chemical mechanisms in the resistive switching of OxRAM devices. A special focus has been put on correctly incorporating the self-heating effect, which allows for resistive switching in the μs regime at room temperature as well as the long retention up to 24 hours at 500 K. In addition, the model has been demonstrated to predict the trends for different oxide thicknesses and sweep times seen in experiments.

VI. Acknowledgment

This work has been partially funded by ANRT contract 2020/1777.

References

[1] A.Padovani *et al.*, NVMTS (2017), pp. 1-8.
[2] Victory Device User's Manual by Silvaco International, (2023).
[3] A.Padovani *et al.*, JAP 121 155101 (2017).
[4] T.Sadi *et al.*, J.Phys: Cond.Mat. 30 (2018) 084005.
[5] J. McPherson et al., Appl. Phys. Lett, pp. 2121–2123, (2003).
[6] S. Kim et al., Sci.Reports, vol. 3, no. 1680, pp. 1–6, (2013).
[7] S.Deshmuk *et al.*, Sci.Adv. 8, eabkl1514 (2022).
[8] K.Muthuseenu *et al.*, SISPAD (2019), pp. 1-4.
[9] B.Traoré et al., J.Phys.Chem.C 120, 43, pp. 25023–25029 (2016).
[10] W.Goes *et al.*, IMW 2021, pp. 17-20.