MOX-Report No. 36/2022

Physics-based Compact Modelling of the Analog Dynamics of HfOx Resistive Memories

Vaccaro, F.; Brivio, S.; Perotto, S.; Mauri, A.G.; Spiga, S.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it http://mox.polimi.it

Physics-based Compact Modelling of the Analog Dynamics of HfO_x Resistive Memories

F. Vaccaro[1,2], S. Brivio[1], S. Perotto[2], A. G. Mauri[3] and S. Spiga[1]

May 15, 2022

[1] CNR-IMM, Unit of Agrate Brianza, via C. Olivetti 2, Agrate Brianza, Italy
[2] MOX– Dipartimento di Matematica, Politecnico di Milano, Piazza L. da Vinci, 32, Italy
[3] Dipartimento di Matematica, Politecnico di Milano, Piazza L. da Vinci, 32, Italy

Abstract

Resistive random access memories (RRAMs) constitute a class of memristive devices particularly appealing for bio-inspired computing schemes. In particular, the possibility of achieving analog control of the electrical conductivity of RRAM devices can be exploited to mimic the behaviour of biological synapses in neuromorphic systems. With a view to neuromorphic computing applications, it turns out to be crucial to guarantee some features, among which a detailed device characterization, a mathematical modelling comprehensive of all the key features of the device both in quasi-static and dynamic conditions, a description of the variability due to the inherently stochasticity of the processes involved in the switching transitions. In this paper, starting from experimental data, we provide a modelling and simulation framework to reproduce the operative analog behaviour of HfO_x-based RRAM devices under train of programming pulses both in the analog and binary operation mode. To this aim, we have calibrated the model by using a single set of parameters for the quasi-static current-voltage characteristics as well as switching kinetics and device dynamics. The physics-based compact model here settled captures the difference between the SET and the RESET processes in the I-V characteristics, as well as the device memory window both for strong and weak programming conditions. Moreover, the model reproduces the correct slopes of the highly non-linear kinetics curves over several orders of magnitudes in time, and the dynamic device response including the inherent device variability.

1 Introduction

Resistive random access memories (RRAMs), belonging to the broader class of memristive devices, are two-terminal devices able to settle in different resistance states upon proper voltage application [1]. These devices have recently achieved a massive interest for non-traditional computing schemes, such as in-memory and bio-inspired computing [2, 3]. Indeed, RRAMs can implement hardware synaptic elements in neural networks, by exploiting the possibility to achieve multiple conductance levels. As a matter of fact, neural

network applications have been growing in recent years, thus posing a severe problem of energy consumption, especially when dealing with the network training process. In this specific regard, training protocols are generally implemented by exploiting the switching dynamics of the devices, i.e., the analog or multilevel operation in RRAMs in response to repeated identical stimuli [4, 5, 6, 7, 8, 9, 10]. Such programming strategy is the main challenge that distinguishes the use of memory devices for neuromorphic computing with respect to their employment in non-volatile memory application, where only the static values of the resistance are strictly relevant. On the other hand, the engineering and understanding of analog dynamics in RRAM devices is still ongoing, with the specific need to join experiments with a device modelling able to address all the key features of analog RRAM devices.

The aim of the present work is therefore to provide such a missing piece of information, i.e., the combination of an experimental characterization with a physics-based modeling for the dynamic behaviour of RRAM devices when programmed under trains of pulses. We consider RRAM devices based on filamentary switching in oxide, which have reached good maturity as binary storage elements, and which have been recently studied also as analog memories for neuromorphic computation [6, 11, 12]. The switching mechanism is related to the formation and the dissolution of conductive filaments in oxide shorting two metal electrodes. The formation and the dissolution of the conducting filaments is usually ascribed to the motion of oxygen vacancies [13, 14, 15]. For this class of RRAMs, also named valence change RRAMs in the literature, various compact models have been proposed in order to reproduce the experimental quasi-static characteristic [16], the switching kinetics (a measure of the switching speed) and the endurance of RRAMs with their statistical behavior [17, 18]. In turn, few works provide a comprehensive description of the device dynamics which includes also analog resistance modulation [19, 20]. Moreover, the fitting of the experimental dynamics is usually performed by means of analytical behavioral models [21, 7, 22], or using higher dimensional models [23, 24]. In this paper, we show a detailed characterization of the device under various programming conditions. The physics-based compact model we propose proves to be able to comprehensively reproduce the quasi-static device behavior, the magnitude, the speed, and the time evolution of the device resistance variation during the switching, which improves the existing literature with particular reference to [16]. The variability in response to repeated identical stimuli in binary and analog switching is correctly described by the model. The presented equations allow to model the complex interplay between resistance and temperature dynamics, that eventually leads to the resistive switching. The novelty of the work consists in highlighting that the same description of the feedback processes involved during the quasi-static operational mode can be exploited to reproduce the device behavior also in a pulse regime, in a wide range of time scales, and using a single set of model parameters. In particular, the actual shape of the time evolution of the resistance can be explained and reproduced with the same mechanisms, even in presence of variability.

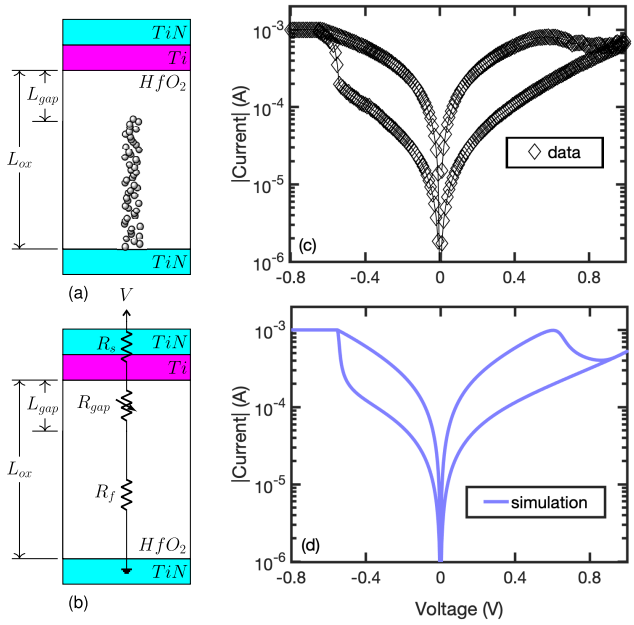

Figure 1: TiN/Ti/HfO_2/TiN device: (*a*) device structure; (*b*) equivalent circuit model; (*c*) experimental I-V characteristic; (*d*) simulated I-V characteristic. In (*a*) the sketch of the filament is representative of the partially dissolved state after RESET.

2 Experimental details and modelling framework

2.1 Device fabrication and testing

The devices considered in this work, both for experimental measurements and simulations, are based on the 40 nm TiN (top electrode)/10 nm Ti/5.5 nm HfO_2/40 nm TiN (bottom electrode) structure (see figure 1(a) for a sketch). HfO_2 is grown at 300∘C by atomic layer deposition (in a Savannah 200 reactor, Cambridge) using the $MeCp_2HfMe(OMe)Hf$ precursor as Hf source, and H_2O as oxygen source. The Ti and TiN layers are deposited by RF magnetron sputtering using only Ar and mixed Ar/N_2 environment, respectively. The device area is patterned by lift-off process. Device electrical testing, either in quasi-static conditions or under applied trains of pulses, has been performed through a Keysight B1500A semiconductor parameter analyser. During the forming process, a quasi-static voltage ramp from 0 to -2 V, with a current compliance of 1 mA (data not shown), creates a conductive filament shorting the two electrodes. Then, RESET and SET operations partially dissolve and re-instate the filament, respectively [25]. Figure 1(c) shows a representative experimental quasi-static current-voltage (I-V) characteristic (with sweep rate $|SR|$ = 0.1 V/s) with negative voltage SET transition, from high resistance state (HRS) to low resistance state (LRS), and positive voltage RESET transition, from LRS to HRS. It is worth noting that the devices we consider can be operated in bipolar switching mode, namely with negative SET/positive RESET voltages or vice versa, thanks to a complemen-

tary switching behaviour [16, 4, 26]. In particular, except for the data in figure 4, we refer to a negative SET/positive RESET voltage case throughout the whole paper. The devices have been then characterized with trains of identical pulses by a variable amplitude and width, as described in sections 3.2 and 3.3, and in the first section of the supplementary material.

2.2 Description of the modelling framework

The electrical experimental data have been simulated by resorting to a compact model implemented in MATLAB. Before analyzing the simulation results into details, we provide a description of the adopted modelling framework. It is based on the equivalent circuit (figure 1(b)) of the TiN/Ti/HfO_2/TiN device (figure 1(a)). According to [16, 27], we assume that the oxide of a filamentary RRAM can be divided into two regions, i.e., the gap, the part of the oxide near the active electrode where the switching is supposed to take place, and the filament, which represents an extension of the electrode in the oxide. The filament region is supposed to act as a reservoir of oxygen vacancies during the switching, and to well conduct for the whole duration of the switching processes [17]. The gap region is modeled as a variable resistor which depends on the oxygen vacancies density N, while the filament region is modeled as a resistor with a fixed oxygen vacancies density N_f. Moreover, we suppose that the electron conduction mechanism in the filament is a thermally-enhanced ohmic conduction [28], which can be described as conduction in band with temperature-dependent mobility [16, 24]. It follows that the gap and the filament resistances are defined by

$$R_{gap} = R_{gap}(N,T) = \frac{L_{gap}}{e\mu_{n0}\exp(\frac{-\Delta E_{ac}}{k_B T})NA}, \tag{1}$$

$$R_f = R_f(T) = \frac{L_f}{e\mu_{n0}\exp(\frac{-\Delta E_{ac}}{k_B T})N_f A}, \tag{2}$$

respectively, where T is the temperature, $L_{gap} = L_{ox} - L_f$ is the length of the gap, with L_{ox} the length of the oxide and L_f the length of the filament, e is the elementary charge, μ_{n0} is the electron mobility, ΔE_{ac} is the electron activation energy, k_B is the Boltzmann constant and $A = \pi r_f^2$ is the section area of the conductive filament, with r_f the filament radius. We consider also a series resistance R_s to take into account ohmic losses at the contacts and potential resistive parasitics (we refer to figure 1(b) for a sketch of all the resistances involved in the proposed model).

The dynamics of the two state variables, N and T, is described through the system of ordinary differential equations

$$\begin{cases} \dfrac{dN}{dt} = -\dfrac{1}{z_{V_0} e A L_{gap}} I_{ion}(N,T) \\ \dfrac{dT}{dt} = \dfrac{I_{el}(N,T) V_{R_{gap}}}{C_{th}} - \dfrac{T - T_0}{C_{th} R_{th}} \end{cases} \tag{3}$$

where the equation for the state variable T coincides with the Newton's law of cooling [29], whereas a standard rate equation [16] is adopted to describe the dynamic of the oxygen va-

cancies density N. In more detail, z_{V_0} denotes the charge number of the oxygen vacancies, I_{ion} and I_{el} are, respectively, the ion and the electrical currents (see figure 1(b)), $V_{R_{gap}}$ is the voltage drop across the gap variable resistor, T_0 is the room temperature, while R_{th} and C_{th} denote the thermal resistance and the capacitance of the oxide, respectively. R_{th} is defined by $L_{gap}/(k_{th}A)$, where k_{th} is the thermal conductivity of the oxide here described by a piece-wise constant model, with a higher value during the RESET phase (k_{th}^{RESET}) with respect to the SET transition (k_{th}^{SET}). This choice takes into account that the accumulation of oxygen vacancies in the gap region enhances the heat conduction during the transition from the HRS to the LRS. Concerning I_{el} flowing in the series of resistances, it is evaluated through the Ohm's law, by dividing the voltage V applied to the RRAM device (see figure 1(b)) by the sum of the resistances, R_{gap}, R_f, R_s. Following [16], we define I_{ion} as the sum of the drift and the diffusion contributions to the ion hopping conduction, with an electric field-induced barrier lowering effect, being

$$I_{ion} = I_{ion,drift} + I_{ion,diffusion} \tag{4}$$

$$= AC\left(\bar{N}\sinh\left(\frac{z_{V_o}eaE}{2k_BT}\right) + \frac{a}{2}\frac{dN}{dx}\cosh\left(\frac{z_{V_o}eaE}{2k_BT}\right)\right), \tag{5}$$

with

$$C = 2z_{V_o}ea\nu_0\exp\left(-\frac{\Delta E_{ac,ion}}{k_BT}\left(\sqrt{1-\gamma^2}+\gamma\arcsin\gamma\right)\right), \tag{6}$$

$$\gamma = \frac{z_{V_o}eaE}{\pi\Delta E_{ac,ion}}, \tag{7}$$

where $\bar{N}$ is the geometric mean between N and N_f, a is the hopping distance, E denotes the electric field, ν_0 is the attempt-to-escape frequency, and $\Delta E_{ac,ion}$ is the activation energy for ion hopping. The factor γ takes into account the effect of the electric field when the hopping barrier is modified. For the sake of a compact implementation of the model, we approximate the ion concentration gradient dN/dx in equation (4) by a finite difference ratio,

$$\frac{dN}{dx} \approx \frac{N_f - N}{L_{gap}}. \tag{8}$$

Finally, the electric field E is defined to be polarity-dependent, since we assume that the band bending occurs entirely on the gap region during the SET transition, and along the whole extension of the oxide during the RESET transition [27].

The discussed model has been validated on several experimental data. Table 1 summarizes the best model parameters achieved after a thorough calibration of the model on experimental data. We resort to a unique set of parameters to deal both with the quasi-static and the dynamic behaviour. The experimental results and the corresponding simulation will be discussed into details in the following sections.

Table 1: Table of the model parameters used in the simulations.

Symbol	Value	Symbol	Value
$\|SR\|$	$0.1\,V/s$	$\Delta E_{ac,ion}$	$0.95\,eV$
R_s	$200\,\Omega$	a	$0.35\,nm$
r_f	$35\,nm$	ν_0	$1\cdot 10^{12}\,s^{-1}$
N_f	$5\cdot 10^{21}\,cm^{-3}$	L_{gap}	$1.4\,nm$
z_{V_0}	$+2$	L_{ox}	$5.5\,nm$
T_0	$300\,K$	k_{th}^{SET}	$0.9\,W/(m\cdot K)$
ΔE_{ac}	$0.06\,eV$	k_{th}^{RESET}	$1.1\,W/(m\cdot K)$
μ_{n0}	$0.6\cdot 10^{-1}\,cm^2/(V\cdot s)$	C_{th}	$1\cdot 10^{16}\,Ws/k$

3 Results and Discussion

3.1 Quasi-static I-V: experimental data and simulation

First, we show that the considered model is able to capture the main features of the quasi-static experimental I-V characteristics (see figure 1(c)). In particular, figure 1(d) shows the simulated I-V curve which correctly reproduces the SET and RESET voltages, as well as the difference between the abrupt SET and the gradual RESET transition. It is worth noting that, at low voltages, the device conduction is symmetric with respect to the origin in both the HRS and LRS. As a consequence, and differently from [16, 27], we do not need to include a metal/oxide interface barrier in the model. The SET transition occurs sharply around -0.55V, due to the triggering of a positive feedback between resistance decrease and Joule heating [7]. During the measurements, the SET transition is limited by the current compliance which prevents possible overshoots [30, 31]. The RESET transition starts from about 0.55V and gradually continues due to the set up of a negative thermal feedback and the attainment of an equilibrium between the drift and the diffusion of oxygen vacancies in the gap [7, 23, 32]. In particular, the diffusion process counteracts the drift, which tends to reduce the oxygen vacancies concentration [33]. Therefore, maximum and minimum concentration values have to be evaluated self-consistently through at least one complete SET and RESET cycle.

3.2 Device operation under pulse programming: binary and analog switching

In the following, we discuss the device operation under pulse programming and the corresponding simulation results (see figure A of the supplementary material for details on the program/read scheme). In particular, we refer to figure 2 that reports the resistance window R_{HRS}/R_{LRS} for various combination of pulse voltages and pulse widths. The panels (a) and (b) in figure 2 show the experimental resistance memory windows R_{HRS}/R_{LRS} for both SET and RESET operations. To obtain a single experimental memory window, corresponding to a single pixel in the figure, the following procedure is followed. For the SET transition, the device is set to the HRS, in the $3\,k\Omega$ range. Then, 300 identical programming pulses at

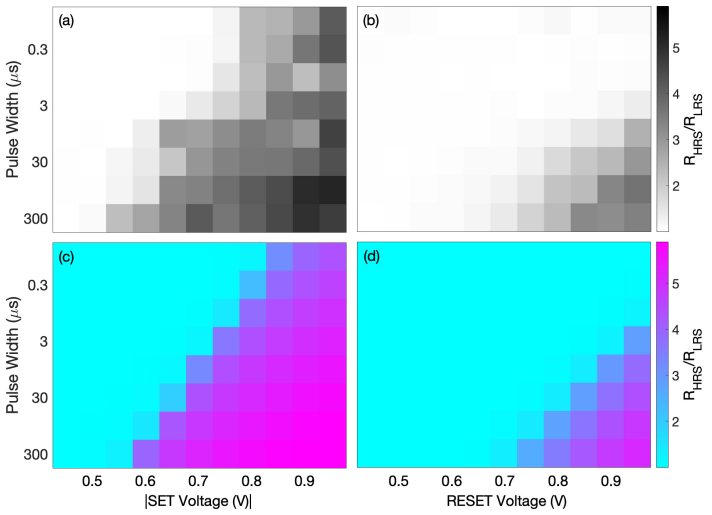

Figure 2: Experimental (a)-(b) and simulated (c)-(d) memory windows for the SET (left) and the RESET (right) transition. For each pixel, we provide the extracted R_{HRS}/R_{LRS} total variation, after applying 300 identical programming pulses at negative, (a)-(c), or positive, (b)-(d), voltage polarity, for each pulse width/amplitude condition, following the procedure described in the text.

negative voltage polarity, spaced out by reading phases at 100 mV, are applied to the device. The final resistance is read after the application of all the pulses. Finally, the R_{HRS}/R_{LRS} total variation is evaluated as the ratio between the initial to the final device resistance, i.e., before and after the application of the 300 programming pulses. The procedure is repeated, starting from the same initial condition, for different combinations of pulse amplitudes (from 0.35 V to 0.95 V in absolute value) and pulse widths (from 0.1 μs to 300 μs), thus obtaining a different memory window for each pulse width/pulse amplitude combination, corresponding to a single pixel of figure 2(a). Figure 2(b) shows the results associated with the RESET case. The adopted procedure is essentially the same as for the SET, starting from positive programming pulses. The R_{HRS}/R_{LRS} total variation for the RESET case is evaluated as the ratio between the final to the initial device resistance, i.e., after and before the application of the 300 programming pulses. The experimental behaviour is accurately reproduced by the model as highlighted in the panels (c) and (d) in figure 2. We notice that, while experimentally we apply trains of separate pulses, the simulations are carried out by resorting to a single voltage pulse with a duration equal to the total experimental time span (i.e., the sum of the width of all the applied pulses).

Figure 2 emphasizes that the maximum resistance variation after 300 identical pulses, i.e., the maximum memory window that we can set, is related to the programming condition (pulse amplitude/pulse width), and that the maximum achievable (R_{HRS}/R_{LRS}) is slightly above 5. In particular, we observe that only if the pulse voltage amplitude and the pulse width are sufficiently large (bottom right corners of the panels in figure 2), a resistance variation close to the maximum resistance window can be obtained. In this

case we can refer to a strong programming condition, where the resistance variation is driven to the maximum value by the first or by the first few pulses. This operation mode is the one usually employed in binary switching for storage applications. Otherwise, for intermediate values of the pulse voltage amplitude and of the pulse width, the resistance variation after 300 pulses equals only a fraction of the maximum resistance window, i.e., a weak programming condition is reached [4]. Under the latter programming condition, each pulse produces a small resistance variation and therefore multiple resistance states are achieved. The devices exhibit a gradual resistance dynamics as a function of the number of applied pulses. This feature representing the key functionality for neuromorphic computing. Finally, for applied voltage and pulse width below a certain threshold, no variation of resistance is measurable (actually, the values for R_{HRS}/R_{LRS} in figure 2 are close to zero).

To summarize, the proposed model captures the device resistance variation for all the investigated programming conditions in a pulse regime, and figure 2 defines the existence of programming windows both for analog and binary switching.

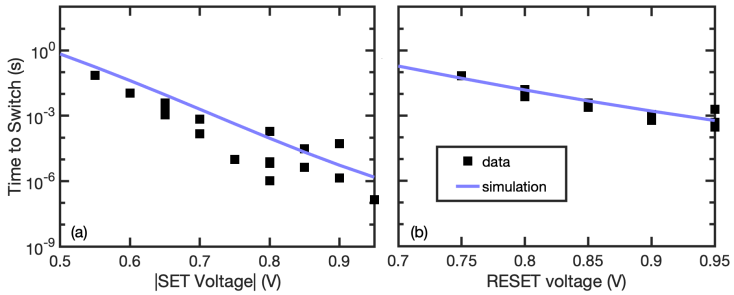

Figure 3: Comparison between experimental (black squares) and simulated (blue line) kinetics for SET (*a*) and RESET (*b*) processes.

3.3 Switching kinetics, dynamics and variability of RRAM devices

A different way of describing the switching over different voltages and times is through the switching kinetics, i.e., the time required by the device to get a certain resistance change, for a definite applied voltage. In the present case, we consider a resistance change from the initial condition equal to a resistance window with $R_{HRS}/R_{LRS} = 2$, namely we halve and double the initial resistance value in the SET and RESET experiments, respectively. Figure 3 confirms the very good agreement between the experimental (symbols) and the simulated (line) switching kinetics, over several orders of magnitude of switching times. The simulations correctly reproduce the experimental slopes, by returning values of about 80 mV/dec and 100 mV/dec for SET and RESET, respectively. In figure 3, each experimental value for the switching time (black squares) corresponds to the total time taken by n pulses (at a fixed amplitude and time) when applied to the device in order to achieve the specified memory window. Notice that trains of n pulses characterized by the same amplitude but with a different pulse width can lead to the same total switching time, depending on the number of pulses. Therefore, the observed variability in the switching time

for a certain voltages depends on the combined effect between the typical switching variability of RRAM devices programmed in a weak pulse regime [4, 34], and the difference in time resolution between series of applied pulses. In particular, the variability turns out to be higher for the SET process than for the RESET one. This finding can be partially ascribed to the time resolution, limited by the pulse width, which affects more the fast SET transition rather than the smooth RESET phase. Moreover, apart from the experimental measurement error, it is well-known that RRAM devices exhibit a significant intrinsic variability, ascribed to random changes in the filament geometry or to the density of oxygen vacancies therein, when filaments are continuously formed and dissolved through SET and RESET operations [17, 35, 36].

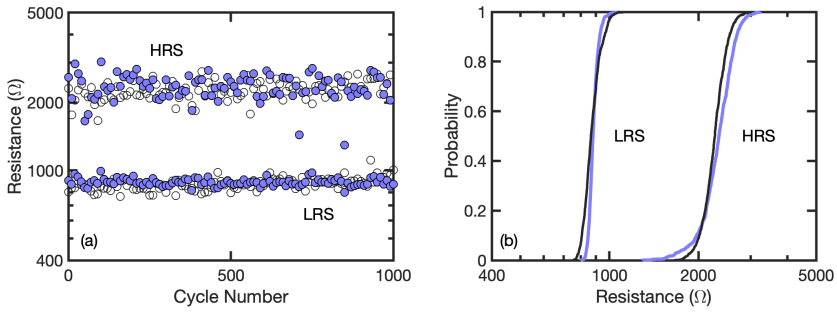

Figure 4: (*a*) Experimental (empty circles) versus simulated (filled circles) data for 1000 cycles of consecutive digital SET and RESET transitions (V_{SET} = 0.9 V, V_{RESET} = -1.2 V, pulse width = 10 μs); (*b*) Experimental (black lines) and simulated (blue lines) cumulative distribution functions for the HRS and the LRS.

The developed compact model allows us to correctly include variability. This has been firstly validated on a binary switching operation, and for the parameters r_f, N_f and L_{gap}. In particular, figure 4(a) shows the variability for parameter L_{gap}, by comparing experimental (empty symbols) with simulated (filled symbols) repetitive binary switching between a high and a low resistance states. Each parameter value is extracted from a normal distribution, with mean equal to the value used in the deterministic simulations (see Table 1) and a relative standard variation of 2.1%. The variation adopted for L_{gap} appropriately describes the measured experimental variability also for the cumulative distributions considered in figure 4(*b*). Indeed, the simulated resistance spread, both in the LRS and in the HRS, well reproduces the experimental values, with a larger deviation from the average value for the HRS than for the LRS, in agreement with other studies [37, 34]. In fact, after the RESET, R_{gap} dominates the series of resistances, thus enhancing the effect of the variability in the L_{gap} parameter (see equation (1)).

Finally, we discuss the RRAM device dynamics when programmed in the analog regime under weak programming conditions. This operation mode represents the most challenging and useful kind of programming for neuromorphic computing. Figure 5 shows representative dynamics curves which are yielded by different sequences of trains of identical pulses at a fixed pulse amplitude (-0.85 V for the SET and 0.9 V for the RESET) but with diverse pulse widths. In particular, for a given pulse amplitude, we first acquire the evolution of

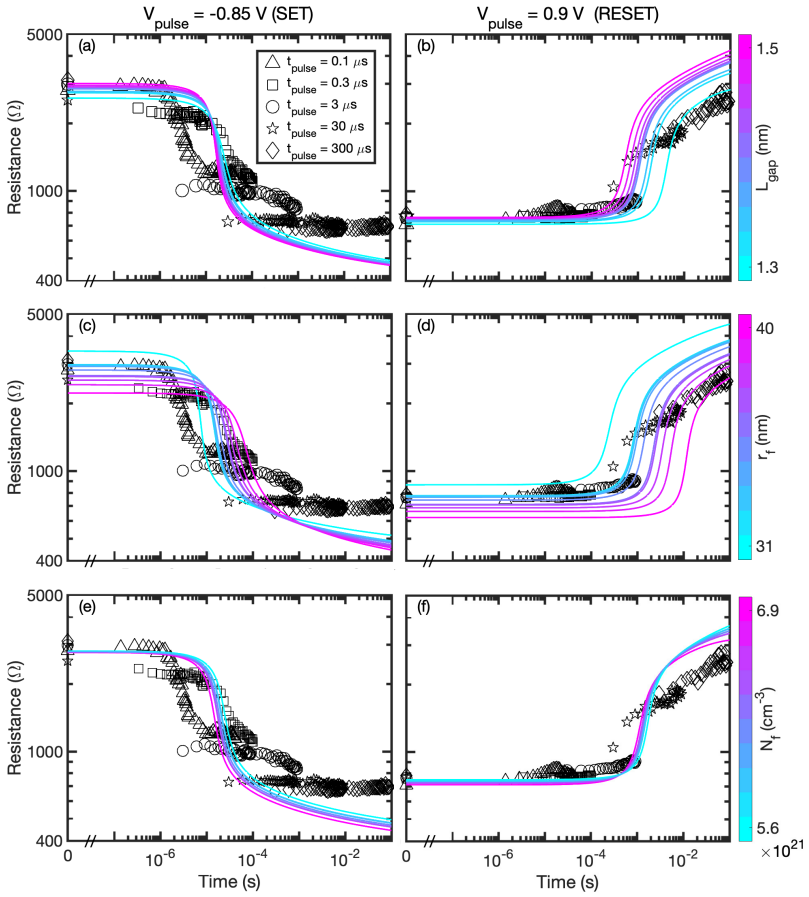

Figure 5: Experimental (black markers) and simulated (coloured lines) dynamics. The simulated dynamics are obtained by introducing variability in three model parameters, namely (a)-(b) L_{gap}, (c)-(d) r_f and (e)-(f) N_f.

the device resistance under the application of 300 consecutive pulses at a fixed pulse width equal to 0.1 us. Then we plot the R values as a function of time, by considering the total time of voltage application equal to the product between the pulse width and the number of pulses, which, e.g., for 1 μs pulse width leads to 300 μs. Then, the device is reset to the initial resistance value by a controlled DC sweep. This procedure is repeated for the same pulse amplitude but different pulse widths equal to 0.3, 3, 30 and 300 μs. In this way we can assess how the resistance changes over a large time scale, but, at the same time, we have enough resolution also to track small variations for shorter pulse width. It is worth mentioning that this procedure is necessary since, for each pulse amplitude/pulse width, the maximum resistance variation that we can achieve is limited as shown in figure 2. Hence, it would not be possible to reproduce the entire resistance dynamics simply by setting a specific pulse width and applying more than 300 pulses. Finally, all these data are plotted in the same panel to achieve the overall R evolution from 100 ns to 100 ms. Such

estimation of the overall dynamics is correct under the assumption that the voltage/overall time pair solely determines the device behavior, meaning that the pulse rate is not affecting the device response, as previously demonstrated in [38]. Notice that each sequence of experimental data, related to a different pulse width, is represented in the figure by a different marker. Concerning the experimental data variability observed for the resistance dynamics curves (i.e., the different resistance values measured after a given time), we can distinguish two different origins. First of all, dynamics variability is typical of a weak programming regime and is related to the intrinsic stochasticity of the filament formation/rupture processes, as previously discussed. In addition, dynamics variability can be caused by the lack of time resolution when applying pulses with a large pulse width. In fact, in such a case, the switching transition could occur before the ending of the pulse, since we read the device resistance after the pulse takes place. This fact can lead to an over-estimation of the time. The variability is higher for the SET configuration with respect to the RESET phase. Figure 5 superimposes the simulated curves for the device dynamics (coloured lines) to the experimental data (highlighted by different markers). It is evident that the proposed model overall well reproduces the functional shape of the resistance evolution, as a function of time. In both SET and RESET transitions, the resistance curve changes slowly at the beginning, until a positive feedback between resistance change and Joule heating sets up. At this point, the resistance varies quickly, changes concavity and finally saturates to a LRS or to a HRS. The saturation of the resistance value during the SET process occurs when the gap resistance becomes negligible with respect to those of filament and series contributions. The RESET saturation occurs due to a negative thermal feedback and to the reached equilibrium between the ion drift and the ion diffusion in the gap, as for the quasi-static case. Even though the simulated resistance dynamics curves succeed in replicating the overall experimental behavior of the resistive switching over several orders of magnitude in time, some discrepancies with experimental data can be noticed, especially for the saturation part of the SET. Indeed, while the experimental dynamics curves saturate to the LRS, the simulated resistances continue to decrease in value, even if with a considerably reduced trend with respect to the transition slope. The sharp experimental saturation at the end of the SET can be ascribed to the complementary resistive switching (CRS), exhibited by the investigated device [32]. The CRS is caused by the presence of a TiO_xN_y interlayer formed in correspondence of the TiN/HfO_2 interface. The TiO_xN_y interlayer acts as a sink/source of oxygen vacancies during the SET/RESET, inducing an opposite transition that supports the resistance saturation. The phenomenon is particularly evident for the SET transition in the pulse regime, as the combination of short pulses with the CRS behavior prevents the device breakdown without current limitation [32]. The inclusion of the effect of the TiO_xN_y interlayer in the model, as a variable resistor dependent on the oxygen vacancies density of the interlayer, is a possible solution to improve the overall quality of the fitting of the experimental dynamics curves, in particular for the saturation part of the LRS. Beside the general shape of the resistance evolution curve, we studied the influence of intrinsic variability of the switching process on the device dynamics. To this aim, in figure 5 we superimpose various simulated dynamic curves to the data, as a function of some of the geometric parameters characterizing the filament, namely the gap length L_{gap} (figure 5(a)-(b)), the filament radius r_f (figure 5(c)-(d)), and the filament ionic density

N_f (figure 5(e)-(f)). We can draw the following conclusions. Both the SET and RESET transitions occur earlier for large values of L_{gap}, even if the advance is more evident for the RESET. Concerning the speed of the transition, the SET transitions are faster in the case of large values for L_{gap}, while the RESET transitions occur with about the same resistance slopes. Further details and explanations on the role of the gap length in the dynamics can be found in the second section of the supplementary material (see figures C - J). Similarly to the case of digital transitions in figure 4, the difference between the curves becomes more evident as the device resistance approaches the associated HRS. For large values of r_f, the dynamics curves are translated vertically, towards lower resistance values, and the transitions are delayed. Finally, for high concentrations N_f of ions, the transitions occur early, and the curves saturate to lower resistance values.

4 Conclusions

In conclusion, modelling and simulation of the operative analog behaviour of HfO_2-based devices have been presented. The discussed model succeeds both in replicating the dynamical behaviour (analog and digital) and in highlighting all the main characteristics of the device (abrupt SET and gradual RESET in quasi-static regimes, the memory window in weak and strong pulse regimes, highly nonlinear kinetics with the correct slope over several orders of magnitude in switching times). The stochasticity involved in the filament formation and dissolution has been also considered in the model, due to its relevant effect in dynamic conditions.

Supplementary material

Further details about the program/read scheme, the influence of the variability of the gap length on the resistance dynamics, and a discussion about the SET saturation can be found in the supplementary material.

Data availability

The experimental data and the MATLAB scripts developed for the numerical simulations are available from the authors upon reasonable request.

References

[1] Rainer Waser, Regina Dittmann, Georgi Staikov, and Kristof Szot. Redox-based resistive switching memories–nanoionic mechanisms, prospects, and challenges. *Advanced materials*, 21(25-26):2632–2663, 2009.

[2] Mohammed A Zidan, John Paul Strachan, and Wei D Lu. The future of electronics based on memristive systems. *Nature electronics*, 1(1):22–29, 2018.

[3] Sabina Spiga, Abu Sebastian, Damien Querlioz, and Bipin Rajendran. Role of resistive memory devices in brain-inspired computing. In Sabina Spiga, Abu Sebastian, Damien Querlioz, and Bipin Rajendran, editors, *Memristive Devices for Brain-Inspired Computing*, Woodhead Publishing Series in Electronic and Optical Materials, chapter 1, pages 3–16. Woodhead Publishing, 2020.

[4] Stefano Brivio and Stephan Menzel. Resistive switching memories. In Sabina Spiga, Abu Sebastian, Damien Querlioz, and Bipin Rajendran, editors, *Memristive Devices for Brain-Inspired Computing*, chapter 2, pages 17–61. Woodhead Publishing, Oxford, 2020.

[5] S Brivio, E Covi, A Serb, T Prodromakis, M Fanciulli, and S Spiga. Gradual set dynamics in HfO_2-based memristor driven by sub-threshold voltage pulses. In *2015 International Conference on Memristive Systems (MEMRISYS)*, pages 1–2. IEEE, 2015.

[6] Erika Covi, Stefano Brivio, Alexander Serb, Themis Prodromakis, Marco Fanciulli, and Sabina Spiga. Analog memristive synapse in spiking networks implementing unsupervised learning. *Frontiers in neuroscience*, 10:482, 2016.

[7] Felix Cüppers, S Menzel, C Bengel, A Hardtdegen, M Von Witzleben, U Böttger, R Waser, and S Hoffmann-Eifert. Exploiting the switching dynamics of HfO_2-based ReRAM devices for reliable analog memristive behavior. *APL materials*, 7(9):091105, 2019.

[8] Jun-Woo Jang, Sangsu Park, Geoffrey W Burr, Hyunsang Hwang, and Yoon-Ha Jeong. Optimization of conductance change in Pr_{1-x} Ca_x MnO_3-based synaptic devices for neuromorphic systems. *IEEE Electron Device Letters*, 36(5):457–459, 2015.

[9] Jiyong Woo, Kibong Moon, Jeonghwan Song, Sangheon Lee, Myounghun Kwak, Jaesung Park, and Hyunsang Hwang. Improved synaptic behavior under identical pulses using AlO_x/HfO_2 bilayer RRAM array for neuromorphic systems. *IEEE Electron Device Letters*, 37(8):994–997, 2016.

[10] Jaesung Park, Myunghoon Kwak, Kibong Moon, Jiyong Woo, Dongwook Lee, and Hyunsang Hwang. TiO_x-based RRAM synapse with 64-levels of conductance and symmetric conductance change by adopting a hybrid pulse scheme for neuromorphic computing. *IEEE Electron Device Letters*, 37(12):1559–1562, 2016.

[11] Dennis Valbjørn Christensen, Regina Dittmann, Bernabe Linares-Barranco, Abu Sebastian, Manuel Le Gallo, Andrea Redaelli, Stefan Slesazeck, Thomas Mikolajick, Sabina Spiga, Stephan Menzel, and al et. 2022 roadmap on neuromorphic computing and engineering. *Neuromorphic Computing and Engineering*, 2022.

[12] Jiyong Woo, Andrea Padovani, Kibong Moon, Myounghun Kwak, Luca Larcher, and Hyunsang Hwang. Linking conductive filament properties and evolution to synaptic behavior of RRAM devices for neuromorphic applications. *IEEE Electron Device Letters*, 38(9):1220–1223, 2017.

[13] S Brivio, G Tallarida, E Cianci, and S Spiga. Formation and disruption of conductive filaments in a HfO_2/TiN structure. *Nanotechnology*, 25(38):385705, 2014.

[14] Andrea Padovani, Luca Larcher, Onofrio Pirrotta, Luca Vandelli, and Gennadi Bersuker. Microscopic modeling of HfO_x RRAM operations: From forming to switching. *IEEE Transactions on electron devices*, 62(6):1998–2006, 2015.

[15] Ying Zhang, Ge-Qi Mao, Xiaolong Zhao, Yu Li, Meiyun Zhang, Zuheng Wu, Wei Wu, Huajun Sun, Yizhong Guo, Lihua Wang, et al. Evolution of the conductive filament system in HfO_2-based memristors observed by direct atomic-scale imaging. *Nature communications*, 12(1):1–10, 2021.

[16] Camilla La Torre, Alexander F Zurhelle, Thomas Breuer, Rainer Waser, and Stephan Menzel. Compact modeling of complementary switching in oxide-based ReRAM devices. *IEEE transactions on electron devices*, 66(3):1268–1275, 2019.

[17] Christopher Bengel, Anne Siemon, Felix Cüppers, Susanne Hoffmann-Eifert, Alexander Hardtdegen, Moritz von Witzleben, Lena Hellmich, Rainer Waser, and Stephan Menzel. Variability-aware modeling of filamentary oxide-based bipolar resistive switching cells using SPICE level compact models. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 67(12):4618–4630, 2020.

[18] Francesco Maria Puglisi, Luca Larcher, Andrea Padovani, and Paolo Pavan. Bipolar resistive RAM based on HfO_2: Physics, compact modeling, and variability control. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 6(2):171–184, 2016.

[19] Linlin Cai, Wangyong Chen, Yudi Zhao, Xiaoyan Liu, Jinfeng Kang, Xing Zhang, and Peng Huang. A physics-based analytic model of analog switching resistive random access memory. *IEEE Electron Device Letters*, 41(2):236–239, 2019.

[20] Ye Zhuo, Rivu Midya, Wenhao Song, Zhongrui Wang, Shiva Asapu, Mingyi Rao, Peng Lin, Hao Jiang, Qiangfei Xia, R Stanley Williams, et al. A dynamical compact model of diffusive and drift memristors for neuromorphic computing. *Advanced Electronic Materials*, page 2100696, 2021.

[21] Jacopo Frascaroli, Stefano Brivio, Erika Covi, and Sabina Spiga. Evidence of soft bound behaviour in analogue memristive devices for neuromorphic computing. *Scientific reports*, 8(1):1–12, 2018.

[22] Sungho Kim, Chao Du, Patrick Sheridan, Wen Ma, ShinHyun Choi, and Wei D Lu. Experimental demonstration of a second-order memristor and its ability to biorealistically implement synaptic plasticity. *Nano letters*, 15(3):2203–2211, 2015.

[23] Astrid Marchewka, Bernd Roesgen, Katharina Skaja, Hongchu Du, Chun-Lin Jia, Joachim Mayer, Vikas Rana, Rainer Waser, and Stephan Menzel. Nanoionic resistive switching memories: On the physical nature of the dynamic reset process. *Advanced Electronic Materials*, 2(1):1500233, 2016.

[24] Stefano Larentis, Federico Nardi, Simone Balatti, David C Gilmer, and Daniele Ielmini. Resistive switching by voltage-driven ion migration in bipolar RRAM—part II: Modeling. *IEEE Transactions on Electron Devices*, 59(9):2468–2475, 2012.

[25] Stefano Brivio, Jacopo Frascaroli, and Sabina Spiga. Role of Al doping in the filament disruption in HfO_2 resistance switches. *Nanotechnology*, 28(39):395202, 2017.

[26] M Saludes-Tapia, S Poblador, MB Gonzalez, F Campabadal, J Suñe, and E Miranda. Tunability properties and compact modeling of HfO_2-based complementary resistive switches using a three-terminal subcircuit. *IEEE Transactions on Electron Devices*, 68(12):5981–5988, 2021.

[27] Camilla La Torre, Alexander F Zurhelle, and Stephan Menzel. Compact modelling of resistive switching devices based on the valence change mechanism. In *2019 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pages 1–4. IEEE, 2019.

[28] Simon M Sze, Yiming Li, and Kwok K Ng. *Physics of semiconductor devices*. John wiley & sons, 2021.

[29] Suhas Kumar, John Paul Strachan, and R Stanley Williams. Chaotic dynamics in nanoscale NbO_2 Mott memristors for analogue computing. *Nature*, 548(7667):318–321, 2017.

[30] Kentaro 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 NiO_x brought about by reducing a parasitic capacitance. *Applied Physics Letters*, 93(3):033506, 2008.

[31] S Brivio, J Frascaroli, and S Spiga. Role of metal-oxide interfaces in the multiple resistance switching regimes of Pt/HfO_2/TiN devices. *Applied Physics Letters*, 107(2):023504, 2015.

[32] S Brivio, E Covi, A Serb, T Prodromakis, M Fanciulli, and S Spiga. Experimental study of gradual/abrupt dynamics of HfO_2-based memristive devices. *Applied Physics Letters*, 109(13):133504, 2016.

[33] Stefano Brivio, Jacopo Frascaroli, Erika Covi, and Sabina Spiga. Stimulated ionic telegraph noise in filamentary memristive devices. *Scientific reports*, 9(1):1–9, 2019.

[34] Stefano Ambrogio, Simone Balatti, A Cubeta, A Calderoni, N Ramaswamy, and Daniele Ielmini. Understanding switching variability and random telegraph noise in resistive RAM. In *2013 IEEE International Electron Devices Meeting*, pages 31–5. IEEE, 2013.

[35] Christoph Baeumer, Richard Valenta, Christoph Schmitz, Andrea Locatelli, Tevfik Onur Mentes, Steven P Rogers, Alessandro Sala, Nicolas Raab, Slavomir Nemsak, Moonsub Shim, et al. Subfilamentary networks cause cycle-to-cycle variability in memristive devices. *ACS nano*, 11(7):6921–6929, 2017.

[36] Simone Balatti, Stefano Ambrogio, David C Gilmer, and Daniele Ielmini. Set variability and failure induced by complementary switching in bipolar RRAM. *IEEE electron device letters*, 34(7):861–863, 2013.

[37] Daniele Garbin, Elisa Vianello, Quentin Rafhay, Mourad Azzaz, Philippe Candelier, Barbara DeSalvo, Gerard Ghibaudo, and Luca Perniola. Resistive memory variability: A simplified trap-assisted tunneling model. *Solid-State Electronics*, 115:126–132, 2016.

[38] Erika Covi, Stefano Brivio, Jacopo Frascaroli, Marco Fanciulli, and Sabina Spiga. Analog HfO_2-RRAM switches for neural networks. *ECS Transactions*, 75(32):85, 2017.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

35/2022 Perotto, S.; Bellini, G.; Ballarin, F.; Calò, K.; Mazzi, V.; Morbiducci, U.
Isogeometric hierarchical model reduction for advection-diffusion process simulation in microchannels

34/2022 Antonietti, P.F.; Vacca, G.; Verani, M.
Virtual Element method for the Navier-Stokes equation coupled with the heat equation

33/2022 Africa, P.C.; Salvador, M.; Gervasio, P.; Dede', L.; Quarteroni, A.
A matrix-free high-order solver for the numerical solution of cardiac electrophysiology

32/2022 Barbi, C.; Menafoglio, A; Secchi, P.
An object-oriented approach to the analysis of spatial complex data over stream-network domains

30/2022 Bonetti S.; Botti M.; Antonietti P.F.
Discontinuous Galerkin approximation of the fully-coupled thermo-poroelastic problem

31/2022 Bortolotti, T; Peli, R.; Lanzano, G; Sgobba, S.; Menafoglio, A
Weighted functional data analysis for the calibration of ground motion models in Italy

29/2022 Fumagalli, I.; Polidori, R.; Renzi, F.; Fusini, L.; Quarteroni, A.; Pontone, G.; Vergara, C.
Fluid-structure interaction analysis of transcatheter aortic valve implantation

28/2022 Ciarletta, P.; Pozzi, G.; Riccobelli, D.
The Föppl–von Kármán equations of elastic plates with initial stress

25/2022 Cavinato, L; Gozzi, N.; Sollini, M; Kirienko, M; Carlo-Stella, C; Rusconi, C; Chiti, A; Ieva, F.
Perspective transfer model building via imaging-based rules extraction from retrospective cancer subtyping in Hodgkin Lymphoma

24/2022 Cappozzo, A.; McCrory, C.; Robinson, O.; Freni Sterrantino, A.; Sacerdote, C.; Krogh, V.; Pan
A blood DNA methylation biomarker for predicting short-term risk of cardiovascular events