

A Consistent Model for Short-Term Instability and Long-Term Retention in Filamentary Oxide-Based Memristive Devices

Nils Kopperberg,* Stefan Wiefels, Sergej Liberda, Rainer Waser, and Stephan Menzel



Cite This: *ACS Appl. Mater. Interfaces* 2021, 13, 58066–58075



Read Online

ACCESS |



Metrics & More



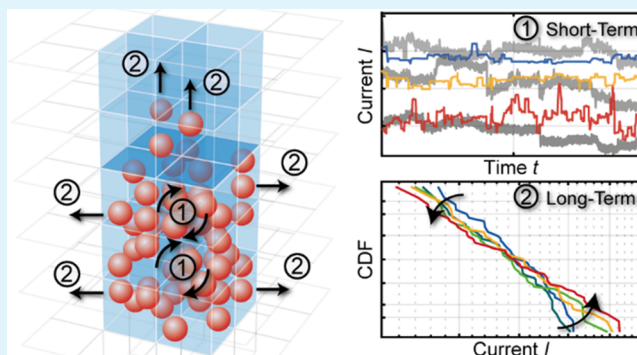
Article Recommendations



Supporting Information

ABSTRACT: Major challenges concerning the reliability of resistive switching random access memories based on the valence change mechanism (VCM) are short-term instability and long-term retention failure of the programmed resistance state, particularly in the high resistive state. On the one hand, read noise limits the reliability of VCMs via comparatively small current jumps especially when looking at the statistics of millions of cells that are needed for industrial applications. Additionally, shaping algorithms aiming for an enlargement of the read window are observed to have no lasting effect. On the other hand, long-term retention failures limiting the lifetime of the programmed resistance states need to be overcome. The physical origin of these phenomena is still under debate and needs to be understood much better. In this work, we present a three-dimensional kinetic Monte Carlo simulation model where we implemented diffusion-limiting domains to the oxide layer of the VCM cell. We demonstrate that our model can explain both instability and retention failure consistently by the same physical processes. Further, we find that the random diffusion of oxygen vacancies plays an important role regarding the reliability of VCMs and can explain instability phenomena as the shaping failure as well as the long-term retention failure in our model. Additionally, the results of the simulations are compared with experimental data of read noise and retention investigations on ZrO_2 -based VCM devices.

KEYWORDS: ReRAM, valence change memory, HfO_2 , oxygen vacancy, simulation, kinetic Monte Carlo, instability, retention



INTRODUCTION

Due to the sharply rising demand on portable electronic devices such as smartphones and being concomitant with that, the need for research on highly scaled non-volatile memories strongly increases.^{1–3} Due to their promising operation features like scalability, non-volatility, reliability, complementary metal–oxide–semiconductor (CMOS) compatibility, and short switching times,⁴ resistive switching random access memories (ReRAMs) are considered to succeed flash technology in the future.⁵ Additionally, ReRAMs are considered promising candidates for several neuromorphic and machine learning applications.^{6,7}

One extensively studied type of ReRAM is bipolar switching valence change-based memories (VCMs) that typically consist of a transition metal oxide layer like HfO_2 , ZrO_2 , or Ta_2O_5 sandwiched between two metal electrodes.^{8–10} By applying an external voltage, oxygen can be extracted from the initially insulating oxide layer at the chemically active electrode. The so-called oxygen vacancies (V_O) left behind in the oxide build a conductive filament significantly changing the resistance of the device. Depending on the polarity of the externally applied voltage, this filament can be ruptured and rebuilt consecutively switching the VCM device between a high resistive state

(HRS) and a low resistive state (LRS).¹¹ These resistance states are non-volatile, can be read out non-destructively, and are used to store the information as logical “0” and “1”.

Regardless of the numerous great properties of ReRAMs mentioned above, there are issues concerning the stability of the programmed states warding them from commercial breakthrough so far.^{12–14} One challenge is short-term instability, also called read noise, limiting the read window, which separates the HRS and LRS. Here, program-verify or shaping algorithms were proposed to overcome these limitations¹⁵ but have been reported to not have a lasting effect.¹³ Instead, the shaped resistance (or read current) distributions always revert back toward their intrinsic statistics. These intrinsic distributions are observed to be very stable on short time scales.¹⁶ However, on long time scales (>1 month) as required for typical memory applications, significant

Received: August 2, 2021

Accepted: October 26, 2021

Published: November 22, 2021



resistance changes may occur that are referred to as retention failure.¹⁷ Both failure mechanisms are so far not fully understood, and to our knowledge, a physical simulation model has not yet been demonstrated to consistently cover both. Puglisi *et al.* explained a random telegraph noise via charge trapping at slow defects but did not discuss retention effects.¹⁸ On the other hand, Aldana *et al.* stated that unbalanced V_{O} generation and recombination are responsible for the LRS retention failure¹⁹ but the read current instability was not discussed. In the work of Huang *et al.*,²⁰ the state instability was connected to Brownian-like hopping in a kinetic Monte Carlo (KMC) model. Furthermore, they used a second, clearly different analytical simulation model to explain retention phenomena mainly by V_{O} diffusion and recombination.

In this work, we will introduce a physical explanation for the processes taking place in VCM devices that are responsible for the instability and retention failures mentioned before, whereas the switching kinetics have been demonstrated in a previous work.²¹ Therefore, we present a KMC-based simulation model based on the work of Abbaspour *et al.*²¹ that we extended by the implementation of domains limiting the diffusion of V_{O} in the oxide layer. This extension is motivated by the work of Schie *et al.*²² Using molecular dynamics simulations, they found different diffusion regimes for V_{O} in HfO_2 , suggesting that V_{O} may diffuse easily over short distances within certain domains. Within sub-diffusive regions between these domains, the diffusion is expected to require a higher activation energy.²² These different diffusion regimes are a possible explanation for the high discrepancy in the time of occurrence of the different failure mechanisms. With the implementation of these diffusion-limiting domains, we can accurately show and explain several phenomena concerning the above-mentioned instability and retention of the HRS of bipolar VCM devices. Both failure mechanisms can be explained by the same physical origin based on the random configuration and movement of V_{O} . In addition, we present experimental results obtained for VCM ReRAM devices based on ZrO_2 , which is considered nearly identical to HfO_2 with respect to its physicochemical properties.²³ Using these exemplary devices, we demonstrate typical short-term instability and long-term retention failures and compare our simulation model to the experimental findings.

SIMULATION MODEL

The simulation model used in this work is based on the work of Abbaspour *et al.*²¹ and represents a typical VCM ReRAM cell. An oxide layer, e.g., HfO_2 or ZrO_2 , sandwiched between two metal electrodes, is modeled by a $6 \times 6 \times 6 \text{ nm}^3$ lattice structure with a spacing of 0.5 nm between the individual lattice points. Oxygen vacancies (V_{O}) that are of major importance for the current transport in VCM memory devices⁸ can be placed at the lattice points of the oxide structure. There, they locally reduce the lattice cations and induce defect states that act as electron traps, which are responsible for electron transport through the initially non-conducting oxide.^{24,25}

At a given V_{O} configuration in the oxide of the VCM device, an external voltage can be applied to the electronically active top electrode (Pt in our work), while the ohmic bottom electrode (Ti in our work) is ground. Solving the three-dimensional Poisson equation leads to the potential landscape in the whole device. Here, the charge of the V_{O} is considered twice positive reduced by the electron occupation probability

that is between zero and one. Additionally, the heat-flow equation is solved, leading to the local temperature distribution in the cell. Afterward, the current through the device is calculated using a trap-assisted tunneling (TAT) current solver since the HRS of VCM devices with comparably low V_{O} concentrations can be well described by the multiphonon tunneling of electrons via electron traps.²⁶ Here, two main processes are considered, which are the electron tunneling between the electrode and trap and the electron hopping between two traps.²⁷

The electron tunneling rate between the electrode and trap is calculated by

$$R_{\text{T}} = R_0 P_{\text{T}} F \quad (1)$$

where R_0 is the electrode-trap coupling factor. The tunneling probability T_{R} is given by the Wentzel–Kramers–Brillouin (WKB) approximation

$$T_{\text{R}} = e^{-2 \int_{x_1}^{x_2} 1/\hbar \sqrt{2m^*(e\Phi_{\text{B}} - E_{\text{e}})} dx} \quad (2)$$

with x_1 and x_2 being the electron positions in the traps and the electrode, the effective mass for tunneling in HfO_2 $m^* = 0.1 m_0$, the tunneling barrier Φ_{B} , and the electron energy E_{e} . The Fermi factor F gives the number of occupied states in the electrode in the case of electrode-to-trap tunneling and correspondingly the number of unoccupied states in the electrode in the case of trap-to-electrode tunneling and is calculated via

$$F_{\text{E} \rightarrow \text{T}} = \int_{E_{\text{t}}^{\text{e}}}^{+\infty} F(E) \rho(E) dE \quad (3)$$

$$F_{\text{T} \rightarrow \text{E}} = \int_{-\infty}^{E_{\text{t}}^{\text{f}}} (1 - F(E)) \rho(E) dE \quad (4)$$

with the Fermi–Dirac density function $F(E)$, E_{t}^{e} and E_{t}^{f} denoting the energies of empty or filled traps, respectively, and $\rho(E)$ being the density of states of electrons in the metal electrode. For the electron hopping between traps, the Abraham–Miller formula is consulted so that the hopping rate from one trap at position i to another trap at position j is calculated with the following equation

$$h_{ij} = v_0^{\text{e}} \exp\left(-\frac{d_{ij}}{a_0}\right) \left\{ \exp\left(-\frac{e(V_i - V_j)}{k_{\text{B}} T}\right), V_i \geq V_j \right. \\ \left. 1, V_i < V_j \right\} \quad (5)$$

where v_0^{e} denotes the electron vibration frequency, d_{ij} is the distance between the two traps, a_0 is the attenuation length of the localized wave function of the defects, and V_i and V_j are the potentials at the trap sites.

The modeling of the system dynamics is based on a kinetic Monte Carlo (KMC) algorithm, where the probabilities of all possible processes in the device are calculated. Then, randomly but weighted by their probabilities, one of these processes is chosen and fulfilled. In our simulation model, three processes concerning V_{O} are possible. The first and central process in this work is the diffusion of V_{O} from one lattice site to a neighboring one. Due to redox reactions, oxygen exchange can take place at the ohmic (bottom) electrode-oxide interface, leading to the generation or recombination of a V_{O} at the interface. In contrast to other simulation models,^{27–29} no generation of V_{O} in the bulk due to anti-Frenkel pair formation

is possible as they are found to be not stable but recombine nearly immediately.^{30,31} The transition rates are calculated via

$$R_{D/G/R} = \nu_0 \exp\left(-\frac{E_{D/G/R} - e\Delta\Phi}{k_B T}\right) \quad (6)$$

with the characteristic vibration frequency ν_0 , the energy barriers for the respective process $E_{D/G/R}$, and the potential difference between the two states $\Delta\Phi$, where the local potential, the self-potential, and the image potential due to the Coulomb interaction are considered. After choosing one process, the time of the simulation is updated using

$$\Delta t = -\frac{\ln(r_1)}{R_{\text{sum}}} \quad (7)$$

with a random number r_1 between 0 and 1 and the sum of all transition rates R_{sum} . All parameters used in our simulation model that are relevant for this work are given in Table 1.

Table 1. Relevant Parameters Used in the KMC Simulation Model

R_0	ν_0^e	a_0	ν_0	E_D
2×10^{16} Hz	2×10^{12} Hz	0.33 nm	1×10^{12} Hz	0.7 eV

Following the results of Schie *et al.*²² who found different diffusion regimes via molecular dynamics simulations for V_{O} in HfO_2 , we introduced domains limiting the V_{O} diffusion in our model aiming for a simplified approach to model the different energy barriers for the V_{O} diffusion. This is consistent with their findings that a simple random distribution of energy barriers is not sufficient to describe the diffusion behavior. Therefore, the oxide layer of the VCM cell is divided into several cubic boxes, as exemplarily sketched in Figure 1. Inside these boxes, the V_{O} can move as described before, whereas for the V_{O} diffusion from box to box, the diffusion energy barrier E_D is increased to 1.2 eV. For simplification, all boxes have the same diffusion barriers and the same size that can be varied for different simulations.

In general, all simulations in this work are based on a VCM cell programmed to the HRS. This initial state is modeled by a fixed number of V_{O} being placed randomly in a predefined filament region with a relatively large gap between the filament and active electrode. The filament boundaries and the number of V_{O} are the same for all cells of a simulation frame so that the initial variations from cell to cell only originate from the random placement of the V_{O} in the filament region. The initial filament always contains an integer number of cubic boxes in all three dimensions. Additionally, individual V_{O} can also be randomly placed in the layer of boxes above the filament with respect to the non-abrupt transition from the filament to the gap after the RESET process. It should be mentioned that our initial state represents the HRS at a sufficiently large time after the programming so that possible relaxation effects observed in a previous experimental work¹⁶ can be neglected. A typical HRS distribution can be seen in Figure 1.

EXPERIMENTAL SECTION

The experimental results in this work were obtained for VCM cells with a lateral size of $7 \mu\text{m} \times 7 \mu\text{m}$ and a stack of 30 nm Pt/5 nm ZrO_2 /20 nm Ta/30 nm Pt, arranged in a 32×1 cell line array structure. Onto the Pt bottom electrode, covering the whole substrate, 30 nm SiO_2 was deposited via electron beam evaporation. Using UV

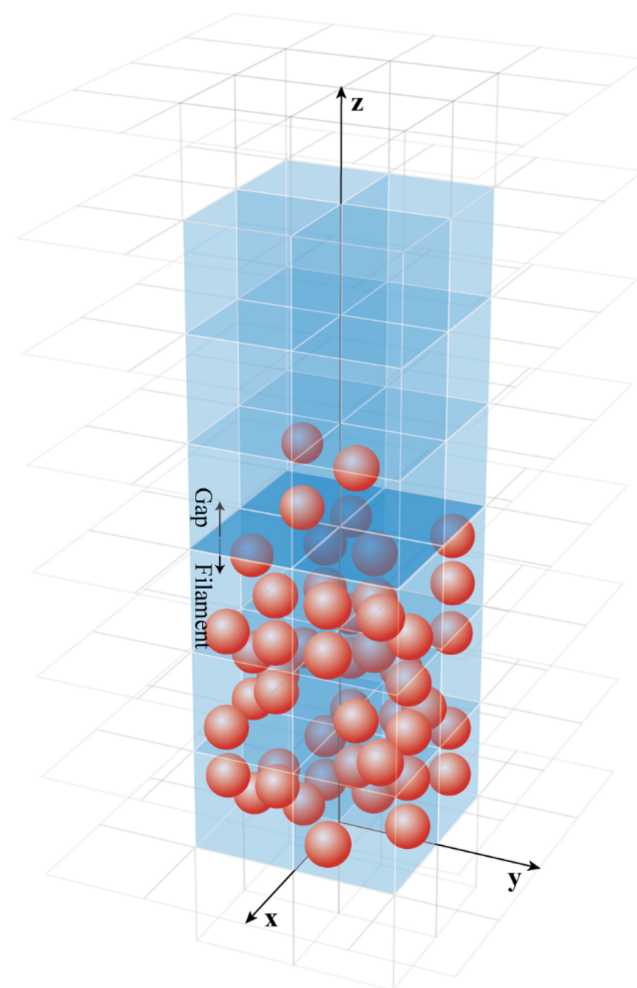


Figure 1. Sketch of the cubic boxes limiting the V_{O} (red balls) diffusion in the oxide layer of the VCM cell. Exemplary shown is a typical defect distribution in the HRS, where the boxes of the filament and gap region are highlighted in blue and divided by the dark blue interface. The ohmic bottom electrode lays at $z = 0$ in the xy plane. For $(1 \text{ nm})^3$ boxes, the active top electrode could be found on top of the blue highlighted boxes.

lithography, the line structure was transferred to the SiO_2 layer. Subsequently, $7 \mu\text{m}$ broad lines were etched down to the Pt layer. It may be noted that the SiO_2 layer does not participate in resistive switching but is used to separate the line arrays. After removal of the photoresist, 5 nm ZrO_2 was deposited via reactive RF sputtering (80 W) and another UV lithography step was applied to structure the top electrode. For the latter, 20 nm Ta was deposited via RF sputtering (12.5 W) and in situ covered by 30 nm Pt (80 W) to prevent oxidation. Further details about the device structure and test setup may be found in ref 32.

The experiments were conducted using a custom array tester by aixACCT Systems, which was connected to the array structure via a wedge card with 32 probes. The tester provides an arbitrary waveform generator, a virtual ground to record the resulting current, and a switchbox that maps the applied signal to the cells of the line array. Using the switchbox, different resistances can be put in series with the VCM cell to limit the switching current. During the initial forming operation, which is required to build the conducting filament, a series resistance of $R_s = 10 \text{ k}\Omega$ was used. For SET operations, R_s was reduced to $1 \text{ k}\Omega$ and RESET operations were performed without series resistance. The Pt bottom electrode (active electrode) was connected to ground, and all voltages were applied to the Ta top electrode (ohmic electrode). However, throughout this paper, all voltages are given with reference to 0 V at the Ta top electrode.

To program the devices, read-verify algorithms were used. For the forming operation, a voltage ramp with a length of $t_{\text{RAMP}} = 20$ ms toward a maximum voltage V_{max} was applied. After each attempt, the resulting cell resistance was determined by a read pulse (50 μs at -0.2 V), and in the case of failed electroforming, the operation was repeated with increased V_{max} . Here, V_{max} ranged from -2.0 to -4.0 V. The SET operation was performed analogously with $t_{\text{ramp}} = 5$ ms and V_{max} ranging from -0.6 to -1.4 V. During RESET, the same algorithm was used but with a rectangular pulse (rise/fall time 5 μs) with a length of $t_{\text{pulse}} = 3$ ms and V_{max} ranging from 1.3 to 2.0 V. It may be noted that the typical voltages for a successful operation were -3.2 V for forming, -1 V for SET, and 1.8 V for RESET.

RESULTS AND DISCUSSION

In the first part of the results, we focus on instability effects in VCM cells programmed to the HRS, which is mainly observed at read operations at comparably low voltages and room temperature. In all simulations under these conditions, only the diffusion of V_{O} can be observed. Due to the small read voltage and the low temperature, no V_{O} generation or recombination at the ohmic electrode happens at the time scales of the experiment. For the same reason, the diffusion of the V_{O} only takes place inside the introduced cubic boxes and no box-to-box diffusion takes place.

Typically, VCMs programmed to the HRS show an intrinsic log-normal behavior of their current distribution.^{12,13} As proposed in previous works, random configurations and the movement of V_{O} seem to be the origin of this statistical behavior.^{16,33} Therefore, the model was extended, and the so far static investigations are enlarged to dynamic simulations, allowing us to observe the time evolution of single cells as well as the statistics of multiple cells.

In the first step, the statistical current distributions resulting from the new simulation model are compared with the experimental data. A box size of $1 \times 1 \times 1 \text{ nm}^3$ has been chosen here and in all following simulations. Fifty V_{O} are placed randomly inside a filament volume of $2 \times 2 \times 3 \text{ nm}^3$ as well as five additional V_{O} in the layer of boxes above. In all simulations, a read voltage of 0.35 V is applied. The corresponding current distribution shows the log-normal behavior that is typical for VCM cells in the HRS³⁴ and can be observed in the dark blue graph (Sim I t_1) of Figure 2. As in experiments,¹⁶ this simulated distribution is in general very stable over time under the mentioned conditions of room temperature and low read voltages. This stability is exemplarily shown in Figure 2 by the nearly similar current distributions at $t_1 = 0$ s at the beginning and $t_2 = 1$ s at the end of a read operation. Therefore, it does not make a difference whether a continuous read voltage is applied or single-read pulses with a convenient waiting time in-between are used to determine the current. This stability could not be reached without the introduction of the diffusion-limiting boxes in our model (Figure S1). Except for the general log-normal behavior, the simulated current distribution differs from the experimental data both in the mean current and especially in the width of the distribution. In real devices, as in our presented experimental data, naturally, there is a device-to-device variability due to the fabrication process and a cycle-to-cycle variability.³⁵ Adding a variability to the number of distributed V_{O} in the filament (42–52) and the filament size ($2\text{--}3 \times 2\text{--}3 \times 3\text{--}4 \text{ nm}^3$) in the simulation model leads to a slightly shifted and much wider current distribution, which looks very similar to the experimental one and can also be found in Figure 2 (red, Sim II). This additional variation of parameters defining the

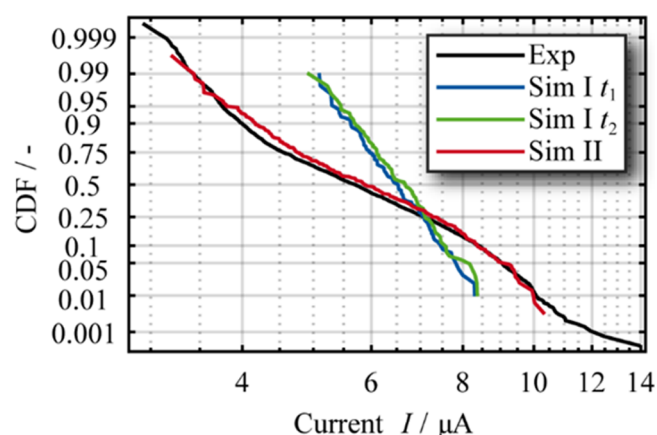


Figure 2. Comparison of experimental read current distribution of the ZrO_2 cells (black) with the simulation results of the KMC model. The blue curves show the initial results (t_1) of a fixed number of V_{O} placed randomly in a predefined filament volume and after 1 s (t_2). Both show log-normal behavior. Adding device-to-device variability modeled by a variable number of V_{O} and filament size, the simulated distribution (red) broadens and resembles the experimental data.

initial state obviously better represents the experimental results. Nevertheless, this variation is left out in the following due to the focus of this work laying on the stochastic processes during read and baking. In addition, excluding these additional parameters leads to a huge reduction of computation time costs.

Although the experimental and simulated current distributions are stable over time, the current of a single cell changes over time. Following our previous work,¹⁶ where we proposed random jumps of V_{O} to be the origin of the observed current jumps, we now investigate the dynamic behavior with the simulation model. In Figure 3a, the current evolution over time is presented for three exemplary experimental (gray) and simulated (colored) current traces each. In general, both, experimental and simulated traces, show discrete jumps of different heights as well as no great changes in the mean current. The traces mainly differ in the high-frequency noise with low amplitude that is observed in the experimental data but not in the simulations. This can be easily explained by the fact that there is no source of electronic noise implemented in our simulation model, which may be the origin of the high-frequency noise in the experimental data, often called the random telegraph noise (RTN) in the literature.^{12,36} In an enlarged view of the simulated data, as exemplarily shown in Figure 3b for the red curve (S1), the highest jumps in the current can be identified. In most cases, these large current jumps are caused by vertical jumps of V_{O} at or close to the filament-gap interface. The next smaller current changes correspond to vertical V_{O} jumps farther from the interface, whereas horizontal V_{O} jumps usually only have a small effect on the read current. Thus, the results of the dynamic simulations with the new box model are in perfect agreement with our previous results, where we found V_{O} jumps from the filament to the gap and vice versa to be the most important to discuss read variability in the HRS of VCM ReRAM devices.¹⁶

One common way to analyze and characterize noise is the power spectral density (PSD) plot.³⁷ In Figure 4, the PSD plots of the experimental and simulated current traces are shown. The magnitude of their PSD values is very similar, and both show a $1/f^2$ dependency, which corresponds to a

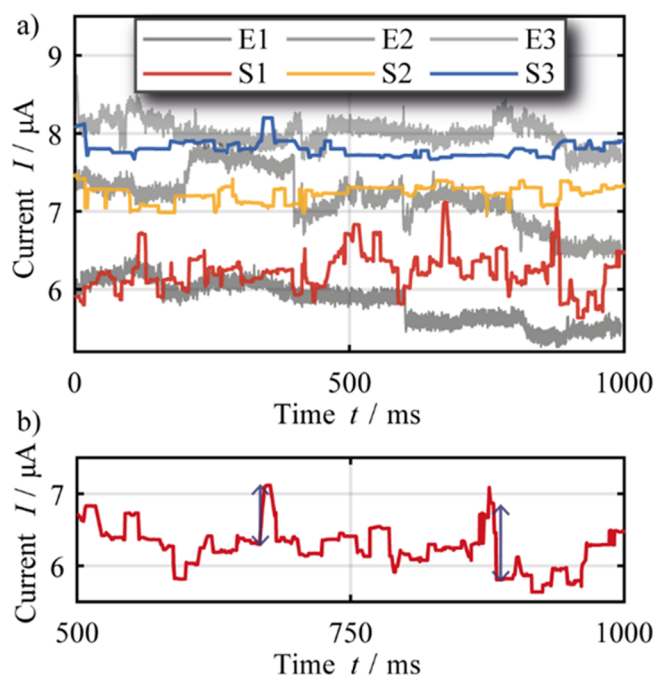


Figure 3. (a) Current evolution over time of exemplary experimental (gray) and simulated (colored) VCM cells in the HRS under applied read voltage. In the zoom-in of the red curve (S1) in panel (b), the highest current jumps (blue arrows) can be identified as vertical jumps of V_{O} close to the filament-gap interface.

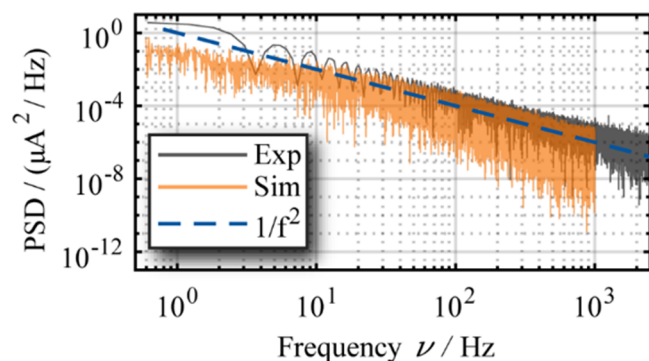


Figure 4. PSD plot of experimental (gray) and simulated (orange) current traces. Both data show a Lorentzian noise behavior with a $1/f^2$ dependency, which is shown by the dashed blue line. The high-frequency noise at the right side is missing in the simulated data.

Lorentzian noise. This accordance is further evidence that our basic assumption of randomly moving V_{O} originating read noise phenomena in VCM ReRAMs is valid. As already mentioned before, electronic noise is not covered by our simulation model. In a previous work, it was found that the current jumps caused by electron trapping and detrapping are much smaller than the effect of V_{O} diffusion.³⁸ Therefore, the high-frequent jumps with very low amplitude, which can be seen in the lower right of the experimental data of Figure 4, are not covered by the simulations. Since we asserted before that the highest jumps in the current are the most problematic for the reliability of the devices, it is reasonable to focus only on the current jumps caused by V_{O} movement.

A major problem of HRS instability in VCM ReRAMs is the so-called shaping failure.¹⁵ With program-verify or shaping algorithms, where cells with high currents are reprogrammed

or removed, an enlargement of the read window between the HRS and LRS shall be achieved. Unfortunately, these approaches are reported to have no lasting effect.¹³ The reasons for this failure are so far not fully understood but shall also be explained by the random movement of V_{O} in our simulation model. Therefore, the time evolution of 100 cells in the HRS generated, as described for Sim I in Figure 2, is investigated. To enlarge the read window, all cells with a current value above $7.5 \mu\text{A}$ at the start time (0 ms) are removed. The resulting current distribution is shown in black in Figure 5. Subsequently, the distribution relaxes over time

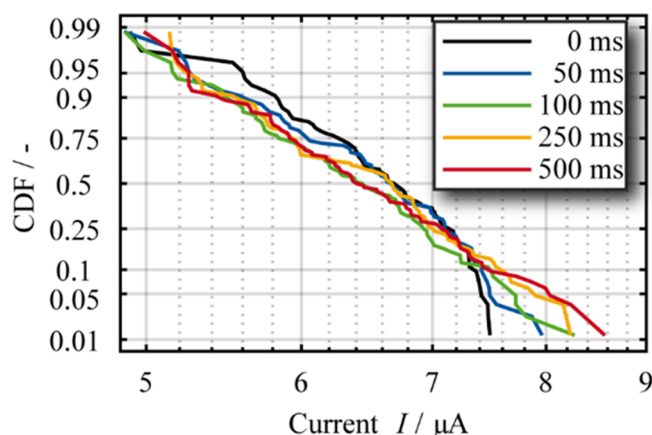


Figure 5. Simulated current distribution of 100 cells. At $t = 0$ ms (black), all cells with a current higher than $7.5 \mu\text{A}$ are removed. Due to the V_{O} movement in the individual cells, the distribution relaxes over time until the intrinsic log-normal distribution is reestablished after 500 ms (red).

and the log-normal behavior emerges from the shaped distribution so that the shaping effect is gone after about 500 ms. Subsequently, the log-normal distribution is stable again. The relaxation time observed in our simulations furthermore fits nicely to experimental results of comparable cells in ref 16, where the recovery of the log-normal shape took about 750 ms.

To sum up our results concerning the investigations of the instability of VCM ReRAMs in the HRS at room temperature and under read conditions, the basic assumption of randomly moving V_{O} can explain different phenomena. It is responsible for the intrinsic log-normal current distribution, the read noise in individual cells, and the shaping failure. Here, the introduced diffusion-limiting boxes are of major importance. Without the boxes, the two scenarios that are presented in Figure S1 are possible depending on the diffusion energy barrier for V_{O} in the oxide layer. Assuming comparably high diffusion energy barriers, no V_{O} would move at room temperature without or under low applied voltages. Therefore, the read noise as well as the shaping failure could not be explained. Assuming lower diffusion energy barriers as in our simulations and as it is also done in other simulations³⁹ would not lead to stable distributions without the limiting boxes. V_{O} would randomly diffuse through the whole oxide layer, and no stable filament would be possible. Already after seconds, the current of all cells would decrease and no longer resemble the stability of real devices.

In the second part of the results, we want to focus on the retention failure typically occurring in VCM ReRAM devices at long time scales. With our newly introduced simulation model with diffusion-limiting boxes, we want to give a suggestion for

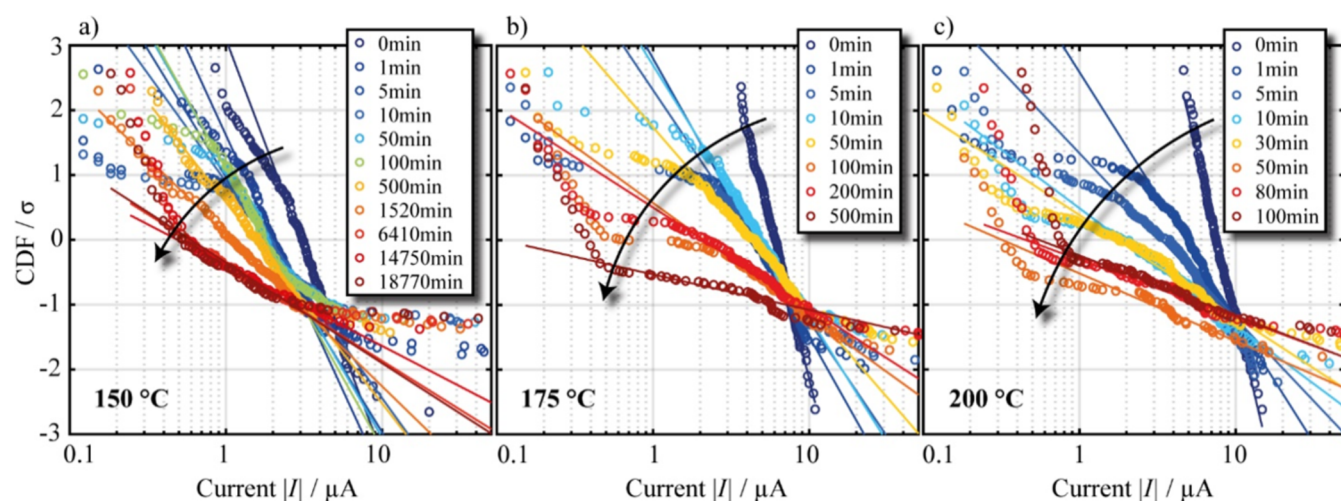


Figure 6. Experimental retention data. On three separate samples, a minimum of 200 cells each is programmed into the blue distribution (0 min). The samples are baked at (a) 150 °C, (b) 175 °C, or (c) 200 °C. At the increments stated in the legends, each sample is taken out of the oven and all cells are read out at 0.2 V (200 μ s). In all cases, the distribution broadens significantly (increasing standard deviation σ) and shifts toward the lower read current (decreasing median μ).

the processes leading to the change of the initially very stable current distributions at large time scales. Again, we propose the random movement of V_O to play a major role here. Originally, retention phenomena can be observed at very large time scales of several years, where the programmed resistance state of the cells changes without any external interference except the always present room temperature. Since it is impractical to perform experimental investigations with a duration of several years, “baking” experiments are used to simulate the thermally excited processes usually taking place over years in much shorter time scales of minutes up to a few hours. In the simulations, a comparable dodge has to be exerted by strongly increasing the temperature. Therefore, the long-time retention phenomena can be investigated with an admissible amount of computational resources. This strong increase in temperature leads to very short time scales on which the retention processes take place in the simulation but is necessary to depict as much V_O from box to box as possible.

To perform thermally accelerated retention experiments, the resistance or read current could be monitored by continuously reading at a low but constant voltage, while the sample was kept at elevated temperatures. This approach, however, has two disadvantages: On the one hand, it adds an additional voltage stress, which could affect the result. On the other hand, this limits the number of cells that can be tested within a reasonable measurement time and thus lowers the achievable statistics. As we demonstrated before, it is crucial to evaluate retention characteristics regarding the intrinsic statistics of the devices.³³ Thus, we prepared three samples as explained in the Experimental Section. On each sample, a minimum of 200 cells is electroformed, pre-cycled 20 times, and then programmed into the blue HRS distributions shown in Figure 6 (0 min). Afterward, the samples are stored in an oven at 150, 175, or 200 °C. After the bake times stated in the legends of Figure 6, the samples are taken out of the oven and cooled down to room temperature and the state of all prepared cells is determined by a read pulse of 200 μ s at 0.2 V.

The resulting read current distributions in Figure 6 are impacted by the bake in two ways: The distributions are (i) shifted toward the lower read current with increasing bake time, resulting in a decrease in the median current μ , and (ii) a

significant broadening of the distributions is observed, i.e., increasing standard deviation σ . Both trends are extracted from the recorded data and displayed in Figure 7. Here, the impact of the thermal acceleration becomes clear as both effects are increased with the bake temperature. Meanwhile, a decrease in μ might be beneficial for the application, and as it drives the HRS state further away from the LRS, the latter reduces the read window due to the observed broadening, which we have also demonstrated on industrial VCM ReRAMs.³³ In

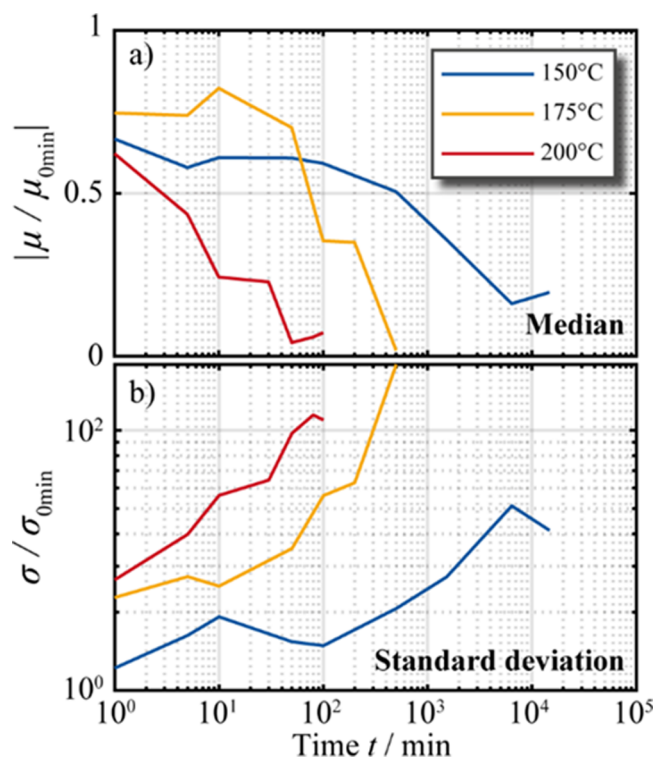


Figure 7. Trend of the characteristic parameters over the bake time, extracted from the read current distributions in Figure 6, and normalized to their initial value. (a) Decreasing median μ . (b) Increasing standard deviation σ .

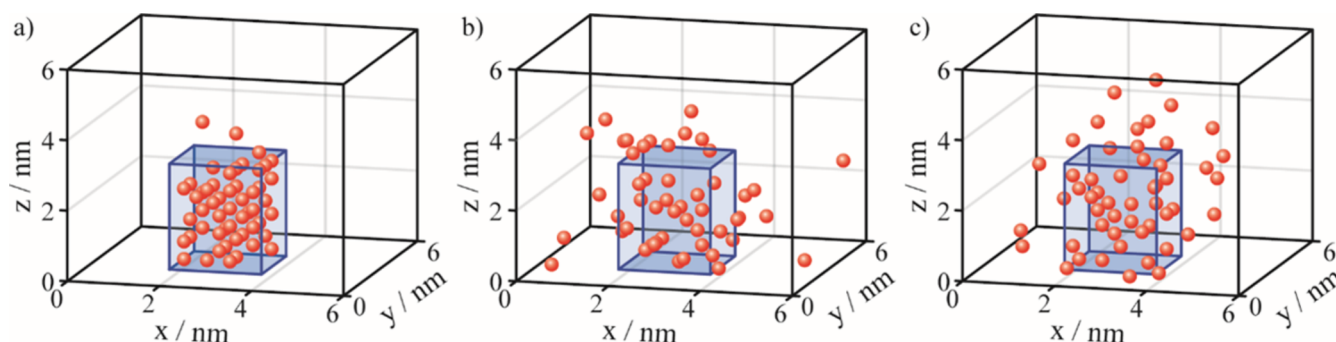


Figure 8. Exemplary V_O (red) configurations in the oxide layer. For simplification, the diffusion-limiting boxes of the presented simulation model are not shown here but only the filamentary limits (blue) of the initial HRS configuration. In panel (a), a typical start HRS configuration is shown. The other configurations exemplarily show final states after the baking simulations. In panel (b), radial diffusion was dominant, whereas in panel (c), the vertical diffusion toward the gap prevailed.

conclusion, the experimental retention results identify a broadening of the read current distribution as a major retention challenge. Thus, our model aims toward a deeper understanding of the underlying physical mechanisms, which should be compatible with the already discussed short-term instability.

In general, the simulations concerning the retention phenomena are very similar to the simulations concerning instability before. The size and the barriers of the diffusion-limiting boxes have not changed as well as the initial random placement of V_O modeling a VCM ReRAM in the HRS. As mentioned before, the temperature is highly increased to 1000 K to simulate the long-term retention behavior at more manageable time scales. Again, the V_O jump inside the diffusion-limiting boxes, as already observed in the instability simulations. Additionally, the strongly increased temperature allows V_O to overcome the barriers of the boxes so that also box-to-box jumps are possible. Due to the V_O concentration gradient created by the high number of V_O in the filament compared to the gap or surrounding region with no or only few V_O , the box-to-box jumps in total have a predominant direction that leads to structural changes of the V_O configuration in the oxide layer of the VCM cell. Generation and recombination processes of V_O at the metal/oxide interfaces, which would also be possible under these conditions, are neglected in our simulations concentrating only on the effect of V_O diffusion. We will demonstrate that this process alone is sufficient to model typical retention characteristics. In total, two main processes can be observed. On the one hand, V_O can jump toward the active electrode, thereby closing the current limiting gap between the filament and active electrode. This leads to higher currents flowing through the cell during readout. On the other hand, V_O can diffuse in the radial direction. Thus, the filament widens, and the read current becomes lower. Both processes take place randomly, have a different strong influence on the read current, and superimpose each other. Therefore, no clear differentiation between the processes can be done. In Figure 8, three exemplary V_O configurations are presented illustrating the possible effects of the box-to-box jumps. In Figure 8a, a typical initial V_O configuration is shown. The V_O are represented by the red balls, and the initial HRS filament region is shown by the blue box. The diffusion-limiting boxes are not shown here for a clearer view. In Figure 8b,c, two exemplary V_O configurations after baking are presented. In Figure 8b the radial diffusion of V_O was the dominating effect, whereas in Figure 8c, the

diffusion of V_O toward the active electrode preponderated. Both configurations are rare and rather extreme cases. Most of the simulated final states show a more balanced ratio between both processes.

Again, the statistical behavior of multiple cells gives more information than looking at a single cell. Therefore, the baking simulations at 1000 K were performed for 50 cells each. The results of two exemplary baking simulations differing in their initially programmed HRS are presented in Figures 9 and 10. The cells of the baking simulation presented in Figure 9 have initially been programmed to an HRS with a comparatively wide filament ($3 \times 3 \times 3 \text{ nm}^3$) containing 50 randomly placed V_O and three V_O in the box layer above. Here, the effects of the radial diffusion and the diffusion into the gap region are well

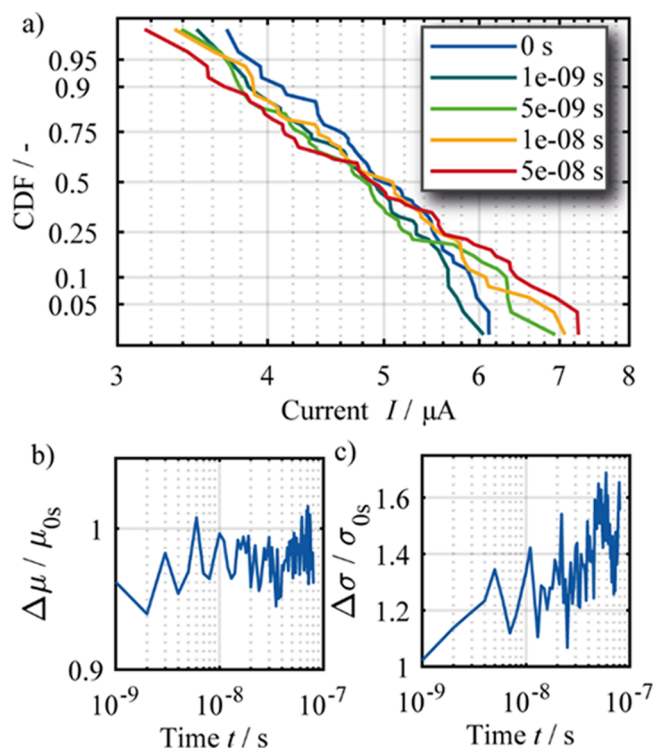


Figure 9. (a) Evolution of the simulated current distribution of 50 cells during baking at 1000 K. Initially, 50 V_O are placed in a filament volume of $3 \times 3 \times 3 \text{ nm}^3$ with three additional V_O in the box layer above. The mean current in panel (b) is mainly constant, whereas the standard deviation of the distribution in panel (c) increases over time.

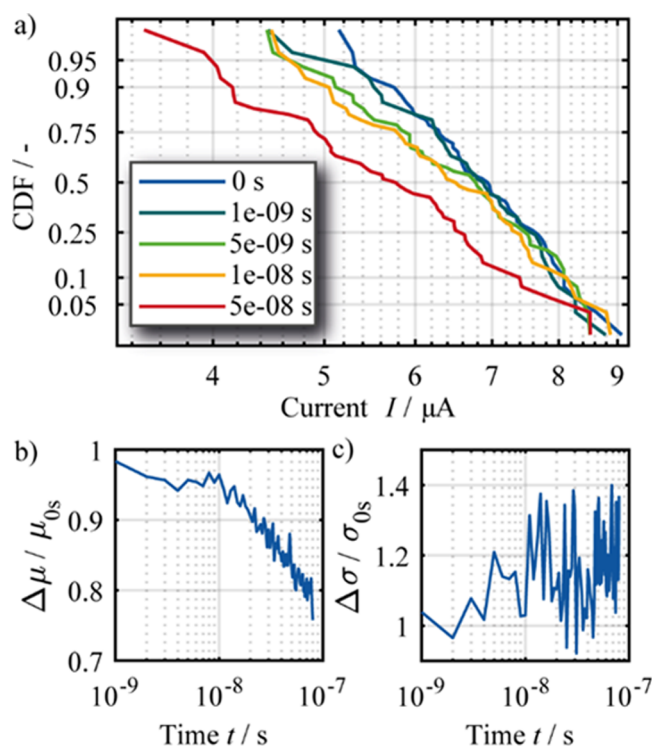


Figure 10. (a) Evolution of the simulated current distribution of 50 cells during baking at 1000 K. Initially, 50 V_O are placed in a filament volume of $2 \times 2 \times 3 \text{ nm}^3$ with five additional V_O in the box layer above. The mean current in panel (b) decreases, whereas the standard deviation of the distribution in panel (c) increases over time.

balanced. Thus, the mean current (Figure 9b) of the 50 cells is mainly constant except for the random statistical fluctuations due to the low number of simulated cells compared to the experimental data. On the contrary, the current distribution clearly widens due to both diffusion processes away from the initial filament. This widening can be further confirmed by looking at the increase in the standard deviation of the current distribution in Figure 9c. This increase in the standard deviation fits to the experimental results in Figure 7b as well as to the observations of other groups.^{20,40}

The cells of the baking simulation presented in Figure 10 have initially been programmed to a narrower filament ($2 \times 2 \times 3 \text{ nm}^3$) containing 50 randomly placed V_O and five V_O in the box layer above. Here, the increase in the standard deviation of the current distribution can be observed as well as in Figure 10c. In contrast to the results of the simulations in Figure 9, the mean current is not constant, but clearly decreasing (Figure 10b), due to the dominant radial diffusion. The combination of both trends, the decreasing mean current and the increasing standard deviation, is in good agreement with the experimental results presented in this work.

In addition to the two presented examples, also, other boundary conditions are conceivable, where the diffusion of V_O into the gap is dominating, leading to an increase in the mean current of the devices during baking. These considerations fit to the fact that, in the literature, decreasing as well as increasing median currents have been observed as HRS retention failure for VCM ReRAM devices.^{33,40,41} Thinking of entropy, the increase in the standard deviation of the current distributions in all simulations is also reasonable. Additionally, it should be mentioned that, in addition to the diffusion of V_O ,

the generation or recombination of V_O at the ohmic electrode could also play a role at large time scales or at the elevated temperatures in baking experiments. However, as shown before, the diffusion of V_O in our simulation model alone is sufficient to explain the general behavior (tilt and shift) of the current distribution during baking. The broadening of the read current distributions observed in retention experiments necessarily involves single cells with increasing read current. Existing models that neglect the diffusion of V_O could only explain a current increase by the stress-induced generation of oxygen vacancy/interstitial pairs in the gap region.^{28,42} However, such pairs have been demonstrated to be unable to form without any externally applied field and even if would recombine immediately.^{30,31} During retention, there is also no current flow that could stabilize the formation of anti-Frenkel pairs.⁴³ In contrast, we present a model that is based on the diffusion of V_O , which explains the experimentally observed retention effects as well as the short-term instability consistently.

CONCLUSIONS

We have presented a 3D KMC simulation model explaining several instability and retention failure mechanisms in filamentary VCM ReRAMs. Therefore, we introduced diffusion-limiting domains to the oxide layer of the device in this work. The main physical process originating from the presented instability and retention phenomena was found to be the random diffusion of V_O . The simulated results have been compared with experimental data of ZrO_2 -based devices.

The log-normal behavior of the HRS read current distribution can be explained by the randomness of the V_O configuration. Typical read noise behavior with current jumps at short time scales in single devices at the one hand and the overall stable current distribution of multiple devices on the other hand have been depicted by our simulation model. Furthermore, the shaping failure has been explained accurately in the same manner. Finally, retention failure like the tilt and shift of the current distribution of the programmed HRS in baking experiments has also been explained by the diffusion of V_O in our model. Due to the introduction of the domains in our model, both the short-time instability and the long-time retention phenomena can be explained and simulated in the same model based on the same physical origin of V_O diffusion.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsami.1c14667>.

Additional simulation results demonstrating the need of the introduced diffusion-limiting domains to obtain stable resistance states and to observe read noise (PDF)

AUTHOR INFORMATION

Corresponding Author

Nils Kopperberg – Institut für Werkstoffe der Elektrotechnik II (IWE2) and JARA-FIT, RWTH Aachen University, Aachen 52074, Germany; orcid.org/0000-0001-5422-7372; Email: kopperberg@iwe.rwth-aachen.de

Authors

Stefan Wiefels – Peter-Grünberg-Institut 7 (PGI-7), Forschungszentrum Jülich GmbH, Jülich 52425, Germany

Sergej Liberda – Institut für Werkstoffe der Elektrotechnik II (IWE2) and JARA-FIT, RWTH Aachen University, Aachen 52074, Germany

Rainer Waser – Institut für Werkstoffe der Elektrotechnik II (IWE2) and JARA-FIT, RWTH Aachen University, Aachen 52074, Germany; Peter-Grünberg-Institut 7 (PGI-7) and Peter-Grünberg-Institut 10 (PGI-10), Forschungszentrum Jülich GmbH, Jülich 52425, Germany

Stephan Menzel – Peter-Grünberg-Institut 7 (PGI-7), Forschungszentrum Jülich GmbH, Jülich 52425, Germany;

orcid.org/0000-0002-4258-2673

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsami.1c14667>

Author Contributions

N.K. performed the simulations, interpreted the data, and wrote the manuscript. S.W. conducted the experiments and co-wrote the manuscript. S.L. co-conducted the experiments. R.W. supervised the research. S.M. conceived the idea and initiated and supervised the research.

Funding

This work was supported in part by the Deutsche Forschungsgemeinschaft under Project SFB 917 and in part by the Federal Ministry of Education and Research (BMBF, Germany) in the Project NEUROTEC under grants 16ES1134 and 16ES1133K. It is based on the Jülich Aachen Research Alliance (JARA-FIT).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the computing time for the project cjpgi70 granted by the JARA Vergabegremium and provided on the JARA Partition part of the supercomputer JURECA at Forschungszentrum Jülich.⁴⁴ Prof. Christoph Jungemann and Dr. Elhameh Abbaspour are acknowledged for making their kinetic Monte Carlo simulation tool available for us within the SFB 917 project.

REFERENCES

- (1) Waser, R.; Dittmann, R.; Staikov, G.; Szot, K. Redox-Based Resistive Switching memories - Nanoionic Mechanisms, Prospects, and Challenges. *Adv. Mater.* **2009**, *21*, 2632–2663.
- (2) Chen, Y. ReRAM: History, Status, and Future. *IEEE Trans. Electron Devices* **2020**, *67*, 1420–1433.
- (3) Slesazek, S.; Mikolajick, T. Nanoscale Resistive Switching Memory Devices: A Review. *Nanotechnology* **2019**, *30*, 352003.
- (4) Pan, F.; Gao, S.; Chen, C.; Song, C.; Zeng, F. Recent progress in resistive random access memories: materials, switching mechanisms and performance. *Mater. Sci. Eng., R* **2014**, *83*, 1–59.
- (5) Aldana, S.; García-Fernández, P.; Romero-Zalaz, R.; González, M. B.; Jiménez-Molinos, F.; Gómez-Campos, F.; Campabadal, F.; Roldán, J. B. Resistive switching in HfO₂ based valence change memories, a comprehensive 3D kinetic Monte Carlo approach. *J. Phys. D: Appl. Phys.* **2020**, *53*, 225106.
- (6) Burr, G. W.; Shelby, R. M.; Sebastian, A.; Kim, S.; Kim, S.; Sidler, S.; Virwani, K.; Ishii, M.; Narayanan, P.; Fumarola, A.; Sanches, L. L.; Boybat, I.; Le Gallo, M.; Moon, K.; Woo, J.; Hwang, H.; Leblebici, Y. Neuromorphic computing using non-volatile memory. *Adv. Phys.: X* **2017**, *2*, 89–124.
- (7) Li, C.; Wang, Z.; Rao, M.; Belkin, D.; Song, W.; Jiang, H.; Yan, P.; Li, Y.; Lin, P.; Hu, M.; Ge, N.; Strachan, J. P.; Barnell, M.; Wu, Q.; Williams, R. S.; Yang, J. J.; Xia, Q. Long Short-Term Memory

Networks in Memristor Crossbar Arrays. *Nat. Mach. Intell.* **2019**, *1*, 49–57.

(8) Waser, R.; Bruchhaus, R.; Menzel, S. Redox-Based Resistive Switching Memories. In *Nanoelectronics and Information Technology*; 3rd ed.; Wiley-VCH: Hoboken, NJ, USA, 2012; pp. 683–710

(9) Fukuyama, S.; Maeda, K.; Matsuda, S.; Takeuchi, K.; Yasuhara, R. Suppression of Endurance-Stressed Data-Retention Failures of 40nm TaOx-Based ReRAM. 2018 IEEE International Reliability Physics Symposium (IRPS); IEEE, 2018, pp. P-MY.4–1-P-MY.4–5, DOI: 10.1109/IRPS.2018.8353677

(10) Wu, E.; Ando, T.; Kim, Y.; Muralidhar, R.; Cartier, E.; Jamison, P.; Wang, M.; Narayanan, V. Filamentary Statistical Evolution from Nano-Conducting Path to Switching-Filament for Oxide-RRAM in Memory Applications. 2019 IEEE International Electron Devices Meeting (IEDM); IEEE, 2019, pp. 30.3.1–30.3.4, DOI: 10.1109/IEDM19573.2019.8993494

(11) Menzel, S.; Hur, J.-H. Modeling the VCM- and ECM-Type Switching Kinetics. In *Resistive Switching*; Wiley-VCH: Hoboken, NJ, USA, 2016; pp. 395–436

(12) Brivio, S.; Frascaroli, J.; Covi, E.; Spiga, S. Stimulated Ionic Telegraph Noise in Filamentary Memristive Devices. *Sci. Rep.* **2019**, *9*, 6310.

(13) Fantini, A.; Gorine, G.; Degraeve, R.; Goux, L.; Chen, C. Y.; Redolfi, A.; Klima, S.; Cabrini, A.; Torelli, G.; Jurczak, M. Intrinsic Program Instability in HfO₂ RRAM and Consequences on Program Algorithms. 2015 IEEE International Electron Devices Meeting (IEDM); IEEE, 2015, pp. 7.5.1–7.5.4, DOI: 10.1109/IEDM.2015.7409648

(14) Puglisi, F. M.; Larcher, L.; Padovani, A.; Pavan, P. A complete statistical investigation of RTN in HfO₂-based RRAM in high resistive state. *IEEE Trans. Electron Devices* **2015**, *62*, 2606–2613.

(15) Perez, E.; Grossi, A.; Zambelli, C.; Olivo, P.; Roelofs, R.; Wenger, C. Reduction of the cell-to-cell variability in Hf1–xAlxOy based RRAM arrays by using program algorithms. *IEEE Electron Device Lett.* **2017**, *38*, 175–178.

(16) Wiefels, S.; Bengel, C.; Kopperberg, N.; Zhang, K.; Waser, R.; Menzel, S. HRS Instability in Oxide-Based Bipolar Resistive Switching Cells. *IEEE Trans. Electron Devices* **2020**, *67*, 4208–4215.

(17) Gao, B.; Kang, J. F.; Zhang, H. W.; Sun, B.; Chen, B.; Liu, L.; Liu, X. Y.; Han, R. Q.; Wang, Y. Y.; Fang, Z.; Yu, H. Y.; Yu, B.; Kwong, D.-L. Oxide-based RRAM: Physical based retention projection. 2010 Proceedings of the European Solid State Device Research Conference; IEEE, 2010, pp. 392–395, DOI: 10.1109/ESSDERC.2010.5618200

(18) Puglisi, F. M.; Zagni, N.; Larcher, L.; Pavan, P. Random telegraph noise in resistive random access memories: Compact modeling and advanced circuit design. *IEEE Trans. Electron Devices* **2018**, *65*, 2964–2972.

(19) Aldana, S.; Pérez, E.; Jiménez-Molinos, F.; Wenger, C.; Roldán, J. B. Kinetic Monte Carlo analysis of data retention in Al:HfO₂-based resistive random access memories. *Semicond. Sci. Technol.* **2020**, *35*, 115012.

(20) Huang, P.; Xiang, Y. C.; Zhao, Y. D.; Liu, C.; Gao, B.; Wu, H. Q.; Qian, H.; Liu, X. Y.; Kang, J. F. Analytic Model for Statistical State Instability and Retention Behaviors of Filamentary Analog RRAM Array and Its Applications in Design of Neural Network. 2018 IEEE International Electron Devices Meeting (IEDM); IEEE, 2018, pp. 40.4.1–40.4.4, DOI: 10.1109/IEDM.2018.8614567

(21) Abbaspour, E.; Menzel, S.; Hardtdegen, A.; Hoffmann-Eifert, S.; Jungemann, C. KMC simulation of the electroforming, set and reset processes in redox-based resistive switching devices. *IEEE Trans. Nanotechnol.* **2018**, *17*, 1181–1188.

(22) Schie, M.; Müller, M. P.; Salinga, M.; Waser, R.; De Souza, R. A. Ion migration in crystalline and amorphous HfOx. *J. Chem. Phys.* **2017**, *146*, No. 094508.

(23) Rumble, J. *CRC handbook of Chemistry and Physics*; 98th Edition, CRC Press LLC: Boca Raton, FL, USA, 2017

(24) Funck, C.; Menzel, S. Comprehensive model of electron conduction in oxide-based memristive devices. *ACS Appl. Electron. Mater.* **2021**, *3*, 3674–3692.

- (25) Ielmini, D.; Waser, R. *Resistive Switching: From Fundamentals of Nanoionic Redox Processes to Memristive Device Applications*; Wiley-VCH: Hoboken, NJ, USA, 2016.
- (26) Larcher, L. Statistical simulation of leakage currents in MOS and flash memory devices with a new multiphonon trap-assisted tunneling model. *IEEE Trans. Electron Devices* **2003**, *50*, 1246–1253.
- (27) Guan, X.; Yu, S.; Wong, H.-S. P. On the Switching Parameter Variation of Metal-Oxide RRAM—Part I: Physical Modeling and Simulation Methodology. *IEEE Trans. Electron Devices* **2012**, *59*, 1172–1182.
- (28) Butcher, B.; Bersuker, G.; Gilmer, D. C.; Larcher, L.; Padovani, A.; Vandelli, L.; Geer, R.; Kirsch, P. D. Modeling the Effects of Different Forming Conditions on RRAM Conductive Filament Stability. *2013 5th IEEE International Memory Workshop*; IEEE, 2013, pp. 52–55, DOI: 10.1109/IMW.2013.6582096
- (29) Padovani, A.; Gao, D. Z.; Shluger, A. L.; Larcher, L. A Microscopic Mechanism of Dielectric Breakdown in SiO₂ Films: An Insight from Multi-Scale Modeling. *J. Appl. Phys.* **2017**, *121*, 155101.
- (30) Clima, S.; Chen, Y. Y.; Chen, C. Y.; Goux, L.; Govoreanu, B.; Degraeve, R.; Fantini, A.; Jurczak, M.; Pourtois, G. First-Principles Thermodynamics and Defect Kinetics Guidelines for Engineering a Tailored RRAM Device. *J. Appl. Phys.* **2016**, *119*, 225107.
- (31) Schie, M.; Menzel, S.; Roberston, J.; Waser, R.; De Souza, R. A. Field-Enhanced Route to Generating Anti-Frenkel Pairs in HfO₂. *Phys. Rev. Mater.* **2018**, *2*, No. 035002.
- (32) Wiefels, S.; Von Witzleben, M.; Hüttemann, M.; Büttger, U.; Waser, R.; Menzel, S. Impact of the Ohmic Electrode on the Endurance of Oxide-Based Resistive Switching Memory. *IEEE Trans. Electron Devices* **2021**, *68*, 1024–1030.
- (33) Wiefels, S.; Büttger, U.; Menzel, S.; Wouters, D. J.; Waser, R. Statistical modeling and understanding of HRS retention in 2.5 Mb HfO₂ based ReRAM. *2020 IEEE International Memory Workshop (IMW)*; IEEE, 2020, pp. 1–4, DOI: 10.1109/IMW48823.2020.9108123
- (34) Karpov, V. G.; Niraula, D. Log-normal statistics in filamentary RRAM devices and related systems. *IEEE Electron Device Lett.* **2017**, *38*, 1240–1243.
- (35) Schönhals, A.; Waser, R.; Wouters, D. J. Improvement of SET Variability in TaOx Based Resistive RAM Devices. *Nanotechnology* **2017**, *28*, 465203.
- (36) Puglisi, F. M.; Pavan, P. Guidelines for a Reliable Analysis of Random Telegraph Noise in Electronic Devices. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 1435–1442.
- (37) Yu, S.; Jeyasingh, R.; Wu, Y.; Wong, H.-S. P. Characterization of low-frequency noise in the resistive switching of transition metal oxide HfO₂. *Phys. Rev. B* **2012**, *85*, No. 045324.
- (38) Abbaspour, E.; Menzel, S.; Jungemann, C. Random telegraph noise analysis in redox-based resistive switching devices using KMC simulations. *2017 International Conference on Simulation on Semiconductor Processes and Devices (SISPAD)*; IEEE, 2017, pp. 313–316, DOI: 10.23919/SISPAD.2017.8085327
- (39) Dirkmann, S.; Kaiser, J.; Wenger, C.; Mussenbrock, T. Filament Growth and Resistive Switching in Hafnium Oxide Memristive Devices. *ACS Appl. Mater. Interfaces* **2018**, *10*, 14857–14868.
- (40) Chen, Y. Y.; Komura, M.; Degraeve, R.; Govoreanu, B.; Goux, L.; Fantini, A.; Raghavan, N.; Clima, S.; Zhang, L.; Belmonte, A.; Redolfi, A.; Kar, G. S.; Groeseneken, G.; Wouters, D. J.; Jurczak, M. Improvement of data retention in HfO₂/Hf 1T1R RRAM cell under low operating current. *2013 IEEE International Electron Devices Meeting*; IEEE, 2013, pp. 10.1.1–10.1.4, DOI: 10.1109/IEDM.2013.6724598
- (41) Lin, Y.-H.; Ho, Y.-H.; Lee, M.-H.; Wang, C.-H.; Lin, Y.-Y.; Lee, F.-M.; Hsu, K.-C.; Tseng, P.-H.; Lee, D.-Y.; Chiang, K.-H.; Wang, K.-C.; Tseng, T.-Y.; Lu, C.-Y. A comprehensive study of 3-stage high resistance state retention behavior for TMO ReRAMs from single cells to a large array. *2017 IEEE International Electron Devices Meeting (IEDM)*; IEEE, 2017, pp. 2.5.1–2.5.4, DOI: 10.1109/IEDM.2017.8268313
- (42) Padovani, A.; Larcher, L.; Pirrotta, O.; Vandelli, L.; Bersuker, G. Microscopic Modeling of HfO_x RRAM Operations: From Forming to Switching. *IEEE Trans. Electron. Devices* **2015**, *62*, 1998–2006.
- (43) Bersuker, G.; Gilmer, D. C.; Veksler, D.; Kirsch, P.; Vandelli, L.; Padovani, A. N. D. R. E. A.; Larcher, L.; McKenna, K.; Shluger, A.; Iglesias, V.; Porti, M.; Nafria, M. Metal Oxide Resistive Memory Switching Mechanism Based on Conductive Filament Properties. *J. Appl. Phys.* **2011**, *110*, 124518.
- (44) Krause, D.; Thürnig, P. JURECA: Modular supercomputer at Jülich supercomputing centre. *J. large-scale Res. Facil.* **2018**, *4*, A132.