Investigation of resistance fluctuations in ReRAM: physical origin, temporal dependence and impact on memory reliability

Lucas Reganaz, Damien Deleruyelle, Quentin Rafhay, Joel Minguet Lopez, Niccolo Castellani, Jean-François Nodin, Alessandro Bricalli, Giuseppe Piccolboni, Gabriel Molas, François Andrieu

Investigation of resistance fluctuations in ReRAM: physical origin, temporal dependence and impact on memory reliability

L. Reganaz[1,2], D. Deleruyelle[3], Q. Rafhay[4], J. Minguet Lopez[1], N. Castellani[1], J. F. Nodin[1], A. Bricalli[5], G. Piccolboni[5], G. Molas[5], F. Andrieu[1]

[1]CEA, LETI, MINATEC Campus, 38000, Grenoble, France
[2]Université Grenoble Alpes, 38000, Grenoble, France
[3]INL CNRS, INSA Lyon, 69621, Lyon, France
[4]Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS, Grenoble INP, IMEP-LAHC, 38000, Grenoble, France
[5]Weebit Nano Ltd

2023 IEEE International Reliability Physics Symposium (IRPS) | 978-1-6654-5672-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/IRPS48203.2023.10117882

***Abstract*— We investigate the impact of ReRAM resistance fluctuations in a 16kb memory array. ReRAM retention and read current fluctuations are the main factors limiting the reliability of the array. A KMC-based 3D simulation framework is introduced for a complete physical description of the observed mechanisms. After individual cell relaxation (up to one second at room temperature) and resistance distribution stabilization, single cell level fluctuations still occur, especially in HRS. Oxygen vacancy migration and recombination, RTN and 1/f noise components contribute to the dynamic evolution. In a first phase, higher and faster current fluctuations are measured due to RTN and V_O low energy migration (1-10min at 25°C). In a second phase, the contribution of V_O migration tends to decrease as they neutralize in clusters or diffuse (>10min at 25°C). Finally, the impact of individual cell fluctuations on the variability and reliability of memory array is analyzed.**

***Keywords*- Non-Volatile memory, ReRAM, RTN, Kinetic-Monte-Carlo simulations, Current fluctuations**

I. Introduction

Resistive random access memory (ReRAM) is considered a promising technology for future memory applications due to its high density, fast read and write speeds, low power consumption, good ability to withstand high temperatures, and its potential for co-integration with traditional CMOS technology [1]-[7]. ReRAM devices operation relies on resistance change between low (LRS) and high (HRS) resistive states. For oxide-based devices, as considered in this work, the switching mechanisms are based on dielectric breakdown and a redox process for the Set and Reset operations, respectively. During the reset operation, the oxygen vacancy (V_O) based conductive filament (CF) is ruptured. The HRS state is particularly sensitive to oxygen ion and vacancy migration [8]-[10] and random telegraphic noise (RTN), the latter being induced in

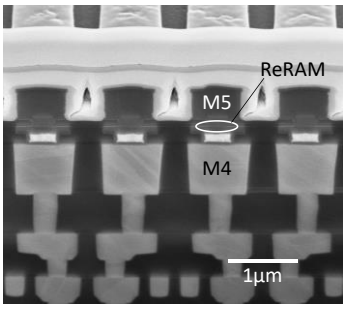

Figure 1 SEM cross section of Metal/Oxide/Metal ReRAM stack structure integrated in the BEOL of 130nm CMOS process into a 16kb 1T1R array [13].

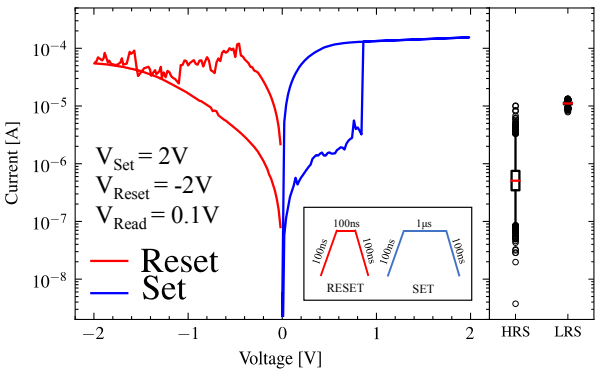

Figure 2 Typical I-V characteristics of Set and Reset operations. Operation voltages and pulse times are displayed. The right box depicts the read current distributions (@25°C) at V_{Read}=0.1V from the 16kb array for HRS and LRS (red dash: median; box: 50%; whiskers: 99.7%; markers: outliers).

HRS by electron capture/emission causing a momentary deactivation of a V_O [11], [12]. In addition to the variability in the intrinsic resistance of a ReRAM array arising from the programming operation, these physical phenomena pose a challenge to the performance evaluation and validation of the ReRAM product operation, as large resistance fluctuations and dynamic errors occur. This work aims to clarify the nature of resistance fluctuations and their impact on the statistical and dynamic behavior of ReRAM.

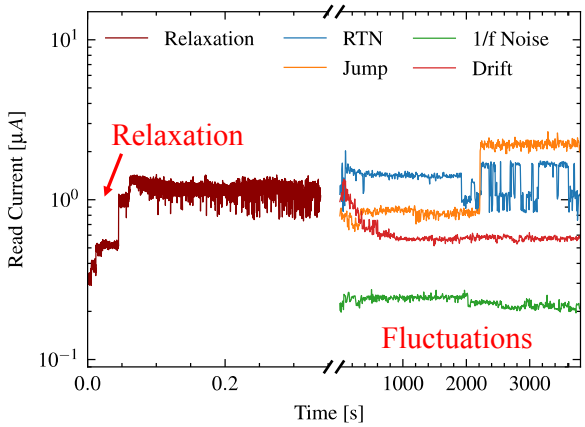

Figure 3 Time evolution of the ReRAM resistance after Reset in the High Resistive State. Left part illustrates a single cell initial relaxation (ms range). Right part illustrates four typical current fluctuations (s – hour range). Monitoring is performed at room temperature (25°C) at V_{Read}=0.1V.

II. Experimental Section

Experimental data are extracted from a 16kb ReRAM array (**Fig.1**) [13]. A typical I-V characteristic from the ReRAM is shown in **Fig.2**. The resistance monitoring after programming is performed on the entire array for 1 hour, every 6s, at ambient temperature and at V_{Read}=0.1V. The right box in **Fig.2** shows the range in read current for HRS and LRS for the 16kb array from which the read current fluctuations take place.

Fig.3 depicts a typical resistance time evolution through read current monitoring showing an early relaxation followed by a fluctuation-dominated phase with typical current fluctuation characteristics. In order to precisely identify the nature of the current fluctuations and their respective occurrence timeframe, a kinetic Monte Carlo (KMC) based simulation tool is used to replicate the conductive filament (CF) relaxation and V_O migration through time (**Fig.4**) [14].

III. Kinetic Monte Carlo Simulations

The simulation flowchart of KMC-based ReRAM, shown in **Fig.4** (a), is a multi-physics framework with several modules as described in reference [14]. It combines parameters from references [8]-[10] to simulate the motion, recombination, and generation of oxygen ions and vacancies during programming and for the next hour.

A key aspect of the simulation is the low energy V_O^{2+}/V_O^{+} motion and recombination, which leads to CF reorganization and the early post-programming relaxation effect (as reported in [15]). The simulation also shows that V_O density near the residual CF tip in HRS is lower and these traps are more likely to be charged (as reported in [16]). This results in a lower migration energy barrier [8]. The simulation allows for adjusting the V_O charge and migration activation energy based on the V_O's local density,

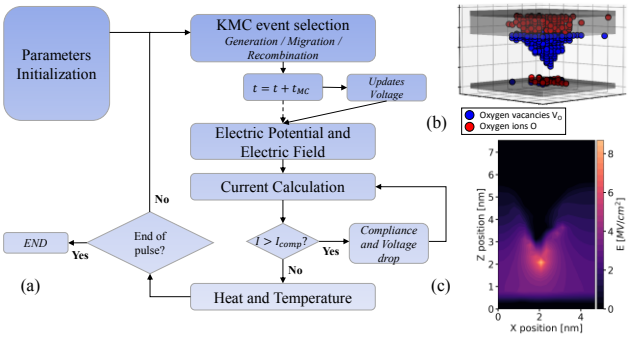

Figure 4 (a) Flowchart description of the KMC framework for the simulation of the ReRAM [14]. (b) Oxygen ions and vacancies representation of a conductive filament after a Reset operation. (c) Corresponding 2D cross-section electric field map during Reset.

emphasizing the importance of these factors in current fluctuation dynamics.

For this study, one hundred devices have been simulated through Forming, Set, and Reset processes under the same programming conditions as were used experimentally. In addition, the Set operation is a 1μs pulse at 2V, while the Reset operation performs a 100ns pulse at -2V. Current reading is also measured for V_{Read}=0.1V.

Fig.5 shows the different mechanisms in the ReRAM leading to current fluctuations and their respective main timeframe of occurrence from which three main phases are drawn. The mechanisms with the lowest energies, thus the first to occur, are oxygen ions and doubly charged V_O migrations. The recombination energy is usually negligible, leaving this event's dynamic mainly driven by oxygen interstitial diffusion with an energy of approximately 0.6eV [3]. The simulation showed that most of the recombination events occur within one second (**Fig.5** (a)), with negligible contribution on longer timescales. With low V_O^{2+} migration energy, these two fast mechanisms dominate the relaxation effect in the range of a few seconds.

Charged V_O diffusion and neutralization effects dominate the first post-relaxation fluctuation phase. **Fig.5** (b) shows the simulated $\Delta\log(I)_{60s}$ (total current range covered during 60s, c.f. **Fig.6**) for V_O migration located in the oxide gap or in the vicinity of the residual CF tip (V_O^{+}/V_O^{2+}). These vacancies still have a low migration energy that allows for a high pace random walk either away from the CF tip (diffusion) or onto it (neutralization). With an energy of up to 1.1eV, the diffusion and neutralization mechanisms contribute to the current fluctuations over a few minutes. Moreover, as these vacancies are potential RTN contributors, the RTN-generating vacancy population tends to decrease, as shown in **Fig.5** (f).

Fig.5 (d) shows that during the second fluctuation phase, vacancy activation does not significantly affect the current fluctuations. The simulation suggests that when a vacancy is located near a cluster, it is less likely for V_O dissociation or activation by charge trapping to occur. Nevertheless, it can still contribute to current fluctuations, even after a short period, approximately ten minutes. This is because it can lead to the diffusion of new vacancies (V_O^{+}/V_O^{2+}) or activate high amplitude RTN (V_O^{-}/V_O).

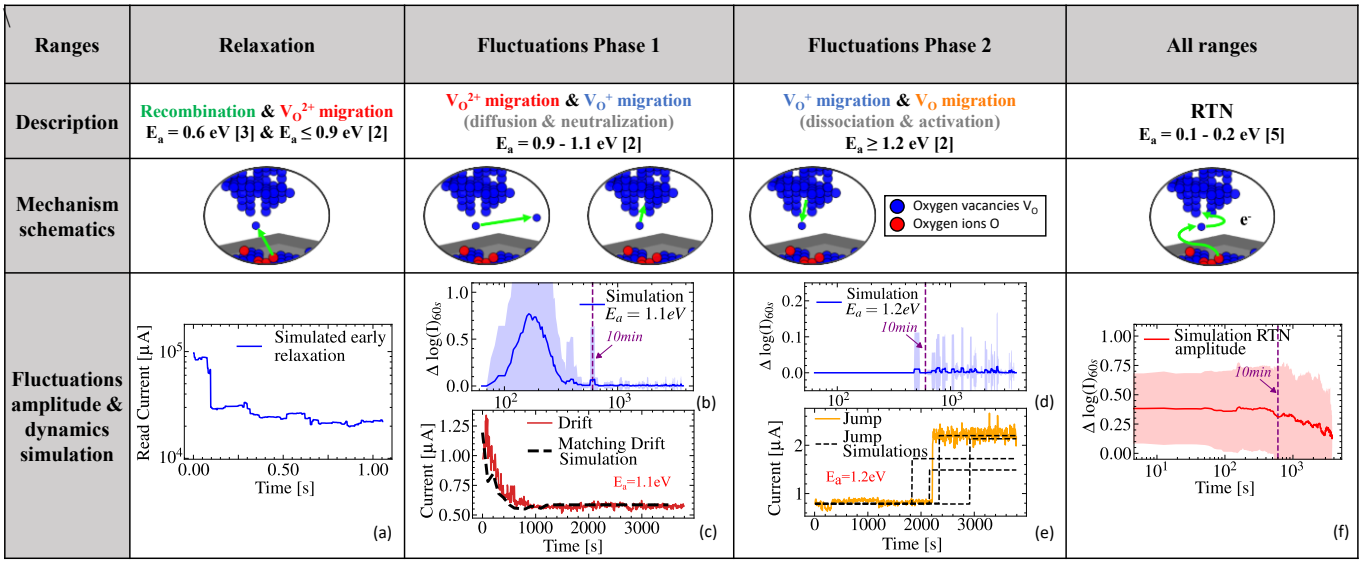

Figure 5 Description of current fluctuation phenomena in the ReRAM. (a) shows the simulation of post-programming relaxation involving V_O^{2+} migration and recombination. (b) shows the $\Delta\log(I)_{60s}$ (see Fig.6) current fluctuation amplitudes for V_O^{+}/V_O^{2+} migration, leading to diffusion away from the CF or neutralization onto it (random walk [6]). The simulation was performed with an energy of 1.1eV, mostly representing singly charged traps. (c) shows one *I(t)* matching diffusion simulation similar to the drift characteristics from Fig.4. (d) shows the simulated $\Delta\log(I)_{60s}$ current fluctuation amplitudes for V_O/V_O^{+} migation from within the CF, simulating potential trap activation. (e) shows a few *I(t)* activation simulations similar to the jump characteristics from Fig.3. (f) shows the simulated potential RTN amplitude. It is calculated from the read current change from the potential conduction deactivation of V_O's located near the tip of the residual CF.

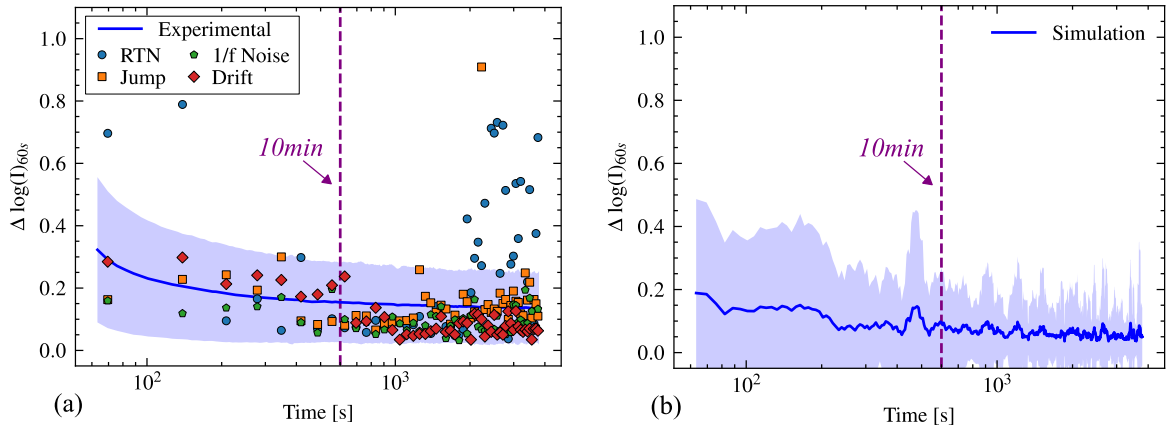

Figure 6 Experimental (a) and simulated (b) evolution of the current fluctuation amplitude calculated with $\Delta\log(I)_{60s} = \log(I_{max-60s}) - \log(I_{min-60s})$ on a 60s sliding window, with average (blue line) and ±1σ standard deviation (blue area). (a) Markers correspond to values of $\Delta\log(I)_{60s}$ from the four *I(t)* of Fig.3 (sampled every 30s for clarity). (b) simulation performed for a V_O migration energy between 1.0eV and 1.2eV, depending on V_O charge [8]-[10].

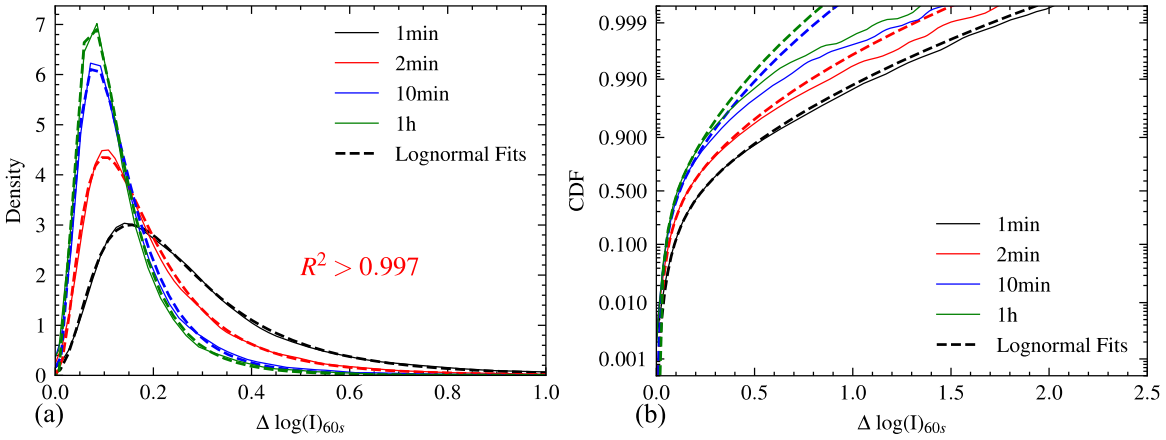

Figure 7 (a) Probability density functions and (b) cumulative distribution functions of $\Delta\log(I)_{60s}$ taken at four different times after Reset from experimental data with V_{Read}=0.1V at room temperature (25°C) with their respective lognormal fitting curves.

Fig.6 depicts the $\Delta\log(I)_{60s}$ for the experimental measurements (a) and the simulation from V_O migration and recombination (b). It shows that the ReRAM array fluctuations amplitude decreases with time and reaches a noise baseline after approximately 10min. The simulation confirmed that the trend of **Fig.6** (a) is due to charged traps in the oxide gap migrating away or neutralizing into the residual CF, leading to fewer high amplitude RTN or V_O migration fluctuations occurrences. **Fig.7** depicts the $\Delta\log(I)_{60s}$ distribution at four different times, showing a log-normal distribution trend for current fluctuation amplitudes. While the noise amplitude baseline is normally distributed (1/f and white noise), the amplitudes of the simulated current fluctuations due to V_O motions and RTN that are located near the residual CF tip have a non-linear dependency on the V_O position due to tunneling current effect. The convolution of both effects gives rise to the lognormal characteristics from **Fig.7**. The distribution tail for high current fluctuations seems to deviate from the lognormal fitting curves after a few minutes at room temperature (**Fig.7** (b)), but no physical explanation has been made for this observation.

IV. RELIABILITY IMPACT

A. Fluctuations and resistance distributions

As shown in **Fig.8**, the resistance distributions measured after 6s do not spread further after one hour. In order to generalize this current fluctuation study for reliability purposes, the $\Delta\log(I)_{60s}$ characteristic has been measured for the same 16kb array after endurance stress (**Fig.9**) for up to 10^5 cycles. $\Delta\log(I)_{60s}$ trends remain identical for both HRS and LRS. Apart from data in **Fig.9**, all experimental measurements have been performed on an approximately 10^2 cycled 16kb array. It can be observed that there is no read disturb effect on the fluctuation phenomenon. This is confirmed by comparing the $\Delta\log(I)_{60s}$ values from resistance monitoring after a 10-minute wait, which match with values from **Fig.6** (a) from 10 minutes to one hour.

The one-hour 16kb ReRAM array resistance monitoring allows one to extract extreme potential cases. In **Fig.10** (a), the red and blue distributions are, respectively, the distributions of the minimum and maximum resistances reached by each of the 16bk array cells. They act as the boundaries from which the HRS distribution has never escaped. A general resistance drift behavior can be noted in **Fig.10** (b) when slicing the resistance distribution into three parts [8], where the average resistance value of the two ends of the resistance distribution tend to move toward the overall average value. **Fig.10** (c) shows the 1-hour resistance fluctuation range for each of the ReRAM cells from the 16kb array. It can be noted that half of the cells travel across less than 3% of the entire distribution. However, no cells are close to moving entirely across it, as the measured maximum resistance relative swing was 44%. The one-hour resistance range each cell traveled across is also not correlated to its position in the initial resistance distribution, showing no specific part of HRS distribution with a tendency for large fluctuations.

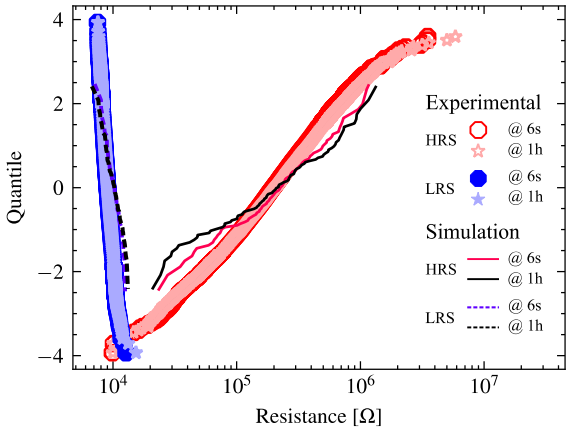

Figure 8 LRS and HRS 6s and 1h resistance distributions experimentally measured on the 16kb array and simulated. Monitoring is performed at room temperature (25°C) and V_{Read}=0.1V.

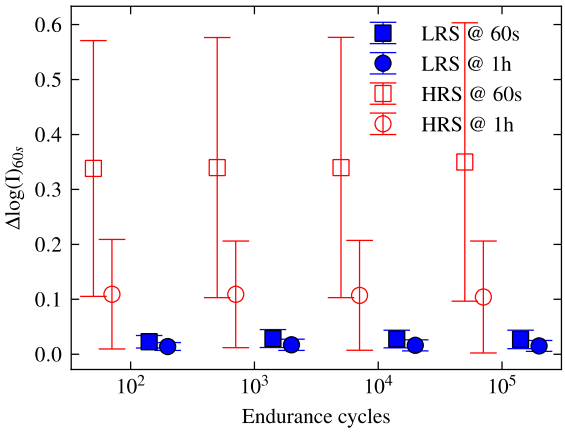

Figure 9 Impact of endurance cycling on current fluctuations. $\Delta\log(I)_{60s}$ values at t=60s and t=1h for LRS and HRS after endurance cycling.

B. Fluctuations and reliability assessment

The reliability assessment is made by setting a resistance threshold in between LRS and HRS levels. Among the three chosen threshold values (20kΩ, 30kΩ, and 40kΩ), the 20kΩ threshold gives an optimized (i.e. lowered) bit error rate for the given LRS and HRS distributions (**Fig.11**). However, the failing bit count analysis for three different threshold allows for understanding the fluctuation dynamic of the resistance distribution's lower tail. The failing bit count for the different resistance thresholds seem to follow the same dynamic, namely an early increase followed by a more stable fluctuation phase (**Fig.11** (a)). It is worth mentioning that this trend is mostly valid for the higher thresholds (green and red lines in **Fig.11**), which aligns with the post-relaxation stability of the resistance distribution's lower tail, as it has been seen in **Fig.8**. The early increase in the failing bit count is a direct result of the initial fluctuation phase, which is characterized by some fast and irreversible processes described in **Fig5** (b). The stable failing bit count during the second fluctuation phase suggests that the fluctuations are mainly driven by a reversible phenomenon known as RTN. The total cumulated failing bit count depicted in **Fig.11** (b), which is the number of cells that have crossed the resistance

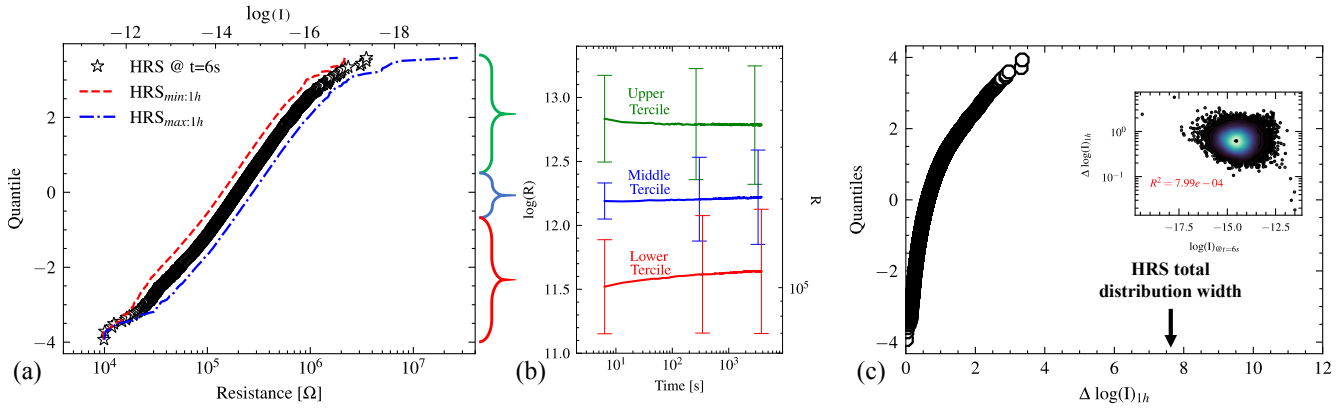

Figure 10 (a) Initial (6s) HRS resistance ReRAM distribution from experimental data. $HRS_{min:1h}$ ($HRS_{max:1h}$) represent the distribution of the minimum.(maximum) resistance of each cell seen over the one hour monitoring. (b) Logarithm of the resistance versus time averaged and ±1σ spread for lower (0-33%), middle (33%-66%) and upper (66%-100%) terciles of initial resistance after Reset. (c) Distribution of the total range of the logarithm of the read current during the one hour monitoring for the ReRAM cells of the 16kb array, at room temperature (25°C) and V_{Read}=0.1V. The median and maximum one hour full amplitude are, respectively, 3% and 44% of the distribution width in logarithmic scale. The inset depicts the relationship between the one hour full amplitude read current swing with respect to the logarithm value of initial read current at t=6s. It shows no linear correlation.

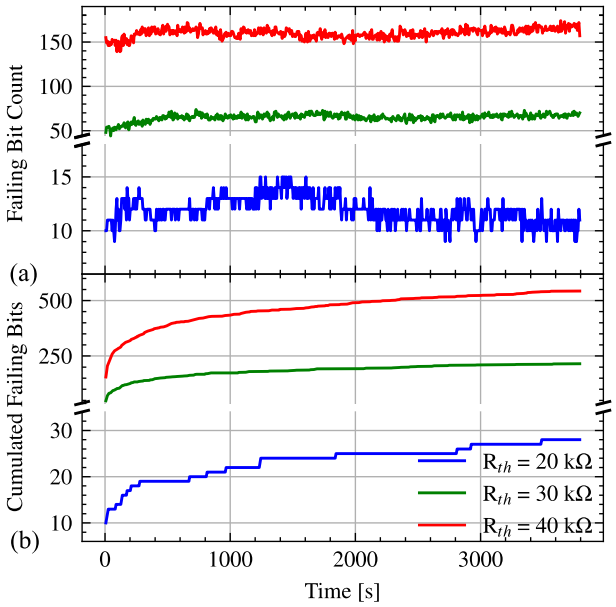

Figure 11 (a) Number of failing cells over time (fail ⇔ R<R_{th}) for three resistance thresholds. (b) Cumulative number of cells from the 16kb array that failed at least once.

TABLE I Fluctuation phases after Reset and reliability impact @25°C.

	Relaxation	Fluctuations Phase 1	Fluctuations Phase 2
Time Range	<<1s	1s-10min	>10min
Events	T°↘, RTN, low E_a migration (V_O^{2+}), recombinations	RTN, medium E_a migration (V_O^{+})	RTN, large E_a migration (V_O)
σ(R)	↗	→	→
Failing bits	↗	Low R_{th}: → High R_{th}: ↗	→

threshold at least once, continues to rise even after an hour, indicating that the RTN-dominated second fluctuation phase has a localized shuffling effect on the resistance distribution.

Table I summarizes the fluctuation phases timeframe, the involved physical mechanisms, and the impact on ReRAM array reliability. It shows the fluctuation trends that the array trimming for threshold value adjustment will depend on. In order to increase consistently the ReRAM array performance over time, the fluctuation dynamics that have been described in this study can be addressed. For instance, performing the ReRAM array failure analysis after the first fluctuation phase can lead to more stable devices and thus better performance over time. Moreover, this study showed that the cumulated failing bit count over time could be greater than the one given by a single read. It is why a multiple readings verification scheme could be more efficient for failure detection. Besides, the performance gain is even further enhanced when constraints are strengthened on the R_{th}. Low and high R_{th} in Table I corresponding to weak and strong constraints, respectively, on the state level discretization threshold.

Ultimately, the $\Delta\log(I)_{60s}$ characteristics and the simulation indicate that the first fluctuation phase timeframe would be significantly shortened at operating temperature (e.g. 85°C). This suggests that customized verification routines as described above can gain efficiency and execution speed even at fast operating regime (<<10min).

V. Conclusions

This work presents results on the statistical and dynamic behavior ReRAM technology and its impact on reliability. Resistance monitoring after programming was performed on a 16kb memory array for one hour at room temperature. A KMC simulation tool was used to replicate the conductive filament relaxation and V_O migration through time. The simulation showed that oxygen ions and V_O^{2+} migrations were the first to occur, and dominated the relaxation effect in the range of a few seconds. Charged V_O diffusion and

neutralization effects dominated the first post-relaxation fluctuation phase that lasted approximately ten minutes at 25°C.

The KMC simulation provided a precise understanding of the current fluctuations and their respective timeframe of occurrence, which is useful in understanding and controlling the reliability of ReRAM technology. As a continuation of this work, the analysis developed in this study will be employed for high temperatures operation, as well as the fluctuations implications on multi-level cell (MLC) application for which state discretization has tighter constraints.

REFERENCES

[1] C. -C. Chou et al., "A 22nm 96KX144 RRAM Macro with a Self-Tracking Reference and a Low Ripple Charge Pump to Achieve a Configurable Read Window and a Wide Operating Voltage Range," *2020 IEEE Symposium on VLSI Circuits*, Honolulu, HI, USA, 2020, pp. 1-2, doi: 10.1109/VLSICircuits18222.2020.9163014.

[2] O. Golonzka et al., "Non-Volatile RRAM Embedded into 22FFL FinFET Technology," *2019 Symposium on VLSI Technology*, Kyoto, Japan, 2019, pp. T230-T231, doi: 10.23919/VLSIT.2019.8776570.

[3] Y. Hayakawa et al., "Highly reliable TaOx ReRAM with centralized filament for 28-nm embedded application," *2015 Symposium on VLSI Circuits (VLSI Circuits)*, Kyoto, Japan, 2015, pp. T14-T15, doi: 10.1109/VLSIC.2015.7231381.

[4] S. Ito et al., "ReRAM Technologies for Embedded Memory and Further Applications," *2018 IEEE International Memory Workshop (IMW)*, Kyoto, Japan, 2018, pp. 1-4, doi: 10.1109/IMW.2018.8388846.

[5] G. Molas et al., "Crosspoint Memory Arrays: Principle, Strengths and Challenges," *2020 IEEE International Memory Workshop (IMW)*, Dresden, Germany, 2020, pp. 1-4, doi: 10.1109/IMW48823.2020.9108143.

[6] P. Jain et al., "13.2 A 3.6Mb 10.1Mb/mm2 Embedded Non-Volatile ReRAM Macro in 22nm FinFET Technology with Adaptive Forming/Set/Reset Schemes Yielding Down to 0.5V with Sensing Time of 5ns at 0.7V," *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, San Francisco, CA, USA, 2019, pp. 212-214, doi: 10.1109/ISSCC.2019.8662393.

[7] C. Peters, F. Adler, K. Hofmann and J. Otterstedt, "Reliability of 28nm embedded RRAM for consumer and industrial products," *2022 IEEE International Memory Workshop (IMW)*, Dresden, Germany, 2022, pp. 1-3, doi: 10.1109/IMW52921.2022.9779300.

[8] N. Capron, P. Broqvist, A. Pasquarello, *Applied Physics Letters* 2007, 91, 192905., doi: https://doi.org/10.1063/1.2807282

[9] D. Duncan, B. Magyari-Köpe and Y. Nishi, "Filament-Induced Anisotropic Oxygen Vacancy Diffusion and Charge Trapping Effects in Hafnium Oxide RRAM," in *IEEE Electron Device Letters*, vol. 37, no. 4, pp. 400-403, April 2016, doi: 10.1109/LED.2016.2524450.

[10] A. S. Foster, A. L. Shluger, and R. M. Nieminen, "Mechanism of Interstitial Oxygen Diffusion in Hafnia," *Physical Review Letters*, 2002, doi:https://doi.org/10.1103/PhysRevLett.89.225901

[11] F. M. Puglisi, P. Pavan, A. Padovani, L. Larcher and G. Bersuker, "Random Telegraph Signal noise properties of HfOx RRAM in high resistive state," *2012 Proceedings of the European Solid-State Device Research Conference (ESSDERC),* Bordeaux, France, 2012, pp. 274-277, doi: 10.1109/ESSDERC.2012.6343386.

[12] S. Ambrogio, S. Balatti, A. Cubeta, A. Calderoni, N. Ramaswamy and D. Ielmini, "Understanding switching variability and random telegraph noise in resistive RAM," *2013 IEEE International Electron Devices Meeting (IEDM)*, Washington, DC, USA, 2013, pp. 31.5.1-31.5.4, doi: 10.1109/IEDM.2013.6724732.

[13] J. Sandrini et al., "OxRAM for embedded solutions on advanced node: scaling perspectives considering statistical reliability and design constraints," *2019 IEEE International Electron Devices Meeting (IEDM),* San Francisco, CA, USA, 2019, pp. 30.5.1-30.5.4, doi: 10.1109/IEDM19573.2019.8993484.

[14] L. Reganaz *et al*., "Elucidating Postprogramming Relaxation in Multilevel Cell Resistive Random Access Memory by Means of Experimental and Kinetic Monte Carlo Simulation Data," *Physica Status Solidi (A)*, 2022, doi: https://doi.org/10.1002/pssa.202100753

[15] A. Fantini et al., "Intrinsic program instability in HfO2 RRAM and consequences on program algorithms," *2015 IEEE International Electron Devices Meeting (IEDM),* Washington, DC, USA, 2015, pp. 7.5.1-7.5.4, doi: 10.1109/IEDM.2015.7409648.

[16] J. Lee, *et al.*, "Charge Transition of Oxygen Vacancies during Resistive Switching in Oxide-Based RRAM," *ACS Applied Materials & Interfaces* 2019 11 (12), 11579-11586 DOI: 10.1021/acsami.8b1838