Endurance of 2 Mbit Based BEOL Integrated ReRAM

NILS KOPPERBERG1, STEFAN WIEFELS2, KARL HOFMANN3, JAN OTTERSTEDT3, (Member, IEEE), DIRK J. WOUTERS1, (Member, IEEE), RAINER WASER2,1,4, (Member, IEEE), AND STEPHAN MENZEL2, (Senior Member, IEEE)

1Institut für Werkstoffe der Elektrotechnik II (IWE II) and JARA-FIT, RWTH Aachen University, 52074 Aachen, Germany
2Peter-Grünberg-Institute 7 (PGI-7), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
3Infineon Technologies AG, 85579 Neubiberg, Germany
4Peter-Grünberg-Institute 10 (PGI-10), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Corresponding author: Nils Kopperberg (kopperberg@iwe.rwth-aachen.de)

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) under Grant SFB 917, in part by the German Federal Ministry of Education and Research (BMBF for its initials in German) through the projects “Neuro-inspired Artificial Intelligence Technologies for the Electronics of the Future” (NEUROTEC) under Project 16ME0398K and Project 16ME0399, and “Neuromorphic Hardware for Autonomous Artificial Intelligence Systems” (NeuroSys) under Project 03ZU1106AA and Project 03ZU1106AB and is based on the “Jülich Aachen Research Alliance – Fundamentals of Future Information Technology” (JARA-FIT).

ABSTRACT In this work, we experimentally characterize the endurance of 2 Mbit resistive switching random access memories (ReRAMs) from a 16 MBit test-chip. Here, very rare failure events where the memory cells become stuck in the low-resistive state (LRS) are observed. As this failure mechanism is the limiting one concerning the endurance of this ReRAM implementation, extensive investigations are conducted and presented. The experimental findings are detailed via a voltage divider model, illustrating why memory cells can become stuck in the LRS. It is proposed, that an insufficient voltage dropping over the cell due to an unfavorable combination of cell- and transistor resistances is responsible for stuck-at-LRS bits. Furthermore, predictions for the origin of these suboptimal combinations are given. Additionally, a one-dimensional Kinetic Monte Carlo (KMC) model that allows a statistical investigation of large numbers of cells with regard to rare random events has been developed. Here, our proposed explanation for the observed failure mechanism is fortified by the simulation and evaluation of the switching process of the memory. All simulations are in very good agreement with the experimental data. Finally, based on our findings, we give suggestions for the improvement of switching algorithms.

INDEX TERMS Endurance, KMC, OxRAM, ReRAM, RESET failure, stuck at SET, VCM.

I. INTRODUCTION

Regarding the strongly increasing demand for highly scaled, non-volatile memories in many modern applications like smartphones, the research in this field has grown strongly over the last years [1], [2], [3]. Here, resistive switching random access memories (ReRAMs) are very promising candidates for future industrial applications as well as for neuromorphic computing [4], [5], [6]. ReRAMs show a lot of great operation features like scalability, fast switching, reliability and complementary metal-oxide-semiconductor (CMOS) compatibility [7], [8], [9] and thus are expected to replace flash technology in the future [10], [11], [12].

The associate editor coordinating the review of this manuscript and approving it for publication was Giambattista Gruosso.
SET, the opposite process is called RESET. The present state of the cell can be easily read out non-destructively by applying a small voltage to the device. Independent of the specific application of the ReRAMs, reliability is a major issue that has to be investigated and optimized [18], [19], [20]. Especially, endurance meaning the ability of huge numbers of faultless consecutive switching cycles is of great interest and will therefore be in the focus of our work [21], [22].

In the following, experimental results of 2 MBit VCM cells are presented, predominantly showing great endurance. Nevertheless, looking at the huge statistics with many cells after many switching cycles, a few cells become stuck at the LRS. This rare failure event is the main effect limiting the endurance and is therefore further investigated experimentally. Whereas Yang et al. [23] are connecting this kind of failure mechanism to the presence of a second filament, we present an alternative explanation. In a simple voltage divider model, we propose a too low voltage dropping over the cell to be programmed but also acts as current limiting element, increasing $V_{WL}$ allows for higher current through the device. A good control of this current is crucial during electroforming and SET, as it strongly affects the resulting resistance of the cell [25], [26], [27]. In contrast, $V_{WL}$ is desired to be comparatively high during RESET in order to allow sufficient current for this process. A current limitation is not needed for this operation, because the increasing cell resistance naturally limits the current.

III. EXPERIMENTAL RESULTS

A. ENDURANCE

Fig. 2, a) shows the cumulative HRS and LRS distributions read after different cycle numbers up to 500k. SET and RESET are performed using our standard algorithm which comprises multiple steps with increasing pulse length, $V_{BL}$, $V_{SL}$ and $V_{WL}$. This results in a very good endurance in the investigated interval of 500k cycles. Whereas the LRS distribution only slightly broadens upon cycling, the HRS distribution is observed to drift towards lower read current, which has a positive effect on the read window between HRS and LRS. However, at higher cycle numbers (300k-500k) a tail appears in the HRS distribution consisting of bits which were not RESET successfully but stuck in the LRS. Although this affects only few ppm of bits, this failure has to be investigated, which is the main subject of this study.

B. STUCK-AT-LRS FAILURE

The inset in Fig. 2, a) shows the number of failed bits in the SET (red) or RESET (blue) operation over the cycle number. Up to 250k cycles, both SET and RESET are comparatively stable, although each cycle carries the chance of single failed bits. Whereas the SET operation remains stable within the tested cycles, an increasing trend of failed bits is observed for the RESET operation. Here, the number of fail-bits increases via on-chip circuitry. From this test vehicle, the highlighted block of 2 MBit is cycled.

A schematic of a single 1T-1R cell with the respective biases is shown in Fig. 1, b). The select transistor is opened by applying a voltage $V_{WL}$ to its gate. In order to electroform or SET a bit, a rectangular voltage pulse $V_{BL}$ is applied to the bitline which is connected to the ohmic electrode (OE) of the ReRAM cell. For the RESET operation, a rectangular voltage pulse $V_{SL}$ is applied to the source of the transistor (while its drain is connected to the active electrode (AE) of the ReRAM cell). To ensure reliable programming, electroforming, SET, and RESET are performed with a program-verify algorithm [24]. Here, the resistance of each bit is determined by a read pulse after each programming step and compared to a target resistance. If the target is reached, the programming operation is terminated and considered successful. Otherwise, the programming step is repeated with increased pulse length, $V_{BL}$, $V_{SL}$ or $V_{WL}$. Here, increasing $V_{BL}$ or $V_{SL}$ provides a higher voltage across the ReRAM element during SET or RESET, respectively. As the transistor not only selects the cell to be programmed but also acts as current limiting element, increasing $V_{WL}$ allows for higher current through the device.

II. EXPERIMENTAL DETAILS

Our experimental endurance study was conducted on a test-chip with 16 Mbit VCM-type ReRAM, shown in Fig. 1, a). The cells are integrated back-end-of-line (BEOL) in a 1-transistor-1-resistor (1T1R) configuration in a 28 nm CMOS technology. Programming and read out are performed

![FIGURE 1. a) Image of the 16 MBit test-chip. The cycled block of 2 MBit is highlighted in green. b) Sketch of the typical 1T-1R structure of our VCM cells, with the ReRAM element on the left and the transistor on the right. The AE of the element is connected to the transistor and the OE to the BL. The SET polarity is therefore with a positive voltage $V_{WL}$ at the BL.

N. Kopperberg et al.: Endurance of 2 Mbit Based BEOL Integrated ReRAM

VOLUME 10, 2022

122697
linearly with the cycle number to a maximum of 22 at 516k cycles, which accounts for approx. 10 ppm.

Despite the significance of this failure mechanism, it affects only very few individual ReRAM cells in the shown experiment. Further, no clusters of stuck-LRS bits are observed, but a random distribution of failed bits in the memory block. In order to gain deeper insights of the underlying mechanism, the endurance experiment was repeated with a programming algorithm that was slightly altered in a direction which provokes this failure and thus generates more stuck-at-LRS bits. Cumulative distributions of HRS and LRS after 50k-100k cycles with the suboptimal algorithm are shown in Fig. 2, b). Here, from 50k to 51k the read current of each bit in each cycle was measured directly, resulting in consecutive traces containing the states right after SET or RESET over 1000 cycles.

In order to understand the origin of this failure, exemplary traces out of this data set are shown in Fig. 3. In the five columns, selected bits with different numbers of failed cycles are depicted. The first column (Fig. 3, a, b)) shows reference bits for which no failures have been observed. The remaining traces are sorted by the number of cycles that a bit was stuck in LRS within the monitored 1k cycles. This ranges from only a few cycles in Fig. 3, c), d) to more than 600 cycles in Fig. 3, i), j). The traces show that bits can become stuck in LRS spontaneously as in Fig. 3, c) or the HRS is observed to drift towards LRS from cycle to cycle until it becomes stuck (c.f. Fig. 3, h)). Although cells can become stuck over several cycles (> 1000), it is not primarily a permanent device failure. The single cells rather end up in a state from where the successful RESET becomes more unlikely. Thus, bits are typically observed to recover after a few or several cycles. Other bits seem to continuously alternate between functioning and faulty, as in Fig. 3, g). It may be noted that the cases in the different columns are not equally likely. RESET failures with long lifetime are much less likely than failures over single or very few cycles.

However, two general observations are made: Firstly, stuck bits usually comprise a comparatively high HRS current before failing or after recovery. Secondly, the stuck-at-LRS bits typically exhibit rather high read current in the problematic LRS. Whereas the first observation might be an indicator how bits become stuck, the second observation of low LRS resistances hints towards a possible explanation of the failure mechanism: In the interplay of ReRAM cell and series resistance of the periphery (including the access transistor), a too low cell resistance might lead to an insufficient electrical field across the cell as the major part drops across the periphery. In the following section we discuss this mechanism using a simple model before we compare our experimental findings to the results of KMC simulations.

C. PHENOMENOLOGICAL MODEL

The origin of the RESET failure can be explained by a simple model. In Fig. 1 an extract of a single memory element in a typical 1T-1R configuration was already presented. The transistor shown on the right can be controlled by the word line voltage \( V_{WL} \), the ReRAM can be read or switched by the applied bit line and source line voltages \( V_{BL} \) and \( V_{SL} \). During the RESET, the resistance of the transistor is nearly constant. Therefore, the transistor as well as the line resistances and other parts of the circuitry are summed up to a resistance of the periphery \( R_{per} \). Thus, as can be seen in Fig. 4, a series connection of the ReRAM cell with a cell resistance \( R_{cell} \) and the resistance of the periphery \( R_{per} \) is taken to represent the 1T-1R element. This results in a voltage divider where the totally applied voltage \( V_{tot} \) splits up into the voltage dropping over the periphery \( V_{per} \) and the voltage dropping over the ReRAM cell \( V_{cell} \). This voltage \( V_{cell} \) is crucial with regard to the switching process, as the switching time depends exponentially on it [28]. According to the voltage divider, \( V_{cell} \) can be calculated to

\[
V_{cell} = V_{tot} \frac{R_{cell}}{R_{cell} + R_{per}},
\]

(1)
FIGURE 3. Selected traces of HRS (blue) and LRS (red) over 1000 consecutive switching cycles. The data is drawn from cycles 50001-51000 in Fig. 2, b). The traces are sorted column-wise regarding the number of failed RESETs. a), b) show good reference bits which were not observed to be stuck in this interval. c), d) comprise few failed cycles, which increases towards i) with more than 600 and j) with all 1000 failed cycles. Stuck bits can occur spontaneously or appear with a gradually increasing HRS current. The failure is not permanent, but can be recovered. Stuck bits typically come from higher HRS current and comprise high LRS current when they are stuck.

FIGURE 4. Sketch of our model idea, simplifying the typical 1T-1R structure of our VCM cells from Fig. 1 to a voltage divider of two resistances.

FIGURE 5. Cell voltage in dependence of cell resistance and periphery resistance. As an example, a cell A with $R_{\text{per}} = 3600\,\Omega$ and $R_{\text{cell}} = 3500\,\Omega$ is shown. For cell B, with the same $R_{\text{per}}$, but programmed into a low-ohmic state with $R_{\text{cell}} = 2500\,\Omega$, $V_{\text{cell}}$ is lowered by approx. 0.2 V.

which therefore depends on three parameters: The first parameter is the externally applied voltage $V_{\text{tot}}$, which is in good approximation constant for all times and all devices. The second parameter is the resistance of the periphery $R_{\text{per}}$, which is constant over time for a single device, but has a certain variance from device to device. The last parameter is the resistance of the cell $R_{\text{cell}}$ that varies during switching and is different for each device after every switching cycle. Especially, $R_{\text{cell}}$ in the LRS is strongly depending on the preceding SET event. The influence of both varying parameters, $R_{\text{cell}}$ and $R_{\text{per}}$, on $V_{\text{cell}}$ is shown in Fig. 5.

It can be seen that, depending on the possible cell and periphery resistances, $V_{\text{cell}}$ can vary significantly. Realistic values were chosen for the shown ranges of $R_{\text{cell}}$ and $R_{\text{per}}$. It is observed that $V_{\text{cell}}$ can lay in a range from approximately 0.9...1.4 V. Exemplarily, a cell with $R_{\text{per}} = 3600\,\Omega$ is marked for a common cell resistance of $R_{\text{cell}} = 3500\,\Omega$ (A) and a low-ohmic state with $R_{\text{cell}} = 2500\,\Omega$ (B) in Fig. 5. Between these two states, a change in $V_{\text{cell}}$ of about 0.2 V can be observed. Since it has been demonstrated that a decrease of $V_{\text{cell}}$ by 0.1 V can increase the RESET time by one order of magnitude [28], this huge variance in $V_{\text{cell}}$ leads to a strongly varying switching behavior from cell to cell and from cycle to cycle.

Most of the cases are not problematic, as the RESET conditions are nicely adjusted. But in rare cases of a very high periphery resistance $R_{\text{per}}$ in combination with a very low cell resistance $R_{\text{cell}}$, the voltage dropping over the cell $V_{\text{cell}}$ is too low to drive the RESET process within the applied pulse width. This failure process becomes even rarer, as programming algorithms increasing the applied voltage $V_{\text{tot}}$ or the time of the switching pulse are able to switch devices that failed or only partly switched in the first try. Only for very extreme configurations, the devices are stuck in the LRS.

During the SET operation, the periphery resistance determines to a large extent in which resistance the cell ends
up. The resulting LRS is expected to be in the range of $R_{\text{cell}} \approx R_{\text{per}}$ as the applied voltage would drop mainly across the periphery resistance as soon as the cell resistance becomes significantly smaller. For RESET, the transistor is opened further to ensure having a sufficiently large voltage dropping over the cell. However, the body effect has to be considered. Since the access transistor needs to be operated in two polarities (for SET and RESET), its bulk and source cannot be connected. In SET direction, the voltage is applied at the BL with the transistor source connected to ground. However, during RESET the effective source of the transistor is located at the middle node in Fig. 1, a), between the active electrode of the ReRAM element and the transistor channel. This reduces the effective voltage between gate and source and thus increases the resistance of the channel. As this results in a higher periphery resistance during RESET compared to SET, the body effect strengthens our prediction for the cause of the RESET failure.

### D. GENERATION

After the reason for the RESET failure is established, the question arises how the unfortunate combination of low cell resistance and high periphery resistance is generated. One reason is the body effect, as discussed above. During SET, the cell is programmed to a state determined by the series resistance of the periphery. If this resistance increases upon attempted RESET, this process is already more challenging. As another possible origin of the RESET failure, the formation of an additional filament was reported [23]. However, both explanations do not account for the increasing number of failed bits with higher cycle numbers (c.f. inset in Fig. 2, a)). In order to understand the generation of stuck bits, the read currents of the last successful SET and RESET in our experimental data are traced and compared to the probability to become stuck in the following cycle. As depicted in Fig. 6, a) this probability increases for cells with a high read current in HRS, close to the read window. This fits to the qualitative observation in Fig. 3 that problematic bits are often accompanied by high HRS current before becoming stuck or after recovery. Additionally, Fig. 6, a) shows that an LRS read current close to the read window indicates a higher probability to become stuck in the next cycle. This is counter-intuitive, as bits are typically observed to become stuck in a high current LRS. However, note that the figure only contains the last good cycle. Bits that are already in the problematic high current state would not appear here. In both cases (HRS and LRS), the bits close to the read window are likely to become stuck afterwards. This hints towards the applied programming algorithm as a possible cause of bits becoming stuck, as the decision if further programming pulses are executed depends on a threshold current in the read window. It seems likely that in particular those bits that hardly reach the specified threshold endure further attempts by the programming algorithm, where the later steps typically comprise longer pulse width, $V_{\text{BL}}$ or $V_{\text{WL}}$.

To understand the impact of the programming algorithm, an exemplary algorithm for the SET operation is studied experimentally in Fig. 6, b). The algorithm consists of four steps with increasing $V_{\text{WL}}$ and thus increasing maximum SET current. Between the steps, all bits are read and the number of bits which failed to reach the threshold current of a successful SET are evaluated. As expected, it can be seen that this number decreases with each executed step. However, the study also shows that the number of fail-bits increases over the number of cycles. Note that here, fail-bits are those which did not SET successfully. This means that the SET operation becomes more difficult at higher cycle numbers. In order to reach a fixed threshold current, the algorithm would therefore more often execute the later steps with higher $V_{\text{WL}}$, which increases the likelihood to end up in the unfortunate combination of $R_{\text{cell}} < R_{\text{per}}$. Especially, an increasing $V_{\text{WL}}$ seems to be an operation to more likely generate these bad combinations. It is conceivable that over several steps the limiting factor is the switching kinetics and thus the pulse length and height ($V_{\text{BL}}$). If, nevertheless, the gate voltage $V_{\text{WL}}$ is increased, the cell might endure a comparatively high SET current as soon as the switching event occurs.

### IV. 1D KMC SIMULATION

To further support our theory and to investigate the origin and behavior of the presented RESET failure on another level, a one-dimensional kinetic Monte Carlo (KMC) simulation
The physics-based compact model is based on several general assumptions. The model assumes a preexisting filament connecting the two electrodes, which shows a high number of oxygen vacancies. The filament has a time-invariant radius \( r \) and is divided into a plug and a disc region with a uniform oxygen vacancy concentration \( N_{\text{disc}} \) and \( N_{\text{plug}} \) each. The lengths of both regions \( l_{\text{disc}} \) and \( l_{\text{plug}} \) are also constant, whereas the oxygen vacancy concentrations in the regions are the only state variables changing over time. The total number of oxygen vacancies in the cell \( N_{\text{cell}} = N_{\text{disc}} + N_{\text{plug}} \) is constant. Oxygen vacancies are treated as doubly positive charged donors that can be moved via drift in an electric field enhanced by temperature, which is assumed constant over the whole filament. Diffusion and thermodiffusion are neglected. Additionally, the metal/oxide interfaces are modeled as Schottky contact at the active top electrode and ohmic contact at the bottom electrode. An equivalent circuit diagram of the model is presented in Fig. 7. Here, the five elements of our model can be seen, namely the Schottky diode, the disc and plug resistances \( R_{\text{disc}} \) and \( R_{\text{plug}} \), the series resistance \( R_{\text{ser}} \) at the ohmic electrode and the periphery resistance \( R_{\text{per}} \) including the transistor. To make the connection to our phenomenological model that was introduced before clear, the first four elements have been summed up to the total cell resistance \( R_{\text{cell}} \) in Fig. 7. As a central part of the model, Kirchhoff’s law

\[
V_{\text{tot}} - [V_{\text{Schottky}} + I \cdot (R_{\text{disc}} + R_{\text{plug}} + R_{\text{ser}} + R_{\text{per}})] = 0
\]

is solved, with the current being denoted by \( I \). The calculation of \( R_{\text{disc}} \) and \( R_{\text{plug}} \) is based on band conduction mechanism and temperature-dependent mobility via

\[
R_{\text{disc/plug}} = \frac{l_{\text{disc/plug}}}{A \cdot \gamma \nu_0 e N_{\text{disc/plug}} \mu_{\text{d0}}} \exp \left( \frac{\Delta E_{\text{ac}}}{k_B T} \right),
\]

where \( \nu_0 \) denotes the characteristic vibration frequency. For the hopping barrier \( \Delta W_A \), a typical value for the present metal-oxide is used and modulated by the Genreith-Schrieber approach as predicted in [29] and [32] via

\[
\Delta W_{\text{eff}}^{\text{h}} = \Delta W_A \cdot \left( \sqrt{1 - \gamma^2} + \gamma \right)
\]

depending on the jump direction. The factor \( \gamma \) modifies the hopping barrier according to

\[
\gamma = \frac{a z V_0 e E}{\pi \Delta W_A}
\]

with the hopping distance \( a \) and the electric field which is calculated by

\[
E = \frac{V_{\text{disc}} + V_{\text{plug}}}{l_{\text{cell}}}
\]

for the RESET. Then, randomly but weighted by the probabilities, one of the two processes is chosen. The time of the process is calculated as

\[
t_{\text{jump}} = - \frac{\ln \frac{r_1}{R^2 + R^2}}{R^2 + R^2}
\]
with $r_1$ being a random number between 0 and 1. Finally, the simulation time is increased by $t_{jump}$ and the chosen process is executed by updating $N_{disc/plug}$, which directly influences $R_{disc/plug}$ and so the current through the ReRAM cell. The simulation parameters are given in Table 1.

| Symbol    | Value       | Symbol    | Value       |
|-----------|-------------|-----------|-------------|
| $l_{cell}$ | 5 nm        | $\Delta W_A$ | 1.2 eV     |
| $l_{disc}$ | 0.75 nm     | $\nu_0$    | $2 \times 10^{13}$ Hz |
| $l_{plug}$ | 4.25 nm     | $\nu_{Vo}$ | +2          |
| $r$       | 30 nm       | $\mu_{Vo}$ | $5 \times 10^{-6}$ eV |
| $A$       | $\pi \cdot 900 \text{ nm}^2$ | $\Delta E_{ac}$ | 0.08 m$^2$ V$^{-1}$ s$^{-1}$ |
| $N_{cell}$ | 8000        | $R_{pl,eff}$ | 720 $\Omega$ |
| $a$       | 0.25 nm     | $R_{ser}$  | $4.24 \times 10^{-6}$ K$^{-1}$ |
| $T_0$     | 293 K       |           |             |

### B. RESULTS

In the first step, the known parameters $R_{ser}$, $l_{cell}$ as well as several material parameters, are implemented. Furthermore, the unknown values of the filament geometry $l_{disc}$, $l_{plug}$ and the radius $r$ are chosen in a reasonable way. The number of oxygen vacancies in the plug region $N_{plug}$ is chosen with regard to typical oxygen vacancy densities used in simulations for VCM cells [30], [33]. The key parameter is the number of oxygen vacancies in the disc region $N_{disc}$ and is chosen quite high in view of the LRS being investigated in the beginning. Additionally, for the values of $R_{per}$ a variability known from the experiment and for $N_{plug}$ and $N_{disc}$ an estimated variability ($\sigma \approx 25$) are added as the VCM cells typically show a cell to cell and a cycle to cycle variability leading to a certain width of the current or resistance distributions. The RESET is performed by applying a voltage $V_{tot}$ to the active electrode. Here a pulse sequence of a short read pulse with a low voltage of $V_{read} = -0.2$ V, a longer RESET pulse of $V_{tot} = 2.4$ V and again a short read pulse is applied. The initial LRS current distribution and the HRS current distribution after RESET are presented as dark blue curves of Fig. 8 and fit nicely to the experimental data. We are interested in very rare events that only occur, looking at very high numbers of cells and large numbers of switching cycles in the experiment. Due to the high amount of computing resources needed to simulate even more cells, this method would be very inefficient. Furthermore, increasing the variability of several simulation parameters is not reasonable, as the current distributions of LRS and HRS are very stable in total except the single cells becoming stuck at high currents. From our simple voltage-divider model above, the RESET failure is assumed to occur for bad combinations of $R_{per}$ and $R_{cell}$. As the properties of $R_{per}$ are well known from the experiment and expected to be stable for a single cell, $R_{cell}$ has to be changed in our simulations. Here, the parameter of choice is the number of oxygen vacancies in the disc region $N_{disc}$. As can be seen in the color gradient of Fig. 8, $N_{disc}$ has been increased continuously and concomitant with that $R_{cell}$ has been reduced. Thereby, the probability to observe bad combinations of $R_{per}$ and $R_{cell}$ is permanently increased.

As a result, more and more bits can be observed that are not completely switched to the HRS in the light blue and green curves of Fig. 8. In the bottom right corner of the orange and red curves, even many cells appear that did not or just slightly changed their resistance after the RESET pulse.

In Fig. 9, the read current of all cells in the LRS before the RESET pulse $I_{LRS}$ and after the RESET pulse $I_{RESET}$ are presented. It can be seen that the cell with lower $I_{LRS}$ can be easily switched to the HRS. At higher $I_{LRS}$, many cells can be observed that do not or only partly switch towards the HRS. From the color of the data points, $N_{disc}$ can be read out similar to the color gradient in Fig. 8 before. Here again, it is visible that the cells with high $N_{disc}$ are much more prone to the RESET failure.

In a second step, we have a closer look at the properties of the cells that have not or only been partly switched from the
LRS towards the HRS. In Fig. 10, the probability of a cell to not switch properly is presented in dependence on $R_{\text{per}}$ and $N_{\text{disc}}$. $N_{\text{disc}}$ is strongly correlated to the resistance of the cell $R_{\text{cell}}$, leading to a lower $R_{\text{cell}}$ the higher $N_{\text{disc}}$ is. In Fig. 10, a), all cells that did not switch properly to the HRS are presented. It can be seen that at least a combination of a medium $N_{\text{disc}}$ and a high $R_{\text{per}}$ or a high $N_{\text{disc}}$ and a medium $R_{\text{per}}$ is needed to obtain the failure mechanism. In Fig. 10, b) only cells that are stuck very deep in the LRS above a high current level are shown. To obtain these bits that do not or hardly switch, a combination of high $N_{\text{disc}}$ and high or very high $R_{\text{per}}$ is necessary, which fits well to our simple model idea from above. In the next step, we want to have a closer look at typical cells that only partly switched and typical cells that did not switch under regular conditions. Hence, cells with a medium $N_{\text{disc}}$ and a medium to high $R_{\text{per}}$ on the one hand (cell-type A) and cells with a high $N_{\text{disc}}$ and a very high $R_{\text{per}}$ on the other hand (cell-type B) are taken. Now, additionally to the regular switching conditions, the switching behavior of the cells under longer switching pulses or switching pulses with higher $V_{\text{tot}}$ are investigated. Thus, both cell-types are pulsed with a second pulse after the first normal RESET pulse. The second pulse either has an increased pulse time or an increased applied voltage $V_{\text{tot}} = 2.6$ V. The current distributions of the initial LRS distributions as well as the current distributions after the normal and the second longer or stronger RESET pulse are presented in Fig. 11. As expected, it can be seen that the cells of type A (red) can partly be switched towards HRS, whereas the cells of type B (blue) are nearly completely stuck in the LRS after a normal pulse (solid line). With a second, longer pulse (dashed line), the majority of the cells of type A can be switched to the HRS, whereas the cells of type B remain stuck in the LRS. Alternatively, with a second stronger pulse, nearly all cells of type A can be switched to the HRS and even some cells of type B switch towards the HRS. On the one hand, this explains why in the experiment as in Fig. 2, a) nearly no cells occur, that only partly switched towards the HRS. The cells, that initially only partly switch towards the HRS, can be completely switched via additional pulses by a programming algorithm. The cells, that cannot even be switched via additional pulses, are typically stuck at a deep LRS state. On the other hand, the possibility to recover many of the cells initially stuck at the LRS via programming algorithms explains the rare occurrence of this failure mechanism in the experiment. Furthermore, it can be mentioned that the increase of $V_{\text{tot}}$ during the programming algorithm seems to be more effective than increasing the pulse time.

Although we are able to reproduce and explain most of the experimentally observed properties of the stuck-at-LRS failure mechanism, our phenomenological model and our 1D KMC model are limited. So far, our models cannot thoroughly explain the increase of the read current of single cells during RESET when becoming stuck. In the 1D KMC model, an increase of the current can indeed be observed due to single oxygen vacancies jumping randomly from the plug...
to the disc region during the RESET. But, in comparison with the experiment, this increase in current is too low. Thus, we propose to consider and investigate additional effects like diffusion, thermodiffusion or an additional oxygen exchange at the AE in future works [20], [28], [34]. Furthermore, in our model, so far, we assumed $N_{cell}$ to be constant. Nevertheless, e.g., due to oxygen exchange at the interface of the OE and the oxide or thermodiffusion, $N_{cell}$ could increase during cycling. This comes along with an increased probability of reaching unfortunate combinations of $R_{cell}$ and $R_{per}$. Once reaching such a bad combination with high currents, thermodiffusion or oxygen exchange are even more encouraged. This would explain why the observed failure events increase over time in the experiment.

V. CONCLUSION

In our work, we presented the generally great endurance of 2 MBit ReRAM based on VCM in a 1T-1R configuration in 28 nm CMOS technology. Nevertheless, looking at high cell statistics and high switching cycle numbers, a rare failure mechanism is observed after RESET. As the failure events increase over the cycle number, we investigated them in more detail. Hence, a phenomenological model simplifying the ReRAM structure into a voltage divider consisting of a cell and a periphery resistance was introduced. With this simple model, the RESET failure was proposed to occur for very rare and unlucky combinations of low cell and high periphery resistances. In the next step, we dealt with the question of how this unlucky combinations turn up and why they increase over time. Here, the SET event and the dedicated program verify algorithm play an important role. In the last step, a 1D KMC model was introduced. The high performance of compact modelling was combined with the ability to investigate random processes by the integration of KMC methods. This allowed us to simulate the RESET process with huge statistics and with special regard to rare processes. In our simulation, the RESET process could be nicely reproduced, and the failure events could be provoked by increasing the number of oxygen vacancies in the disc region. We saw that, as we assumed, cells with a bad combination of cell and periphery resistance are prone to this failure mechanism. Finally, we looked at the impact of possible programming algorithms and the outcome of longer or stronger RESET pulses. These programming algorithms have to be further optimized to prevent permanent RESET failures of the cells.

REFERENCES

[1] R. Dittmann, S. Menzel, and R. Was er, “Nanoionic memristive phenomena in metal oxides: The valence change mechanism,” Adv. Phys., vol. 70, no. 2, pp. 155–349, Apr. 2021.
[2] Y. Chen, “ReRAM: History, status, and future,” IEEE Trans. Electron Devices, vol. 67, no. 4, pp. 1420–1433, Apr. 2020.
[3] Z. Wang, H. Wu, G. W. Burr, C. S. Hwang, K. L. Wang, Q. Xia, and J. J. Yang, “Resistive switching materials for information processing,” Nature Rev. Mater., vol. 5, no. 3, pp. 173–195, Mar. 2020.
[4] D. V. Christensen et al., “2022 roadmap on neuromorphic computing and engineering,” Neuromorphic Comput. Eng., vol. 2, no. 2, 2022, Art. no. 022501.
[5] S. Aldana, P. García-Fernández, R. Romero-Zaliz, M. B. González, F. Jiménez-Molinos, F. Gómez-Campos, F. Campabadal, and J. B. Roldán, “Resistive switching in HfO$_2$ based valence change memories, a comprehensive 3D kinetic Monte Carlo approach,” J. Phys. D, Appl. Phys., vol. 53, no. 22, May 2020, Art. no. 225106.
[6] G. W. Burr, R. M. Shelby, A. Sebastian, S. Kim, S. Sidler, K. Viriyani, M. Ishi, P. Nazaryan, A. Fumarrala, L. L. Sanche s, I. Boybat, M. L. Gallo, K. Moon, J. Wuo, H. Hwang, and Y. Leblebici, “Neuromorphic computing using non-volatile memory,” Adv. Phys. X, vol. 2, no. 1, pp. 89–124, 2017.
[7] R. Was er, Nanotechnology: Information Technology, vol. 3. Weinheim, Germany: Wiley, 2008.
[8] C.-F. Yang, C.-Y. Wu, M.-H. Yang, W. Wang, M.-T. Yang, T.-C. Chien, V. Fan, S.-C. Tsai, Y.-H. Lee, W.-T. Chu, and A. Hung, “Industrial applicability: Read disturb model and performance on mega-bit 28 nm embedded ReRAM,” in Proc. IEEE Symp. VLSI Technol., Jun. 2020, pp. 1–2.
[9] F. Pan, S. Gao, C. Chen, C. Song, and F. Zeng, “Recent progress in resistive random access memories: Switching mechanisms, and performance,” Mater. Sci. Eng. R, Rep., vol. 83, pp. 1–59, Sep. 2014.
[10] D. S. Jeong, R. Thomas, R. S. Katiyar, J. F. Scott, H. Kohlstedt, A. Petrun, and C. S. Hwang, “Emerging memories: Resistive switching mechanisms and relevant status,” IEEE Prog. Quantum Electron., vol. 75, no. 7, 2012, Art. no. 76502.
[11] M. Kim, D. S. Jeong, and C. S. Hwang, “Nonaffinitative resistive switching in binary oxide system: A review on the present status and outlook,” Nanotechnology, vol. 22, no. 25, 2011, Art. no. 254002.
[12] S. Slezacek and T. Mikolajik, “Nanoscale resistive switching memory devices: A review,” Nanotechnology, vol. 30, no. 35, Aug. 2019, Art. no. 352003.
[13] R. Was er, R. Bruchhaus, and S. Menzel, “Redox-based resistive switching memories,” in Nanoelectronics and Information Technology, 3rd ed. Hoboken, NJ, USA: Wiley, 2012, pp. 683–710.
[14] S. Fukuyama, K. Maeda, S. Matsuda, K. Takeuchi, and R. Yasuhara, “Suppression of endurance-stressed data-retention failures of 40 nm TaO$_x$-based ReRAM,” in Proc. IEEE Int. Rel. Phys. Symp. (IRPS), Mar. 2018, pp. P-MY-4.1–P-MY-4.5.
[15] L. Goux, X. P. Wang, Y. Y. Chen, L. Pantisano, N. Jossart, B. Govoreanu, J. A. Kittl, M. Jurczak, L. Altimime, and D. J. Wouters, “Roles and effects of TiN and Pt electrodes in resistive-switching HfO$_2$ systems,” Electrochem. Solid-State Lett., vol. 14, no. 6, p. H244, 2011.
[16] S. Menzel and J.-H. Hur, Modeling the VCM- and ECM-Type Switching Kinetics. Hoboken, NJ, USA: Wiley, 2016, pp. 395–436.
[17] D. Cooper, C. Baeumer, N. Bernier, A. Marchewka, C. La Torre, R. E. Dunin-Borkowski, S. Menzel, R. Was er, and R. Dittmann, “Anomalous resistance hysteresis in oxide ReRAM: Oxygen evolution and reincorporation revealed by in situ TEM,” Adv. Mater., vol. 29, no. 23, Jun. 2017, Art. no. 1700212.
[18] A. Fantini, G. Giori ne, R. Degraeve, L. Goux, C. Y. Chen, A. Redolfi, S. Clima, A. Cabrini, G. Torelli, and M. Jurczak, “Intrinsic program instability in HfO$_2$ RR AM and consequences on program algorithms,” in IEDM Tech. Dig., Dec. 2015, pp. 7.5.1–7.5.4.
[19] F. Puglisi, L. Larcher, A. Padovani, and P. Pavan, “A complete statistical investigation of RTN in HfO$_2$-based RR AM in high resistive state,” IEEE Trans. Electron Devices, vol. 62, no. 8, pp. 2600–2613, Aug. 2015.
[20] N. Kopperberg, S. Wie fels, S. Liberta, R. Was er, and S. Menzel, “A consistent model for short-term instability and long-term retention in filamentary oxide-based memristive devices,” ACS Appl. Mater. Interfaces, vol. 13, no. 48, pp. 58066–58075, 2021.
[21] M. Lanza et al., “Standards for the characterization of endurance in resistive switching devices,” ACS Nano, vol. 15, no. 11, pp. 17214–17231, Nov. 2021.
[22] S. Fukuyama, A. Hayakawa, R. Yasuhara, S. Matsuda, H. Kinoshita, and K. Takeuchi, “Comprehensive analysis of data-retention and endurance trade-off of 40 nm TaO$_x$-based ReRAM,” in Proc. IEEE Int. Rel. Phys. Symp. (IRPS), Mar. 2019, pp. 1–6.
[23] C.-F. Yang, C.-Y. Wu, M.-C. Shih, M.-T. Yang, M.-H. Yang, Y.-T. Wu, T.-C. Chien, C.-W. Lai, S.-C. Tsai, W.-T. Chu, and A. Hung, “Demonstration of high endurance capability on mega-bit RR AM macro and model of ppm level failures,” in Proc. IEEE Symp. VLSI Technol. Circuits, Jun. 2022, pp. 318–319.
[24] C. Peters, F. Adler, K. Hofmann, and J. Otterstedt, “Reliability of 28 nm embedded RR AM for consumer and industrial products,” in Proc. IEEE Int. Memory Workshop (IMW), May 2022, pp. 1–3.
D. Strukov, F. Alibart, and R. Williams, “Thermophoresis/diffusion as a plausible mechanism for unipolar resistive switching in metal-oxide-metal memories,” J. Appl. Phys., vol. 107, no. 3, pp. 509–518, 2012.

H.-C. Tseng, T.-C. Chang, Y.-C. Wu, S.-W. Wu, J.-J. Huang, Y.-T. Chen, J.-B. Yang, T.-P. Lin, S. M. Sze, M.-J. Tsai, Y.-L. Wang, and A.-K. Chu, “Impact of electroforming current on self-compliance resistive switching in an ITO/Gd:SiO\textsubscript{x}/TiN structure,” IEEE Electron Device Lett., vol. 34, no. 7, pp. 858–860, Jul. 2013.

C. Bengel, A. Kindsmüller, P. Stasner, F. Berg, F. Cüppers, S. Hoffmann-Eifert, R. Waser, S. Menzel, and U. Böttger, “Intrinsic RESET speed limit of valence change memories,” J. Appl. Electron. Mater., vol. 3, no. 12, pp. 5563–5572, Dec. 2021.

C. L. Torre, “Physics-based compact modeling of valence-change-based resistive switching devices,” Ph.D. dissertation, Dept. Elect. Eng. Inf. Technol., RWTH Aachen Univ., Aachen, Germany, 2019.

C. L. Torre, A. F. Zurhelle, and S. Menzel, “Compact modelling of resistive switching devices based on the valence change mechanism,” in Proc. Int. Conf. Simulation Semiconductor Processes Devices (SISPAD), Sep. 2019, pp. 1–4.

S. Menzel, I. Valov, R. Waser, B. Wolf, S. Tappertzhofen, and U. Böttger, “Statistical modeling of electrochemical metallization memory cells,” in Proc. IEEE 6th Int. Memory Workshop (IMW), Taiwan, May 2014, pp. 1–4.

A. R. Genix-Schriever and R. A. De Souza, “Field-enhanced ion transport in solids: Reexamination with molecular dynamics simulations,” Phys. Rev. B, Condens. Matter, vol. 94, Dec. 2016, Art. no. 224304.

C. Bengel, A. Siemon, V. Rana, and S. Menzel, “Implementation of multinary lukasiewicz logic using memristive devices,” in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), May 2021, pp. 1–5.

D. Strukov, F. Alibart, and R. Williams, “Thermophoresis/diffusion as a plausible mechanism for unipolar resistive switching in metal-oxide-metal memristors,” J. Appl. Phys. A. Solids Surf., vol. 107, no. 3, pp. 509–518, 2012.

KARL HOFMANN was born in Mannheim, Germany. He received the Ph.D. degree in physics from RWTH Aachen, in 2001. He joined Infineon Technologies as a Development Engineer. He has published several papers on device reliability. He is currently focusing on the development of emerging non-volatile memories. He is a member of the Technical Committee of the International Memory Workshop (IMW).

JAN OTTERSTEDT (Member, IEEE) was born in Bremen, Germany. He received the Dr.-Ing. degree in electrical engineering from the University of Hannover, Germany, in 1997. Afterwards, he has joined the Semiconductor Group of Siemens, which later became the Infineon Technology AG. Since more than 15 years, he is responsible for concept engineering for embedded non-volatile memories. Since 2006, he Lectures on “Testing Digital Circuits” at the Technische Universität München (TUM).

DIRK J. WOUTERS (Member, IEEE) received the master’s and Ph.D. degrees in electrical engineering from the University of Leuven, Leuven, Belgium, in 1982 and 1989, respectively. In 2014, he joined the Institute of Electronic Materials, RWTH Aachen University, Aachen, Germany, where he focused on the research of metal-oxide-based RRAM.

RAINER WASER (Member, IEEE) received the Ph.D. degree in physical chemistry from the University of Darmstadt, Darmstadt, Germany, in 1984. In 1992, he joined the Faculty of Electrical Engineering and Information Technology, RWTH Aachen University, Aachen, Germany, as a Professor. In 2007, he was the Director of the Institute of Solid State Research, Forschungszentrum Jülich, Jülich, Germany. He was a recipient of the prestigious Gottfried Wilhelm Leibniz Preis in 2014.

STEPHAN MENZEL (Senior Member, IEEE) was born in Bremen, Germany. He received the Diploma degree and the Ph.D. degree (summa cum laude) in electrical engineering from RWTH Aachen University, Aachen, Germany, in 2005 and 2012, respectively. He is currently a Senior Researcher with the Peter-Grünberg-Institut 7, Forschungszentrum Jülich GmbH, Jülich, Germany, where he is also leading the Simulation Group.

...