RRAM Variability and its Mitigation Schemes

Peyman Pouyan, Esteve Amat*, Said Hamdioui, Antonio Rubio*

Department of Quantum Engineering TU Delft Netherlands
*Department of Electronic Engineering UPC, Barcelona Spain
p.pouyan@tudelft.nl

Abstract—**Emerging technologies such as RRAMs are attracting significant attention due to their tempting characteristics such as high scalability, CMOS compatibility and non-volatility to replace the current conventional memories. However, critical causes of hardware reliability failures, such as process variation due to their nano-scale structure have gained considerable importance for acceptable memory yields. Such vulnerabilities make it essential to investigate new robust design strategies at the circuit system level. In this paper we have analyzed the RRAM variability phenomenon, its impact and variation tolerant techniques at the circuit level. Finally a variation-monitoring circuit is presented that discerns the reliable memory cells affected by process variability.**

Index Terms—**RRAM; Reliability; Process Variability; Mitagation; Emerging Memory; Resistive Memory**

I. INTRODUCTION

As Moore's law approaching its ultimate stages, the research for alternative memory technologies has been accelerated to explore the best candidates for future memory systems. One of these candidates is the resistive memory, including different device types, classified by their switching mechanisms. Among them, the two terminal devices that hold the information in the resistance between terminals, such as Phase Change RAM (PCRAM), Spin Transfer Torque RAM (STT-RAM) and Redox Memory (RRAM) are the most important ones.

In PCRAM the resistivity can switch between high and low states because of resistance difference between the amorphous and the crystalline states of chalcogenide glass [1]. STT-RAM can have high or low resistance states in correspondence to its magnetic tunnel junction [2]. RRAM functions based on the change of resistance value on a Metal-Insulator-Metal (MIM) structure [3]. Its resistance varies because of ion migration inside the structure along with some redox processes including electrode/insulator material.

The RRAM has multiple advantages over STT-MRAM and PCRAM such as: 1-It can be better scaled (up to few nm) because of its ion based switching mechanism [3], 2-It is possible to make a memory cell with a single RRAM while the other two need a select device [4], 3-It is compatible with monolithic 3D integration and crossbar architecture to make high density memory [5], 4-It has better endurance than PCM and larger on/off ratio than STT-MRAM [6].

The RRAM itself has been categorized to 4 types based on its switching mechanism [6]: 1-Electrochemical Metallization Bridge Memory (EMB), 2- Metal Oxide Bipolar Filamentary 3- Metal Oxide Unipolar Filamentary 4- Metal Oxide Bipolar Non-Filamentary. Among them, type two has shown promising characteristics and is the most popular RRAM in research and commercial prototyping, chosen for study in this paper as well. However, in spite of all interesting RRAM features, various reliability concerns such as the process variability are reported in literature.

With the device dimensions scaled down to the nanoscale regime, the device variability has taken a crucial relevance into the devices behavior, and consequently on memory reliability and performance. These process fluctuations has become also relevant in RRAM [7], as it is designed in nano-scale sizes for high density embedded memory applications [8]. Therefore, this challenge has motivated the develop of approaches and techniques at various system levels, such as improving the device material, resilient circuits and architectures to somehow avoid or mitigate the variability effect. While improving the device variability is mostly a manufacturing matter, the key tool for improving the yield is considering variation-tolerant circuits, which is studied in this paper.

This paper is organized as following: Section II studies the RRAM variability and its causes; Section III presents some of the techniques to mitigate the variability at circuit level; in Section IV we demonstrate a technique to monitor process variability in RRAM memory; and finally Section V concludes the paper.

II. RRAM VARIABILITY

One of the major sources of unreliability in RRAM memory design is existence of resistance fluctuations in its nominal Low Resistance State (LRS) and High Resistance State (HRS) values [7]. This phenomenon can affect the robustness of the memory operation and reduces the yield. RRAM devices suffer from two types of variation: 1-cycle-to-cycle and 2-device-to-device. In this paper we will only consider the device-to-device variability phenomenon, firstly because it has wider distribution than cycle-

to-cycle deviation and second because unlike cycle-to-cycle which is more related to the device stability and its constructing material the device-to-device variation characterizes the uniformity of memory array and can be treated/mitigated at circuit level.

Figure 1.a shows an example for RRAM device-to-device variability[9] at its LRS and HRS, and Figure 1.b shows the fluctuation of them in resistance switching.
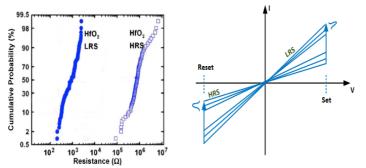

Figure 1. a)Statistical distribution of resistances in HRS and LRS for 100 sweep cycles adapted from [9], b)RRAM I-V characteristic with resistance variability switching

In order to better understand the origins behind the RRAM variability it is worth to consider the RRAM switching mechanism. The RRAM switching mechanism is based on Conductive Filamentary (CF) mechanism where the existing defects (such as oxygen vacancies) result in formation and disruption of CF and change of resistance from LRS to HRS and vice versa [10]. In these devices first an electroforming process with a voltage higher than the normal operation is applied to the device to construct an initial filament between top and bottom electrodes. The first formed filament does not connect the top and bottom electrodes and the device is in its HRS infant mode. To switch the device to LRS, a positive voltage is applied to the top electrode. This will pull out the oxygen ions from the lattice and generate oxygen vacancies which will extend the filament and switch the RRAM from HRS to LRS (SET). To switch the RRAM from LRS to HRS (RESET) a reverse voltage is applied which will bring back the oxygen ions to the lattice and then they will re-combine with the oxygen vacancies and rupture the filament toward HRS.

All these processes of electroforming, SET and RESET if not controlled well can cause considerable variability. Now here we will explain the main sources of fluctuations in RRAMs.
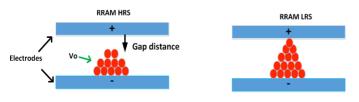
A. Sources of Variability

1-Electroforming

In the forming process generally a quasi-DC sweep on the bit-line up to 3.5V with a step voltage equal to 0.025V is applied. Special current control of the device is needed (usually by a select device) to avoid hard breakdown. However there always exist some deviation in the applied forming voltage that will consequently cause distributions of RRAM'S LRS and HRS [11].

2-Flutuation in CF radius or number of CFs

The stochastic nature of ion migration will cause different CF shapes, sizes (specifically in the radius) and number of CFs [12]. As RRAM resistance value strongly depends on the CF structure, any non-uniformity of CF will be the origin of variability.

3-Tunneling gap distance variation

The distance between the created filament and the top electrode is called the gap distance. Any variation in this parameter can affect the tunneling current and would impose resistance variation in RRAM (mainly HRS variation) [12]. It is noted that the HRS variation (ranging from 5x to 100x) is generally larger than the LRS variation (ranging from 2x to 10x) because of its exponential dependence of the tunneling current affected by gap distance [3]. Figure 2 shows an example for CF in HRS and LRS mode.

Figure 2. RRAM filament in LRS and HRS

Beside these intrinsic RRAM variability sources, the operating parameters can also affect the device variability. For instance lower V_{SET} value and pulse width will increase the LRS variation. Increasing temperature also results in higher resistance variation mainly because of retention loss [7]. Furthermore aging will also deteriorate variation during the cycles, making it essential to take into account the time-dependent parameters in device variability.

The RRAM variability imposes severe concerns into the reliable operation of the memory array and next section presents a brief impact of resistance variation on memory operation.

B. Impact of RRAM variability

The correct operation of memory depends on the precise Read/Write process, nevertheless the existing variability can disturb them and reduce the memory yield [13].

One of the techniques for writing the RRAM devices is called the preset writing in which the selected RRAM is first reset to the LRS state and then by applying an appropriate pulse it is switched to the desired state [14]. Generally a squared pulse voltage with a specific width is applied to the device to change the state of RRAM. However because of existing process variability in the RRAM array, the resistance state can be in its tail points (a very low LRS or a very high HRS) and therefore the voltage pulse would not be sufficient to switch the RRAM to the correct state. This may lead to a memory fault and can be detected with special March tests inside the RRAM array [15].

Generally there are two commonly used read approaches in RRAM memory arrays: 1-Voltage mode and 2-Current mode sensing, where in both of them a sensed voltage/current (V_{cell}/I_{cell}) is compared with a reference voltage/current (V_{Ref}/I_{Ref} generated using reference cells) and the state of the device is recognized. Depending on the RRAM if it is in LRS or HRS state the current/voltage would differ and a ´0´ or ´1´ is read from the cell.

By considering normal distributions for LRS, HRS and the Reference resistance it is possible to analyze the probability of failure in the read procedure of RRAM array. Figure 3 shows the probability of failure when considering Reference=(LRS+HRS)÷2.
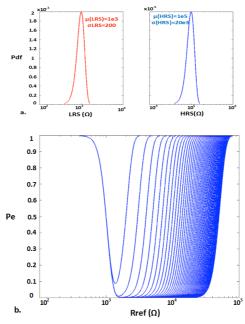

Figure 3. Impact of variability in RRAM Read process

The wide resistance distribution of RRAM cells can impose low read yields and can reduce the device lifetime [16], therefore it is needed to design variability tolerant circuits, presented in the next section.

III. VARIABILITY MITIGATION

There are different methodologies to improve or mitigate the variability of RRAM devices. One of the conventional methods is based on material engineering that tries to decrease the device resistance fluctuations by utilizing enhanced material or interfaces [3]. Another variability-mitigating approach is to use optimized programming (SET/RESET) methods that can constrain the distribution of LRS and HRS states [7]. Ultimately the key final approach which is mainly considered in this paper is based on circuit design innovations inside the RRAM memory array or its CMOS peripherals, to suppress the variability effect. Figure 4 shows these approaches and the circuit-based techniques are presented in the following.
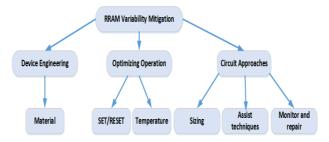
Figure 4. RRAM variability mitigation techniques

A. Sizing

The concept of RRAM sizing is different from the conventional transistor sizing as the device contains two terminals and does not have an active channel. It has been shown that the memory RRAM cell can be constructed by two parallel RRAMs to mitigate the variability fallout [17]. This technique has been relevant by using the concept of making the resistors and consequently the overall resistance value can be more variation-tolerant, though it can impact the performance and area overhead of the memory array.

B. Circuit Assist Techniques

Because of the existing variations in the LRS and HRS and also the variations in reference cells the read process can become incorrect. Therefore it is necessary to design sensing circuits/approaches that can tolerate these fluctuations. One of the such techniques is to reduce the V_{Ref}/I_{Ref} distribution to ensure a correct read process inside the RRAM memory array. This can be done by a parallel-series reference cell (PSRC) scheme that narrows the I_{Ref} distribution and reduces σI_{Ref} against the resistance variation [18]. Figure 5 shows the concept for the failure and clarifies the fact that a narrowed I_{Ref} distribution can help to improve the read margin.
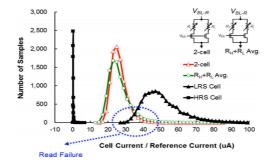

Figure 5. Narrowing I_{Ref} is a solution to improve the read margin adapted from [18]

Another approach to improve the read margin in presence of process variability can be based on utilizing an adaptive reference current scheme to feed the current mode sense amplifier [19]. The adaptive reference current scheme can detect the spatial resistance distribution and generate the optimum reference current for reliable read operation. Figure 6 shows the adaptive reference current scheme to make the correct read operation in a variability aware RRAM memory array.
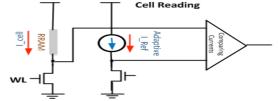

Figure 6. Adaptive reference current to improve RRAM Read margin

In addition to these approaches there are also other circuit assist techniques implemented at the periphery of the RRAM array to enlarge the memory read window and enhance the RRAM memory yield [20]. Such approaches aim for advanced read/write mechanisms to tolerate the existing process variability inside the RRAM arrays.

The existing fluctuation in the resistance distribution of RRAM cells motivates the design for monitoring techniques to measure the HRS to LRS resistance ratio to guarantee a reliable memory operation. One of such monitoring techniques is presented in the next section.

IV. VARIABILITY MONITORING TECHNIQUE

In order to show the mechanism of our monitoring methodology first the RRAM memory operation is described in brief. Then the monitoring circuit is presented and the simulation results according to the variability monitoring are demonstrated.

RRAM memory systems are commonly organized in a matrix-like structure called crossbar. The storage cell in the crossbar can be built with only one RRAM device (1R cell), but due to the well-known problem of sneak-path a select device such as a transistor is utilized in constructing the memory cell (1T1R) [21]. Therefore the 1T1R cell generally consists of an NMOS transistor and a resistive switching device. In this structure the RRAM current is correctly controlled through the crossbar and each device is turned 'on' or 'off' based on the row address in the crossbar. Figure 7 presents an example 2x2 crossbar memory constructed with 1T1R as storage cells. Note that the square boxes in this figure are multiplexers, which govern the appropriate voltage or current to be applied to the cells.

Each 1T1R cell can be written and read by applying the appropriate signals through the bit-line (BL), word-line (WL) and the select line (SL). To write '1' in the cell (also called SET process), first the corresponding WL is activated in the crossbar, and next while the SL is grounded an appropriate voltage is applied at the BL (V_{DD}). Changing the state of RRAM from LRS to HRS, is called the RESET process, and it is equivalent to writing a '0' in the cell. For this process the WL is again activated, when a proper voltage (V_{DD}) is applied at the SL, and while the corresponding BL is grounded, the other BLs in the block are all in high impedance mode.

To read the RRAM device using the voltage sensing mode a V_{Read} voltage (lower than the write voltage) is applied to the corresponding 1T1R cell, and the BL voltage (a voltage division of the read voltage through the RRAM resistance and the resistor Rx), will be detected through the comparator giving the state of RRAM and the bit-cell. A reference cell, consisting of a reference resistance (R_{Ref}), is used to generate a reference voltage to be compared in the comparator. A resistive device in LRS state will produce a higher voltage in comparison with one at HRS mode. It is important that the read voltage (V_{Read}) be much lower than the write voltage V_{DD} in order not to modify the state of RRAM in the read operation. Also note that in this work the read mode based on voltage is chosen, the reason is controlling the applied voltage to the RRAM cell is easier and also the model which we will later use for circuit simulation works better in voltage mode.

To implement the variability-monitoring mechanism, it is considered that the crossbar memory can benefit from idle cycles during its operation and the granularity for monitoring scenario in a crossbar is row-by-row in each phase because it is easier to control the monitoring flow rather than in the column-by-column approach. Also the area overhead can be reduced if the monitoring granularity is chosen to be at row level and not in column level, because less multiplexer would be needed. Here, a monitoring technique based on analyzing the process variability of the RRAM cells during the memory lifetime is demonstrated. This monitoring approach can be applied to the cells, row-by-row one-by-one and can evaluate the relation between HRS and LRS for a robust memory operation. In fact the monitoring includes a set of write and read operations, where in each step resistance state is monitored through the monitoring circuits.

Figure 7 shows the monitoring circuit attached to the 2x2 crossbar. This monitoring circuit is constructed with some switches (SW_M; SW_L; SW_H) to control the correct passing current, two capacitors (C_{LRS}, C_{HRS}) to store respecting voltages to LRS and HRS, a divisor to calculate the ratio between HRS and LRS, and a comparator to compare the division result with the design value of K=HRS÷LRS. The cells will undergo the monitoring phase one-by-one and their reliability will be analyzed.
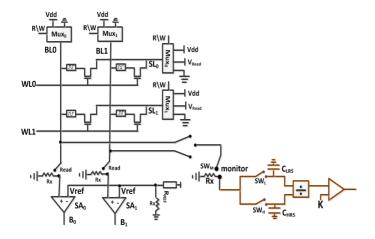

Figure 7. A 2x2 RRAM memory with peripherals along with the variability monitoring circuit

In order to perform the circuit simulations, this work uses the resistive switching (RRAM) Verilog model designed by Stanford University [22]. This model is designed for bipolar metal oxide RRAM devices based on conductive filament switching concept and has no limitations on the size of RRAM cell. We have considered all the default sizes in the model proposed by authors such

as the cell size, which is equal to 10x10 nm^2. The monitoring circuit shown in Figure 7 is simulated by HSPICE for monitoring one RRAM cell, and Figure 8 depicts the circuit simulation obtained results for one RRAM cell.

The simulation is transient and contains first writing a '0' inside the cell assuming LRS as an initial state. In order to write '0' inside the cell without interference with monitoring section we have considered some switches in the monitoring path. At this phase, the switches are as following SW$_M$=open to isolate the monitoring circuits and also SW$_L$, SW$_H$=open. At the moment of writing '0' the RRAM current goes to almost zero, because the resistance state is switched from LRS to HRS and no current can flow inside the device. Then, in the monitoring phase, we are interested to measure the value of HRS. This is done by applying appropriate voltage from the source-line and to close two of the switches in the circuit (SW$_M$=close, SW$_H$=close), while one switch is kept open SW$_L$=open for further monitoring steps. Figure 8 shows in detail that the corresponding voltage to HRS is stored through our monitoring circuit and capacitor (C$_{HRS}$), such that the V$_{CHRS}$ is equal to the voltage divided between Rx and HRS.

Next the objective would be to measure the corresponding voltage according to LRS of the RRAM. Therefore, it is needed to write '1' inside the cell. Again, the isolating switches should be open (SW$_M$, SW$_L$, SW$_H$=open) and a transient HSPICE simulation is done. Figure 8 shows how the RRAM current goes up to 300uA, due to its switching from HRS to LRS, in which more current can pass through the device. After writing '1' it is time to monitor the LRS status of the device by applying the appropriate voltage from the source-line and closing two switches such as (SW$_M$=close, SW$_L$=close) and keeping one switch open (SW$_H$=open) for proper monitoring. Then, accordingly the other capacitor in monitoring circuit (C$_{LRS}$), V$_{CLRS}$ holds the voltage divided between Rx and LRS. After these, the monitoring phase continues with dividing these values and finding out their ratio that is {HRS÷LRS} (around 23 in this example for a fresh RRAM cell in Stanford model). Note that, in this case the value of Rx is important and can have significant impact on the measurement if it is chosen big. In the phase of monitoring if the {HRS÷LRS} value is acceptable (for instance in this specific example 23 > K) the cell is recognized as a healthy cell and if not it is identified as a weak, making the memory repairing techniques essential.
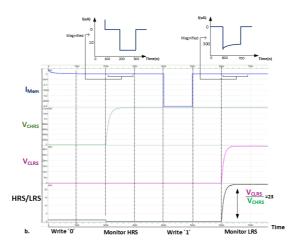

Figure 8. HSPICE circuit simulation using Stanford model to monitor variability by measuring HRS to LRS ratio

V. CONCLUSIONS

In this work RRAM reliability is analyzed from process variability point of view. The main sources of RRAM variability are presented and the impact on RRAM memory operation is analyzed. Then circuit techniques are introduced to mitigate the variability effect and improve the yield of RRAM memory array. Finally, a variability-monitoring scheme is presented to measure the on/off ratio inside the RRAM memory array and to monitor the fluctuations between the HRS and LRS. This monitoring technique can help to differentiate the reliable RRAM memory cells from the weak ones.

ACKNOWLEDGMENT

REFERENCES

[1]   H.-S. P. Wong, et al, "Phase Change Memory," *Proc. IEEE*, vol. 98, no. 12, pp. 2201–2227, 2010.

[2]   D. Apalkov, et al, "Spin-Transfer Torque Magnetic Random Access Memory (STT-MRAM)," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 9, no. 2, 2013.

[3] H.-S. P. Wong, et al, "Metal–Oxide RRAM," *Proc. IEEE*, vol. 100, no. 6, pp. 1951–1970, 2012.

[4] D.J. Wouters, R. Waser, M. Wuttig, "Phase-Change and Redox-Based Resistive Switching Memories," *Proc. IEEE*, vol. 103, no. 8, pp. 1274–1288, 2015.

[5] C. Xu, et al, "Impact of Cell Failure on Reliable Cross-Point Resistive Memory Design ," *ACM Journal on Design Automation of Electronic Systems (TODAES)*, vol. 20, no. 4, 2015.

[6] www.itrs.net

[7] A. Chen, M.R Lin, "Variability of Resistive Switching Memories and its Impact on Crossbar Array Performance," *IEEE IPRS*, pp.10.1–10.4, 2011.

[8] F. T. Chen, "Fast-Write Resistive RAM (RRAM) for Embedded Applications," *IEEE Des. Test Comput.*, vol. 28, no. 1, pp. 64–71, 2011.

[9] S. Yu, et al, "Improved Uniformity of Resistive Switching Behaviors in HfO$_2$ Thin Films with Embedded Al layers," *Electrochem. Solid State Lett.*, vol. 13, pp. H36–H38, 2010.

[10] B. R. Field-, D. Ielmini, "Modeling the Universal Set / Reset Characteristics of Filament Growth," *IEEE TED*, vol. 58, no. 12, pp. 4309–4317, 2011.

[11] C. Zambelli, et al, "RRAM Reliability/Performance Characterization through Array Architectures Investigations," *IEEE ISVLSI*, pp.327–332, 2015.

[12] S. Yu, X. Guan, H.-S. P. Wong, "On the stochastic nature of resistive switching in metal oxide RRAM: Physical modeling, monte carlo simulation, and experimental characterization," *Proc IEDM*, pp. 17.3.1–17.3.4, 2011.

[13] P.Pouyan, E.Amat, A.Rubio; "Reliability Challenges in Design of Memristive Memories" *Proc Vari*, pp 1-6, 2014.

[14] H. Manem, J. Rajendran, G.S. Rose, "Design Considerations for Multilevel CMOS/Nano Memristive Memory," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 8, no. 1, 2012.

[15] C.-Y. Chen, et al, "RRAM Defect Modeling and Failure Analysis Based on March Test and a Novel Squeeze-Search Scheme," IEEE Trans. Comput., vol. PP, no. 1, pp. 1–1, 2014.

[16] Pouyan, E.Amat, A.Rubio; "Statistical Lifetime Analysis of Memristive Crossbar Matrix", *Proc DTIS,* 2015.

[17] J. Rajendran, R. Karri, and G. S. Rose, "Improving Tolerance to Variations in Memristor-based Applications Using Parallel Memristors," *IEEE Trans. Comput.*, vol. 9340, no. c, pp. 1–1, 2014.

[18] R. Schemes, M. Chang, S. Sheu, K. Lin, C. Wu, C. Kuo, P. Chiu, Y. Yang, Y. Chen, H. Lee, C. Lien, F. T. Chen, K. Su, T. Ku, M. Kao, and M. Tsai, "A High-Speed 7.2-ns Read-Write Random Access 4-Mb Embedded Resistive RAM (ReRAM) Macro," *IEEE JSSC*, vol. 48, no. 3, pp. 878–891, 2013.

[19] J-K Park, et al, "Analysis of Resistance Variations and Variance-aware Read Circuit for Cross-Point ReRAM," *Proc IMW*, pp. 112–115, 2013.

[20] X. Xue, et al, "A 0.13 µm 8 Mb Logic-Based Cu$_x$Si$_y$ ReRAM With Self-Adaptive Operation for Yield Enhancement and Power Reduction," *IEEE JSSC*, vol. 48, no. 5, pp. 1315–1322, 2013.

[21] A. Mohammed, H. Aly, H. Fahmy, M. Mustafa, and K. Nabil, "Memristor-based Memory : The Sneak Paths Problem and Solutions," *Microelectronics Journal,* pp. 1–9, 2012.

[22] "Standford RRAM model" [on line], https://nano.stanford.edu/stanford-rram-model