Stochastic Failure Model for Endurance Degradation in Vacancy Modulated HfO_x RRAM using the Percolation Cell Framework

Nagarajan Raghavan[1, ♣], Kin Leong Pey[1], Daniel D. Frey[2] and Michel Bosman[3]
[1]Engineering Product Development (EPD) Pillar, Singapore University of Technology and Design (SUTD), Singapore – 138 682.
[2] Engineering Systems Division, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139.
[3]Institute of Materials Research and Engineering (IMRE), A*STAR, Singapore – 117 602.
♣Ph: (+65) 9862 1185, E-mail: nagarajan@sutd.edu.sg

Abstract — **Endurance is a key performance metric for non-volatile memory devices. For resistive switching random access memory (RRAM) technology, which operates based on the reversible drift of oxygen ions in a bipolar switching scheme, the degradation in endurance can be stochastically modeled by using the percolation cell framework, as the SET and RESET events are akin to the dielectric breakdown and recovery phenomena in logic gate stacks. Based on our understanding of the physical mechanisms postulated for endurance degradation and the electrical test results on the memory window closure, we propose a generic statistical model in this study that can be used to simulate the endurance cycle distribution as a function of the pulse voltage, pulse duration, compliance level, activation energy, dielectric thickness, filament temperature and filament size / shape. Fitting the endurance model to real test data helps in estimating the value of the activation energy for oxygen ion transport. Our focus here is only on the "failure to RESET" mechanism, whereby the filament does not rupture to reach the high resistance state due to a large imbalance in the oxygen ion and oxygen vacancy count.**

Keywords – Endurance, Filament, Oxygen vacancy, Percolation, Reset, Thermochemical model.

I. INTRODUCTION

Non-volatile memory (NVM) technology has undergone significant progress over the past few years with the dominance of Flash memory technology, which is currently being used for data storage applications in all modern commercial electronic products and systems. The 2D planar version of Flash has now given way to 3D NAND Flash [**1**, **2**] with better write performance and reliability. In spite of all these material and architectural evolutions in Flash, it is likely to become a bottleneck soon as further scalability is challenging given that the difference between a binary '0' and '1' in a 60 nm node Flash cell is controlled only by 67 electrons [**3**]. Even if a few of these charges were to leak out, there will be substantial impact on the reliability, variability (noise) and shrinking memory window. Technologies such as phase change memory (PCM) have not shown much promise as they require high power for switching [**4**]. Magnetic spin-transfer torque RAM (STT-RAM) is another potential candidate with supposedly infinite endurance [**5**], however the unconventional material stack makes it expensive for fabrication and non-CMOS compatible as well with lot of effort needed in optimizing process conditions. The most promising technology that is likely to replace 3D Flash in a couple of years is the resistive random access memory (RRAM) [**6**], with a metal-insulator-metal (MIM) stack, which is a simple structure that is easy to design and fabricate. The resistive switching mechanism in RRAM can be governed either by oxygen vacancy and oxygen ions or nucleation / rupture of metallic conducting filaments. Our focus in this study is on OxRAM, a sub-class of RRAM, where switching involves generation of oxygen vacancy defects in the dielectric for SET and their recombination with mobile oxygen ions for RESET.

As in all NVM technologies, there are three key performance metrics for RRAM → endurance, retention and read disturb. Degradation of any of these metrics is a critical reliability issue for the memory. Various electrical tests have shown that endurance degrades due to lowering resistance in high resistance state (HRS) attributed to accumulation of defects and the difficulty to RESET from low resistance state (LRS) as we cycle the device repeatedly [**7**] – [**9**]. Possible mechanisms that have been proposed to explain endurance degradation are oxidation of electrode, increased vacancy count and additional filament generation during stress as well as lateral diffusion of O^{2-} ions away from core filament [**7**].

While a good amount of qualitative knowledge on the physics of degradation now exists for RRAM, there is limited insight on the statistics governing the failure. We are in need of analytical models that can be used to predict the lifetime of the memory device (failure distribution and slope / variance) as a function of the operating conditions such as V_{SET}, V_{RESET}, filament temperature (T_{FIL}) and compliance level for SET (I_{comp}) and design considerations such as dielectric thickness (t_{ox}) and area. Some attempts have been made to model the stochastics of retention failure [**10**, **11**] and read disturb [**12**]. However, as far as we know, there is no statistical model developed to look into endurance failures in OxRAM. Given that the SET operation in OxRAM is similar to dielectric breakdown in high-κ logic stacks and RESET is akin to the recovery of breakdown path, we make use of the percolation cell framework (which is the most standard approach to formulate dielectric failure probability [**13**, **14**]) to model the endurance failure here. The standard material stack comprising HfO_2 and TiN based electrode with high oxygen solubility is considered here for a case study.

The content of this work is streamlined as follows. In Section II, we shall present the percolation framework tailored to

represent the filament configuration in OxRAM and introduce the Markov process to represent the two-stage reset mechanism involving ionic transport and ion-vacancy recombination. The overall physical mechanism for endurance failure is qualitatively discussed in Section III. Using the physical mechanism as the basis, we formulate the endurance failure distribution in Section IV to derive the analytical expression for endurance reliability. We make use of the developed model to simulate the Weibull probability plots for the number of cycles to failure in Section V, examining the dependence of the endurance distribution on reset diffusion activation barrier (E_{aD}), lateral size of filament, I_{comp}, T_{FIL}, t_{ox} and reset pulse parameters (V_{ps}, T_{ps}) where V_{ps} is the negative voltage for RESET and T_{ps} is the duration of this signal. Finally, in Section VI, we shall present our inferences from this study and conclude with possible suggestions for further work.

II. Percolation Cell and Markov Model

The percolation cell schematic for the OxRAM in the LRS is shown in Fig. 1, where the size of the filament (number of rows and columns of vacancies) depends on the compliance chosen for the forming and SET phenomena. In contrast to the conventional percolation models where we begin our analysis on a virgin device and develop an expression for the failure probability considering the column that percolates first with all the cells turned 'ON" for conduction, the approach here is completely reverse whereby we start with the LRS and formulate the probability for the filament to rupture ("reset"). Endurance failure is considered to be the case where reset does not occur (filament does not rupture). We define "*reset*" to have taken place successfully as long as there is at least one row of oxygen vacancies that are annihilated in the percolation path by oxygen ions, thereby introducing a dielectric tunnel barrier in the filament, lowering conductance significantly. Although "failure to SET" has also been observed in some endurance studies [**8**], we do not consider that scenario here as the mechanism behind it is not clear yet.

The *reset* process is considered as a two-stage (three state) sequential mechanism involving O^{2-} ion drift from top electrode into the dielectric (State I → State II), followed by O^{2-} - V_0^{2+}recombination (State II → State III). This can be represented as a Markov Chain in Fig. 2, with the transition rates, λ_D and λ_R for diffusion / drift and recombination respectively. The differential expressions governing time dependent Markov state probability are given by Eqns. (1)-(3), where $p_I(t = 0) = 1$ and $p_{II}(t = 0) = p_{III}(t = 0) = 0$ is the initial condition for the three states. Given a reset pulse with parameters (V_{ps}, T_{ps}), reset probability is the probability of reaching the final state III. Solving the set of differential equations below, the final expression for $p_{REC}(V_{ps}, T_{ps})$ can be expressed by Eqn. (4).

$$\frac{dp_I(t)}{dt} = -\lambda_D \cdot p_I(t) \quad (1)$$

$$\frac{dp_{II}(t)}{dt} = \lambda_D \cdot p_I(t) - \lambda_R \cdot p_{II}(t) \quad (2)$$

$$\frac{dp_{III}(t)}{dt} = \lambda_R \cdot p_{II}(t) \quad (3)$$

$$p_{REC}(V_{ps}, T_{ps}) = 1 - \frac{1}{1 - \lambda_R/\lambda_D} \cdot e^{-\lambda_R T_{ps}} + \frac{1}{\lambda_D/\lambda_R - 1} \cdot e^{-\lambda_D T_{ps}} \quad (4)$$

Figure 1. Schematic showing the filament, vacancy (white and black circles) and oxygen ion (red circles) distribution after the first forming event and after endurance failure. With more switching cycles, some oxygen ions may get partially oxidized at the top electrode (shown by shaded lines), thereby increasing the diffusion barrier. The filament grows in size due to vacancy accumulation leading to "*failure to reset*" as the dynamic balance between complementary V_o^{2+} : O^{2-} pairs for recombination is reduced.

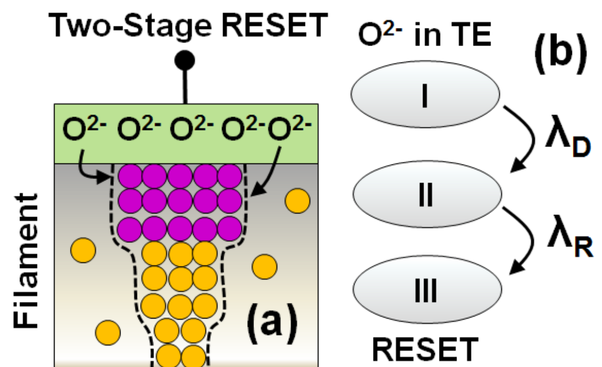

Figure 2. (a) Possible arrangement of the vacancies in the SET state with successful reset being determined by the ability to rupture the filament anywhere within the dielectric. The thickest part of the filament is shown in pink. (b) We model the dynamics of RESET by a two-stage (three-state) process involving defect transport (drift / diffusion) and recombination, represented by the Markov state diagram.

The rate expressions in the Markov Chain for the drift / diffusion (λ_D) [**15**] and recombination processes (λ_R) [**16**] may be expressed by Eqns. (5) and (6) respectively. The terms E_{aD}, E_{aR}, ν, a, d, T_{FIL} and ξ refer to the diffusion and recombination activation barriers, lattice vibration frequency and periodicity, diffusion distance, local filament temperature and electric field. Since we are analyzing the LRS state, it is likely that the localized current through the filament prior to reset can give rise to a much higher internal filament temperature (T_{FIL}), compared to the ambient test temperature (T_0) due to joule heating effects. To account for this thermal excitation, we express T_{FIL} by a steady-state first order approximation as in Eqn. (7), where R_{th} and R_{ON} refer to the thermal resistance of HfO_x (~5-50 × 10^4 K/W [**17**]) and ON resistance respectively, while $T_0 = 298K$.

$$\lambda_D = \frac{\nu}{d} \cdot a \cdot \exp\left(-\frac{E_{aD}}{k_B T_{FIL}}\right) \cdot \sinh\left(\frac{qa\xi}{2k_B T_{FIL}}\right) \quad (5)$$

$$\lambda_R = \nu \cdot \exp\left(-\frac{E_{aR}}{k_B T_{FIL}}\right) \quad (6)$$

$$T_{FIL} = T_0 + \frac{R_{th}}{R_{ON}} \cdot V_{ps}^2 \quad (7)$$

III. PHYSICAL MECHANISM OF ENDURANCE FAILURE

The possible mechanism underlying endurance failure is also exemplified in Fig. 1. With increasing number of switching cycles, additional defects are generated in the vicinity around the formed conductive filament. The oxygen ions liberated from these extra defects drift to the top electrode while the vacancies tend to diffuse towards the filament (akin to grain boundary) which serves as a thermodynamic sink [**18**]. With increasing likelihood of partial oxidation of the electrode (forming metal oxide), the number of mobile O^{2-} ions is progressively limited and the diffusion barrier E_{aD} gets enhanced, while the filament size laterally dilates due to vacancy accumulation. In effect, the reset process becomes *self-limiting* with lower probability of filament rupture due to larger imbalance in the vacancy and mobile oxygen ion density, which eventually leads to endurance failure. Note that the above hypothesis for the physical mechanism encompasses all the degradation processes postulated earlier in Ref [**7**]. Therefore, the physics behind endurance failure is quite complex and involves many processes that are interrelated to each other.

IV. FORMULATING ENDURANCE FAILURE DISTRIBUTION

We shall now use the illustration in Fig. 1 to express the reset failure probability ($P_{RESET}(t)$) and endurance reliability function (R_{END}). As mentioned earlier, we consider reset to be successful as long as "*at least*" one row of vacancies in the filament is fully passivated by the drifting oxygen ions so that a tunnel barrier is introduced, causing the resistance to shoot up by many times. In other words, the "failure to RESET" occurs if and only if every single row in the filament fails to rupture. This is expressed in Eqn. (8), as $F_{RESET}(t) = (F_{ROW}(t))^m$, given that we have m rows in the filament. Here, $F_{RESET}(t)$ and $F_{ROW}(t)$ refer to the failure probability for reset and probability of no rupture in a single row respectively. A single row of vacancies fail to rupture as long as even one of the vacancies remain unpassivated. Therefore, the expression for $F_{ROW}(t)$ can be expressed by the last line in Eqn. (8) where the failure for the row to rupture is one minus the probability that each of the n vacancies is passivated. Here, n refers to the number of columns in the filament. For simplicity, we assume n to be uniform all along the filament, i.e., it is cylindrical in shape. We will later relax this assumption when analyzing the influence of filament shape.

$$\begin{aligned} P_{RESET}(t) &= 1 - F_{RESET}(t) \\ &= 1 - [F_{ROW}(t)]^m \\ &= 1 - \left[1 - (p_{REC}(t))^n\right]^m \end{aligned} \quad (8)$$

Our insight thus far has been constrained to a single reset event. Making use of Eqn. (8), we may now express the overall endurance reliability function (R_{END}) for N_{SW} switching cycles (assuming there is no "failure to SET" event) by Eqn. (9), which is just a simple series product of the reset success probability for each and every cycle. Having obtained R_{END}, we can convert it to the Weibit scale for probability plotting by using the standard conversion expression shown in Eqn. (10). With the complete endurance model now derived, we shall simulate the endurance failure distribution to examine the influence of various material and operational design factors.

$$R_{END} = \prod_{i=1}^{N_{SW}} P_{RESET(i)}(V_{ps}, T_{ps}) \quad (9)$$

$$W_{END} = \ln(-\ln(R_{END})) \quad (10)$$

V. STATISTICAL ENDURANCE SIMULATION AND ANALYSIS

A. *Effect of Gradual Oxidation of Top Electrode*

The endurance distribution for the ideal case where we assume no degradation or wear out during the cycling process is plotted in Fig. 3 for V_{ps} = -1V and T_{ps} = 10μs, given the filament constriction size of $n = 3$ and $m = 3$. Notice that the distribution is highly non-Weibull and appears very skewed and unrealistic because it estimates that some devices may switch only a few ten times, while some may switch almost 10^{10} times within a reasonable percentile range. It is very unlikely for the spread in the endurance cycles to be so large. Another strange trend is that the distribution seems to shift quite a lot even for small ΔE_{aD} variations of ~5meV.

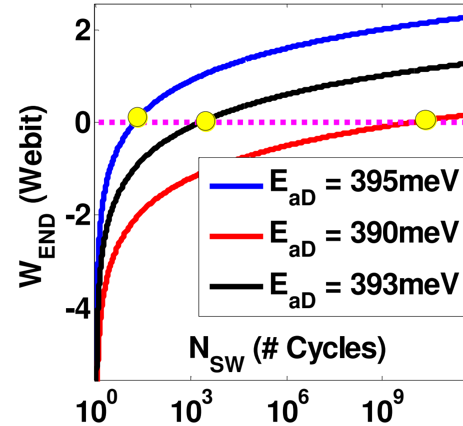

Figure 3. Endurance failure distribution considering no degradation in the material parameter values (E_{aD}) during the cycling and no change in filament size as well.

When we modify the ideal model to empirically factor in the gradual increase in E_{aD} value using Eqn. (11), the resulting distribution plot is shown in Fig. 4 where the trend is increasingly steeper as the scaling factor (SF) increases. The higher the SF, the faster is the oxidation of the metal electrode resulting in increased migration barrier. There is a hump in the distribution at low percentiles. These steep distribution trends appear more realistic and confirm that the gradual wear-out that we postulated in Section III does exist and needs to be accounted for in the modeling process.

$$E_{aD} = E_{aD0} + SF \cdot (\log N_{SW})^2 \quad (11)$$

B. Role of Compliance and Filament Size on Endurance

The endurance distribution plot in Fig. 5 shows the trend for three different lateral sizes of the filament which is accounted for by varying the value of n, which is the number of columns, keeping the number of rows fixed at $m = 3$ (corresponding to t_{ox} ~ 1.8 nm if we assume the defect size to be a_0 ~ 6Å). We observe that the endurance degradation is more severe when we have thicker filaments. While this may sound logical since rupture of thicker filaments requires more recombination events in order to induce a tunnel barrier, experimental reports show a completely opposite trend, whereby the endurance is way better for thicker filaments [**19**]. Note here that we tweak the n value to represent larger filaments which arise when a higher compliance (I_{comp}) is chosen for the forming and SET events. In Fig. 5, we have blindly assumed that the filament temperature is at its ambient value of T_0 = 298K irrespective of its size. This is however not true, as thicker filaments will have more localized current density and associated joule heating. As a result, the T_{FIL} for n = 3, 4, 5 can be markedly different.

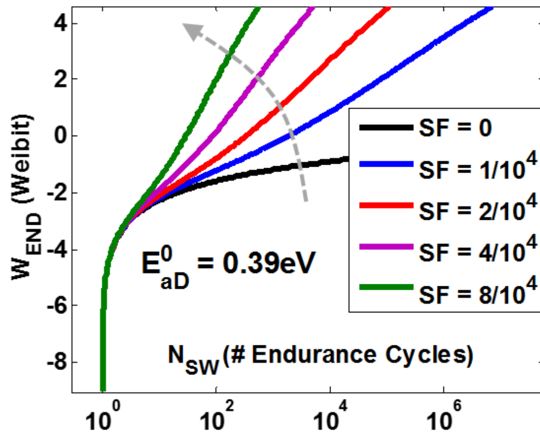

Figure 4. Endurance failure distribution considering a gradual increase in E_{aD} value (represented by the scaling factor, SF, in Eqn. (11)) due to oxidation of the top electrode metal. The simulation here considers an initial value for E_{aD} ~ 0.39eV.

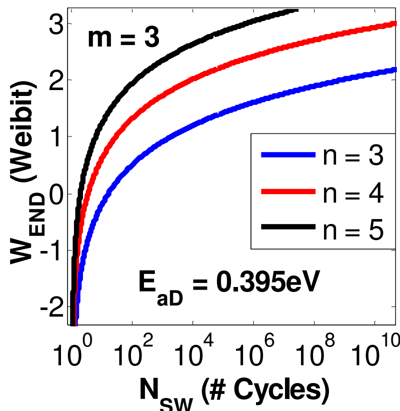

Figure 5. Weibull plot of endurance distribution for different lateral filament size (t_{ox} ~ 1.8 nm), assuming no temperature effect.

When we consider the effect of T_{FIL} by assigning a lower R_{ON} for larger n (higher I_{comp}), we see a complete reversal of the trends in Fig. 6 where in the thicker filaments (higher compliance switching events) show higher endurance. The higher value of T_{FIL} for higher I_{comp} more than compensates for the increased number of vacancies to be annihilated to reach HRS. The confirmation of our simulation results in Fig. 6 with the hypothesis in Ref [**19**] is a clear evidence that the local filament temperature plays a vital role. It can be inferred from this analysis that while we desire to have low power switching by means of setting low I_{comp}, we would have to compromise for poorer endurance lifetime. Improved performance in most cases comes at the expense of sacrificing reliability and vice versa.

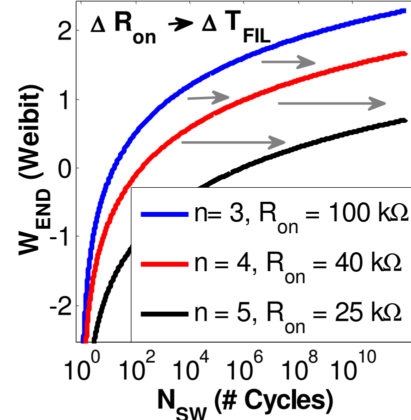

Figure 6. Weibull plot of the endurance distribution for different lateral filament size (t_{ox} ~ 1.8 nm), considering the Joule Heating steady state local temperature in the filament.

C. Role of Lateral Dilation and Wear-Out of Filament

Similar to the wear-out due to metal electrode oxidation, another source of wear-out is the lateral expansion of the filament during repetitive cycling. In order to model this dilation, we construct another empirical relationship for n (number of columns which represents filament size) as a function of N_{SW} (see Eqn. (12)). When the endurance plot is analyzed for different values of SF as shown in Fig. 7, we observe the distribution getting steeper with higher SF and realistic variance. Both Fig. 4 and Fig. 7 support the fact that degradation in E_{aD} and n is intrinsic to endurance failure in all high-κ based M-I-M stacks.

$$n = n_0 + SF \cdot (\log N_{SW})^2 \qquad (12)$$

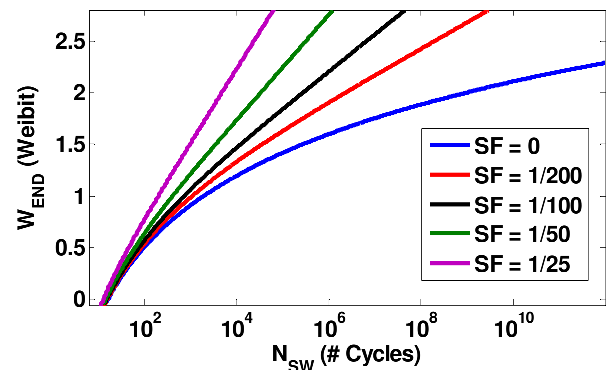

Figure 7. Weibull plot of cycles to failure considering a gradual lateral dilation of the filament (represented by SF in Eqn. (12)) due to increasing V_0 accumulation.

D. Role of Dielectric Thickness on Endurance

It is well known that thicker dielectrics will have a more catastrophic filament formation process as the percolation transient is harder to control. The width of filaments in thicker dielectrics is therefore larger. To represent this scenario, we simulate the endurance distribution for symmetric filaments of cell size 2×2 (t_{ox} ~ 1.2 nm), 3×3 (t_{ox} ~ 1.8 nm) all the way up to 6×6 (t_{ox} ~ 3.6 nm) in Fig. 8, whereby we increase the width proportionally with the height of the filament (t_{ox}). It is interesting to note that thicker dielectrics show far better endurance. This could be physically explained by the fact that there are more mobile O^{2-} ions in thick stacks that are available to initiate a rupture of at least one row of vacancies in the filament. As a simple scenario, reset in 3×3 filament would require 3 vacancies in a row to be passivated by a population of 9 oxygen ions making it a ratio of 3:1. However, for a 6×6 filament, we would require 6 vacancies to be passivated from a large population of 36 oxygen ions, which increases the ratio to 6:1. Therefore, it is natural to expect thicker dielectrics to have a higher endurance. The enhanced T_{FIL} in larger filaments also contributes to this trend.

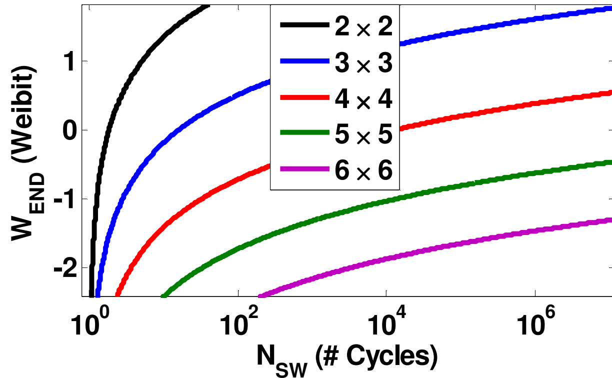

Figure 8. Endurance distribution for "squarish" filaments in RRAM samples of different thickness. We consider filaments with cell sizes ranging from 2×2 up to 6×6, corresponding to dielectric thickness in the range 1.2 – 3.6 nm.

E. Effect of Reset Voltage and Pulse on Endurance

The influence of reset voltage (V_{ps}) and reset pulse period (T_{ps}) in the pulsing scheme on the endurance cycle distribution is represented in Figs. 9(a) and (b) respectively taking t_{ox} ~ 1.8 nm. Even small changes in V_{ps} (T_{ps} ~ 10µs) or T_{ps} (V_{ps} ~ -0.8V) seem to affect the N_{SW} distribution significantly. This is due to the exponential dependence of p_{REC} on these two parameters and the series product of P_{RESET} over many cycles that determines the final endurance failure probability. In Fig. 10, we present a contour plot of the $\{V_{ps}, T_{ps}\}$ values that reproduce exactly the same endurance distribution trend for different values of E_{aD}. We may choose the optimum $\{V_{ps}, T_{ps}\}$ combination taking into account the memory operation speed (frequency) requirements as well as possibility of abrupt hard breakdown which increases when we choose a higher V_{ps}.

F. Effect of Filament Shape and Asymmetry

We have thus far assumed the filament to be 'squarish' and symmetrical. However, considering that the top and bottom electrode in an OxRAM device could be of different materials, it is very likely that the filament may be wider on one end and narrower on the other. To assess the impact that the shape and asymmetry of the filament would have on the endurance statistics, we purposely simulated the endurance for a small filament with 6 vacancies arranged in (a) symmetric, (b) conical and (c) inverse conical fashion as shown by Fig. 11. In this case, we relax our previous assumption so that defects further away from the top electrode – dielectric interface would have lower rates of recombination due to longer ionic drift distance.

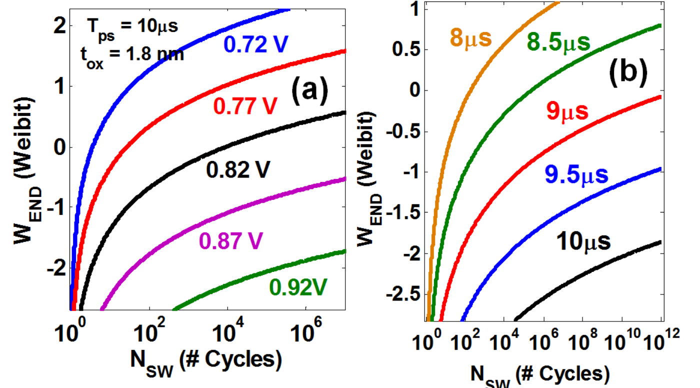

Figure 9. Shift in the endurance failure distribution for (a) change in reset pulse voltage (V_{ps}) and (b) reset pulse period (T_{ps}). Due to the exponential dependencies of reset probability on V_{ps} and T_{ps}, even small changes in the operational condition can cause significant shift in the distribution trends.

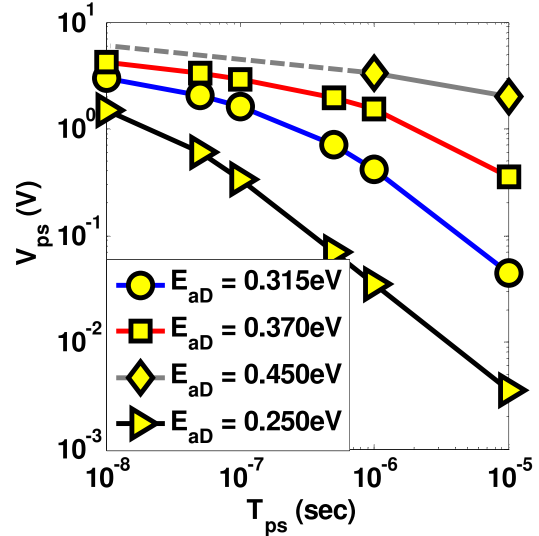

Figure 10. Contour plot of possible combination of V_{ps} and T_{ps} values that give the same endurance distribution, assuming different activation barriers (E_{aD}) ranging from 0.25 – 0.45eV.

From Fig. 11, it turns out that at high percentiles, symmetric filaments show the poorest endurance, followed by inverted cone and then upright conical. In an asymmetric filament, we have very few vacancies on either end of it which makes it more feasible for reset to occur as opposed to symmetric ones where there is no potential "necking" point for easier rupture. It is therefore effective to choose the materials for the OxRAM stack such that filament is asymmetrical during the forming and SET events. This may require the purposeful use of highly oxygen

scavenging films (such as hafnium) on any one side of the electrode to create a wider base of the defects there during the fabrication stage itself.

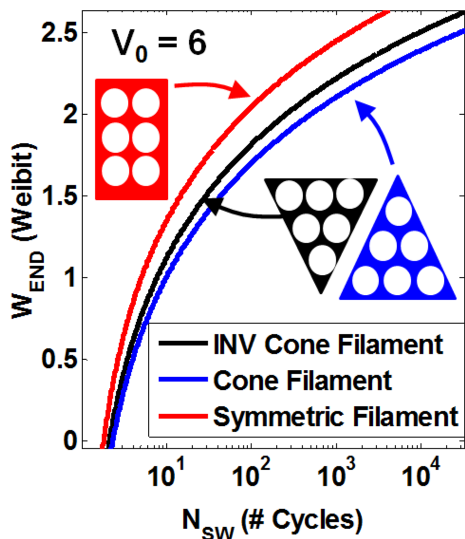

Figure 11. Role of different filament geometry on the endurance distribution considering lower probability of reset for more distant traps from top electrode – dielectric interface. The symmetric filament shows the poorest lifetime.

G. Parameter Extraction for Endurance Test Data

Although most RRAM literature reports include some results on endurance cycling test, it is generally carried out only on very few devices with limited scope for a proper statistical investigation. A recent in-depth endurance study by Y.Y. Chen *et. al.* [**8**] is one of the only studies with a plot of the endurance failure distribution using a reasonably good statistical sample size. With this being the only available source of data, we fit our percolation based endurance model to this test data in Fig. 12 and extracted the values of E_{aD} ~ 0.30-0.35 eV (for HfO_2), t_{ox} = 3.5 nm (post scavenging dielectric thickness), T_{ps} = 30 ns, V_{ps} = 1.66V and T ~ 380K (high I_{comp}). The fit demonstrates very good applicability of our model to experimental data and the value of E_{aD} is close to theoretical predictions [**18**] and the estimated value of V_{ps} and T_{ps} are close to the design value chosen for the endurance pulse test algorithm.

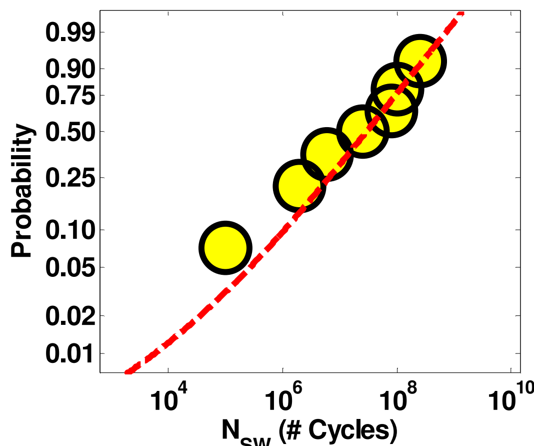

Figure 12. Verification of the proposed endurance failure model with limited statistically significant electrical test data set available in the literature [**8**].

VI. Inferences and Conclusions

Noting that endurance degradation is a key failure mechanism for resistive switching memory devices, we developed a generic, physical and comprehensive statistical model to describe endurance failure using the percolation cell framework as the basis. The developed model was used to investigate the role of various material, operational and architectural factors such as E_{aD}, t_{ox}, T_{FIL}, I_{comp}, filament size, shape, aysmmetry and $\{V_{ps}, T_{ps}\}$ on the shape of the endurance failure distribution and the cycle lifetime. The presence of gradual degradation (wear-out) in the E_{aD} and / or n value helped steepen the distribution, so as to reflect the realistic spread in endurance cycles better. The model was able to fit experimental test data reasonably well and it helped explain the trend of lowering endurance for thin dielectrics and lower I_{comp}.

We observe a trade-off in the endurance reliability and switching power criteria. While low power switching and/or ultra-thin dielectric based memory device is desired, it results in a substanial lowering of the vacancy and oxygen ion count in the LRS, leading to very poor endurance. Engineering the shape of the conductive filament was also shown to have a drastic influence on the endurance lifetime with asymmetric shaped ones faring batter. All practical values of endurance lifetime observed during electrical tests is confined within a short range of E_{aD} values between 0.25eV and 0.45eV.

There are lot more issues to be addressed before the endurance reliability study can be considered holistic. Firstly, we need more experimental data on endurance with large statistical sample size for a wide range of test conditions. This information is needed to calibrate our model and to also extrapolate the endurance failure probability for different DC and AC /pulsed operating conditions. Secondly, we have to examine the variability in endurance trends and find out how this variability can be tweaked and lowered using different test conditions and material design options. This requires the implementation of a Kinetic Monte Carlo (KMC) algorithm to simulate the reset dynamics. Thirdly, we have to assess how and in what way do the measures used to improve endurance lifetime affect retention and read disturb in a positive or negative way. The correlation of endurance, retention and read disturb events (all of which depend on the shape and size of the filament) cannot be ignored and any attempt to design for reliability should collectively assess the impact on all these three potential failure mechanisms in RRAM.

Next, we need to use physical analysis tools to better image the conductive filament in the SET state for an initial device and an endurance degraded one to obtain solid evidence that supports our hypothesis of gradual dilation of filament size and to also confirm the shape of the filament. Moreover, we have not considered the possibility of failure in the HRS state ('failure to SET') where some studies (Ref. [**8**]) have shown that the device ceases to percolate during the cycling process unless the V_{SET} is increased further (risking the possibility of hard breakdown). Finally, OxRAM is only one of the switching device type for resistance switching. The conducting bridge random access memory (CBRAM) [**20**] which operates by

repeated nucleation and rupture of a metallic filament that originates from the anode and involves electrochemical ionic migration has not been discussed here. The physics of switching in CBRAM is completely different and warrants other models (non-percolative) that can be used to describe the kinetics better. We have to develop endurance models that are suited to CBRAM in the near future and carry out a similar study.

ACKNOWLEDGMENT

The authors would like to thank the International Design Center (IDC) at the Singapore University of Technology and Design (SUTD) for fully funding this research work and provision of the travel allowance under Grant No. IDG11300103.

REFERENCES

[1] H.T. Lue, T.H. Hsu, Y.H. Hsiao, S.P. Hong, M.T. Wu, F.H. Hsu, N.Z. Lien, S.Y, Wang, J.Y. Hsieh, L.W. Yang, T. Yang, K.C. Chen, K.Y. Hsieh and C.Y. Lu, "A highly scalable 8-layer 3D vertical-gate (VG) TFT NAND flash using junction-free buried channel BE-SONOS device", *2010 Symposium on VLSI Technology (VLSIT),* pp. 131-132, (2010).

[2] E.S. Choi, H.S. Yoo, H.S. Joo, G.S. Cho, S.K. Park and S.K. Lee, "A novel 3D cell array architecture for terra-bit NAND flash memory", *3rd IEEE International Memory Workshop (IMW),* pp. 1-4, (2011).

[3] T.R. Oldham, M.R. Friendlich, H.S. Kim, M.D. Berg, K.A. LaBel, S.P. Buchner, D. McMorrow, D.G. Mavis, P.H. Eaton and J. Castillo, "Radiation and reliability concerns for modern non-volatile memory technology", (2011).

[4] Y.S. Shin, "Non-volatile memory technologies for beyond 2010", *IEEE Symposium on VLSI Circuits*, pp. 156-159, (2005).

[5] W. Zhao, T. Lionel, G. Yoann, C.V. Luís, L. Yahya, K.O. Jacques, R. Dafine, S. Gilles and C. Claude, "Design of MRAM based logic circuits and its applications", *Proceedings of the 21st edition of the Great Lakes Symposium on VLSI*, pp. 431-436. ACM, (2011).

[6] H. Akinaga and H. Shima, "Resistive random access memory (ReRAM) based on metal oxides", *Proceedings of the IEEE*, Vol. 98, No. 12, pp. 2237-2251, (2010).

[7] B. Chen, Y. Lu, B. Gao, Y.H. Fu, F.F. Zhang, P. Huang, Y.S. Chen, L.F. Liu, X.Y. Liu, J.F. Kang, Y.Y. Wang, Z. Fang, H.Y. Yu, X. Li, X.P. Wang, N. Singh, G.Q. Lo and D.L. Kwong, "Physical mechanisms of endurance degradation in TMO-RRAM", *IEEE International Electron Devices Meeting (IEDM)*, pp.283-286, (2011).

[8] Y.Y. Chen, R. Degraeve, S. Clima, B. Govoreanu, L. Goux, A. Fantini, G.S. Kar, G. Pourtois, G. Groeseneken, D.J. Wouters and M. Jurczak, "Understanding of the endurance failure in scaled HfO_2-based 1T1R RRAM through vacancy mobility degradation", *IEEE International Electron Devices Meeting (IEDM),* pp.20.3.1-20.3.4, (2012).

[9] B. Chen, J.F. Kang, B. Gao, Y.X. Deng, L.F. Liu, X.Y. Liu, Z. Fang, H.Y. Yu, X.P. Wang, G.Q. Lo and D.L. Kwong, "Endurance degradation in metal oxide-based resistive memory induced by oxygen ion loss effect," *IEEE Electron Device Letters*, Vol. 34, No. 10, pp. 1292-1294, (2013).

[10] B. Gao, H. Zhang, B. Chen, L.F. Liu, X.Y. Liu, R. Han, J.F. Kang, Z. Fang, H.Y. Yu, B. Yu and D.L. Kwong, "Modeling of retention failure behavior in bipolar oxide-based resistive switching memory", *IEEE Electron Device Letters,* Vol. 32, No. 3, pp. 276-278, (2011).

[11] L. Zhang, R. Huang, Y.Y. Hsu, F.T. Chen, H.Y. Lee, Y.S. Chen, W.S. Chen, P.Y. Gu, W.H. Liu, S.M. Wang, C.H. Tsai, M.J. Tsai and P.S. Chen, "Statistical analysis of retention behavior and lifetime prediction of HfO_x-based RRAM", *IEEE International Reliability Physics Symposium (IRPS),* pp. MY.8.1-MY.8.5, (2011).

[12] W.C. Luo, J.C. Liu, H.T. Feng, Y.C. Lin, J.J. Huang, K.L. Lin and T.H. Hou, "RRAM SET speed-disturb dilemma and rapid statistical prediction methodology", *IEEE International Electron Devices Meeting (IEDM),* pp. 9.5.1 – 9.5.4, (2012).

[13] J.H. Stathis, "Percolation models for gate oxide breakdown", *Journal of Applied Physics*, Vol. 86, No. 10, pp.5757-5766, (1999).

[14] J. Suñé, "New physics-based analytic approach to the thin-oxide breakdown statistics", *IEEE Electron Device Letters,* Vol. 22, No. 6, pp.296-298, (2001).

[15] D.B. Strukov and R.S. Williams, "Exponential ionic drift: fast switching and low volatility of thin-film memristors", *Applied Physics A*, Vol. 94, No. 3, pp.515-519, (2009).

[16] S. Yu, Y.Y. Chen, X. Guan, H.S.P. Wong and J.A. Kittl, "A Monte Carlo study of the low resistance state retention of HfO_x based resistive switching memory", *Applied Physics Letters*, Vol. 100, No. 4, 043507, (2012).

[17] X. Guan, S. Yu and H.S.P. Wong, "On the Switching Parameter Variation of Metal-Oxide RRAM—Part I: Physical Modeling and Simulation Methodology", *IEEE Transactions on Electron Devices,* Vol. 59, No. 4, pp.1172-1182, (2012).

[18] K. McKenna and A. Shluger, "The interaction of oxygen vacancies with grain boundaries in monoclinic HfO_2", *Applied Physics Letters*, Vol. 95, No. 22, 222111, (2009).

[19] B. Butcher, S. Koveshnikov, D.C. Gilmer, G. Bersuker, M.G. Sung, A. Kalantarian, C. Park, R. Geer, Y. Nishi, P.D. Kirsch and R. Jammy, "High endurance performance of 1T-1R HfOx based RRAM at low (<20μA) operative current and elevated (150°C) temperature", *IEEE International Integrated Reliability Workshop (IIRW),* pp. 146-150, (2011).

[20] Fujiwara, T. Nemoto, M.J. Rozenberg, Y. Nakamura and H. Takagi, "Resistance switching and formation of a conductive bridge in metal/binary oxide/metal structure for memory devices", *Japanese Journal of Applied Physics (JJAP)*, Vol. 47, No. 8, pp.6266-6271, (2008).