Extended Abstracts of the 2014 International Conference on Solid State Devices and Materials, Tsukuba, 2014, pp408-409

Scaling of Resistive Switching Devices

Daniele Ielmini

Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano and IU.NET
Piazza L. da Vinci 32 – 20133 Milano, Italy
Phone: +33-02-2399-6120 E-mail: daniele.ielmini@polimi.it

1. Introduction

Resistive switching memory (RRAM) is receiving large interest as future memory technology for possible use in nonvolatile, embedded and storage class memories [1,2]. The primary requirement for memory application is scalability, where (i) the device size must be scaled down for several technology nodes even below 10 nm, and (ii) the device must allow integration in 3D architectures, to compete with recently introduced 3D Flash technologies [3]. RRAM scaling was addressed from both the technology point of view, demonstrating switching functionality in 10 nm technology [4], and a theoretical point of view, where the atomistic size of the conductive filament (CF) was speculated based on quantum mechanical tunneling calculations [5]. However, the reduction of device and CF sizes has deep reliability implications, which must be carefully addressed for prediction of the RRAM scaling.

This work addresses RRAM scaling from the points of view of energy consumption and size-dependent reliability. Reliability is discussed in terms of switching variability and current fluctuations, both affecting the statistical distributions of resistance R in the RRAM array. Variability is shown to increase at decreasing size of the CF, as a result of the Poisson distribution of the number of defects in the CF. The fluctuation amplitude is also shown to increase at decreasing CF size, which is studied through random telegraph noise (RTN) in RRAM devices. The size-dependent reliability issues can be solved by careful system engineering, such as program-verify operations, and material engineering, to achieve the largest intrinsic resistance window through accurate selection and design of the material stack.

2. Energy scaling of RRAM switching

Nonvolatile memories mainly find application in portable systems, such as tablets and smartphones, thus energy consumption is among the main concerns for memory performance. To reduce the energy of RRAM switching, it is necessary to control the current in both the set and reset operations, while minimizing the programming time. Current reduction can be achieved by controlling the forming and set operations through a MOS transistor in the so-called one-transistor/one-resistor (1T1R) structure [6,7]. Fig. 1 shows the measured current-voltage (I-V) characteristics for a compliance current I_C = 80 μA (a) and 8 μA (b) during the set operation. The RRAM device consisted of a $HfSiO_x$ layer included in TiN electrodes with a Ti cap to serve as oxygen exchange layer [8]. The voltage in the figure only includes the drop across the memory element within the 1T1R structure. Note that the reset current I_{reset} remains in the same range of I_C, namely $I_{reset} = I_C$, thus ensuring a 10x reduction of switching energy from (a) to (b). The smaller current is due to the reduced size of the CF, which in fact displays an increased resistance in Fig. 1b.

Although beneficial for energy control, the CF size reduction leads to a degradation of the switching variability, as shown by the enhanced scattering of I-V curves around the median characteristic in Fig. 1b. This is due to the smaller number of defects causing a larger statistical variability of the number and position of defects in smaller CFs [9]. Fig. 2a shows the relative spread of set-state resistance, defined as the ratio between the standard deviation σ_R and its average value μ_R. Note that σ_R/μ_R increases at decreasing I_C, supporting the size-dependent variability due to discrete defect injection. The figure also shows the relative spread of I_{reset}, namely $\sigma_{Ireset}/\mu_{Ireset}$ (b) and the standard deviation of the reset voltage σ_{Vreset} (c), both indicating an increased variability at decreasing I_C.

3. Modeling

To study the tradeoff between energy reduction and statistical variability and noise, models were developed for the conduction and switching in presence of discrete defects in the device. Switching was described by an analytical model for CF growth and retraction induced by the local temperature and field [10]. Discrete defect injection was described by a Monte Carlo model where each defect (or defect cluster) was assigned a different energy barrier for injection in both set and reset. Calculation results in the figure can account for the size-dependence of switching variability, due to the larger spread of energy barrier at decreasing number of defects in the CF [10]. To address the increased variabil-

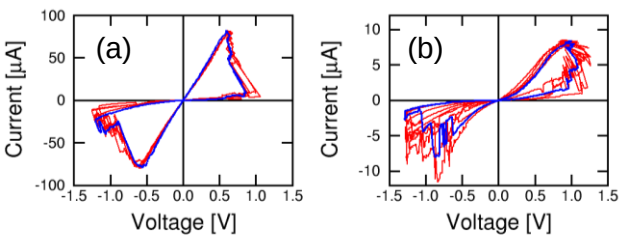

Fig. 1 Measured and calculated I-V curves in a $HfSiO_x$-based RRAM at I_C = 80 μA (a) and 8 μA (b).

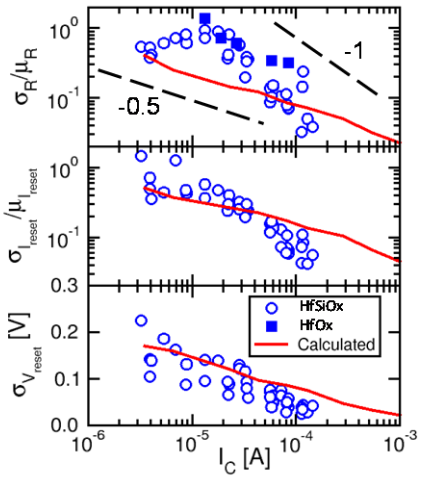

Fig. 2 Measured and calculated σ_R/μ_R (a), $\sigma_{Ireset}/\mu_{Ireset}$ (b) and σ_{Vreset}. Variability increases at decreasing size of the CF.

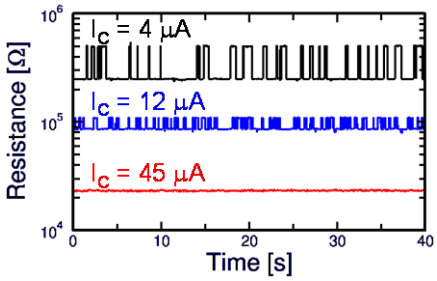

Fig. 3 Measured R as a function of time at variable I_C, indicating size-dependent RTN.

ity at small CF sizes, set/reset operations should be performed through advanced program-verify loops. Also, the resistance window in the device should be improved through a careful design of the oxide stack.

The reduced CF size also leads to a larger amplitude of current noise, such as RTN [10]. Fig. 3 shows the measured resistance as a function of time for a RRAM device affected by RTN and programmed in the set state with 3 different values of I_C. Noise amplitude increases at decreasing I_C, which can be explained by the increased impact of a charged defect on the carrier concentration in the CF. This effect was described by a numerical model capable of solving the Poisson equation in presence of charged defects at the surface of the CF [10]. Fig. 4 shows the calculated density of carriers in the CF for diameter ϕ = 10 nm (a) and 1 nm (b). a negatively-charged defect was assumed at the surface of the CF, resulting in partial (a) or full depletion (b) of electrons. Full depletion leads to a large RTN amplitude, which accounts for the size-dependent RTN in Fig. 3. Fig. 5 shows the measured and calculated relative change of resistance $\Delta R/R$ as a function of R. As R increases, $\Delta R/R$ increases due to the transition from partial to full depletion at decreasing CF size, thus highlighting the increasing impact of RTN in downscaled RRAM technologies.

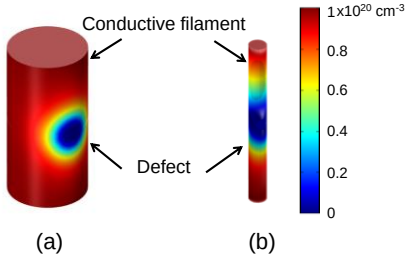

Fig. 4 Calculated carrier concentration for ϕ = 10 nm (a) and 1 nm (b), indicating partial and full depletion, respectively.

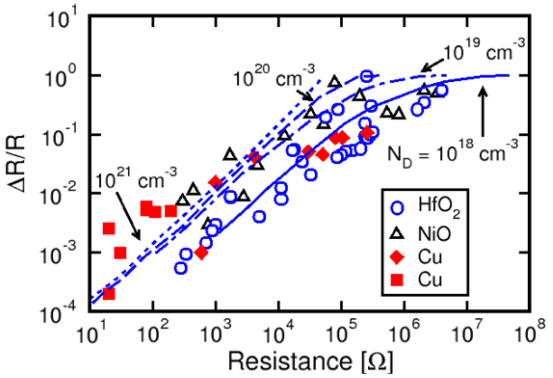

Fig. 5 Measured and calculated $\Delta R/R$ for RTN as a function of R. As R increases, the RTN amplitude increases due to the transition from partial to full depletion. Different materials indicate similar behaviors.

4. Conclusion

RRAM scaling was discussed in terms of the tradeoff between energy reduction and reliability. CF size reduction allows for reduced operation current, however at the expense of an enhanced switching variability and noise amplitude. These effects can be predicted by analytical and numerical models of localized switching /conduction.

Acknowledgements

We acknowledge SEMATECH and Micron for providing samples and collaboration. We are also indebted with S. Ambrogio, S. Balatti, K. Min, K. Hasnat, A. Calderoni and N. Ramaswamy for several discussions.

References

[1] T.-Y. Liu, et al., *ISSCC Tech. Dig.* (2013) 210.
[2] S. Sills, et al., *Symp. VLSI Tech. Dig.* (2014).
[3] H. Tanaka, et al., *Symp. VLSI Tech. Dig.* (2007) 14.
[4] B. Govoreanu, et al., *IEDM Tech. Dig.* (2011) 729.
[5] V. Zhirnov, et al., Nanotechnology **22** (2011) 254027.
[6] H. Y. Lee, et al., *IEDM Tech. Dig.* (2008) 297.
[7] F. Nardi, et al., Solid State Electronics **58** (2011) 42.
[8] A. Calderoni, et al., Proc. *IMW* (2014) 5.
[9] S. Balatti, et al., IEEE EDL **34** (2013) 861.
[10] S. Ambrogio, et al., *IEDM Tech. Dig.* (2013) 782.