Mechanism of electrical shorting failure mode in resistive switching

Andrew J. Lohn,[a] Patrick R. Mickel, and Matthew J. Marinella
Sandia National Laboratories, Albuquerque, New Mexico 87185, USA

(Received 29 May 2014; accepted 8 July 2014; published online 17 July 2014)

The electrical shorting failure mode in resistive switching is characterized by the inability to increase the resistance electrically and is one of the most common failures observed in these devices. We show that vacancy accumulation at the inert electrode is a likely cause of the electrical shorting failure mode. A detailed description is provided of the specific effect of injected oxygen vacancies from the reactive electrode and from the secondary reservoir that is formed at the inert electrode during an electrical shorting failure. We present quantitative theoretical and experimental analysis of the failure mechanism while suggesting approaches and conditions for prevention and recovery. The approach also provides an analytical description of sub-saturation vacancy injection during normal operation while experimentally showing the range of conditions where this behavior dominates. © *2014 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4890635]

I. INTRODUCTION

Despite demonstrating exceptional performance[1–3] and very large information density per device,[4] commercial resistive random access memories (RRAM) are still not widely available. This delay has been, in large part, due to the difficulty in studying the switching filament directly.[5,6] It is commonly believed that, in bipolar switching, oxygen vacancies are injected from a reactive electrode into a conducting filament to decrease its resistance and that those oxygen vacancies are pulled back into the reactive electrode to increase the filament resistance.[7,8] A more detailed quantitative understanding of the mechanisms and behavior of vacancy injection is needed to understand device operation and also to improve reliability. Using terminology borrowed from the transistor gate oxide community, RRAM operation is described as repeated soft breakdown where the resistance can be increased following the breakdown events.[9,10] Often, after repeated switching, the device transitions from soft breakdown to hard breakdown, where the resistance can no longer be increased electrically. This electrical shorting failure mode is the most common acute failure mode we observe in these devices. In moving toward commercialization, a clear understanding of the physics leading to failure is needed so that these types of failures can be avoided or eliminated. Since it is difficult to directly observe the subtle changes in density of an oxidation state over a few nanometers within an enclosed system, approaches that are able to infer information about the structure and chemistry of the device electrically are potentially of much interest.

Many types of mechanisms have been proposed to explain behavior in RRAM devices,[11,12] but few provide a quantitative approach that can be applied to behavior outside that of the normal operating modes that were used to build the model. Recently, solving for heat flow in RRAM has provided a simple analytical model for resistive switching that is quantitatively accurate and can be easily adapted to explain atypical device behavior such as failure.[13] Here, using this model and analytical framework, we demonstrate that this model is capable of identifying the regions of a current-voltage (I-V) hysteresis loop that correspond to sub-saturation vacancy injection and the regions that correspond to post-saturation vacancy injection. We also show that bipolar vacancy injection is a likely physical mechanism for the electrical shorting failure mode. We quantitatively study the effect of this failure mode on the conducting filament and are therefore able to determine the conditions under which recovery from shorting failure is possible and impossible.

II. EXPERIMENT AND THEORY

This study uses tantalum oxide based RRAM because these are among the most likely candidates to achieve commercial success.[5,14,15] We observe shorting failure behavior in a variety of device designs (both CMOS compatible and incompatible) with no significant difference identified between the electrode materials, indicating that these effects are characteristic of the switching filament and not the specifics of manufacturing. The metal-insulator-metal structures are made with a reactive tantalum electrode, a tantalum oxide insulator[16,17] ranging from 5 to 15 nm in thickness, and an inert electrode. Our CMOS compatible devices are fabricated with titanium nitride inert electrodes[18] as small as 350 nm lateral dimensions, and our CMOS incompatible devices use platinum[19] as an inert electrode with lateral dimensions ranging from tens to hundreds of μm. The platinum designs have much larger series resistance and capacitance, leading to larger stored charges that could cause an increased current in a shorting event, but also a slower time constant which may reduce the discharge current.

The analytical model[13] used in this study determined that, over a large range of resistance values, ON and OFF switching are governed by separate state variables, namely, the radius and conductivity of the conducting filament, respectively. During ON switching (lowering device resistance), the conductivity of the filament is constant while the filament radius increases. In contrast, during OFF switching

[a]ajlohn@sandia.gov

(increasing device resistance), the filament radius is constant while the conductivity is decreased. The equations corresponding to these two switching modes relate the electrical power to the resistance of the device for either a changing conductivity or a changing radius. From those equations, it is simple to derive the current-voltage relation for either case

$$I_r = \frac{V_r - A_r(T_{crit} - T_{RT})V_r^{-1}}{R_{min}}, \tag{1a}$$

$$I_\sigma = \frac{V_\sigma + A_\sigma(T_{crit} - T_{RT})V_\sigma^{-1}}{R_{max}}. \tag{1b}$$

The subscript r (radius) or σ (conductivity) indicates which variable is allowed to change. T_{crit} is the critical temperature for ion motion and T_{RT} is the ambient temperature of the device. The factors A_r, A_σ, R_{min}, and R_{max} are constants whose values can be calculated from materials and device design parameters: $A_r = \frac{2k_E d_O}{\sigma_{max} d_E}$, $A_\sigma = \frac{8d_o^2 L_{WF} T_{crit}}{r_{max}^2}$, $R_{min} = \frac{k_E}{4\pi\sigma_{max}^2 L_{WF} T_{crit} d_E}$, and $R_{max} = \frac{4d_o^2 L_{WF} T_{crit} d_E}{\pi r_{max}^4 k_E}$. In the above equations, V is voltage, d_E is electrode thickness, k_E is electrode thermal conductivity, d_O is thickness of the oxide insulator, r_F is filament radius, k_F is filament thermal conductivity, r_{max} is the maximum previous filament radius, σ_{max} is the conductivity corresponding to the saturation concentration, and L_{WF} is the Wiedemann-Franz constant.

In normal operation, the mechanism for resistive switching that is described by the above equations has three phases starting from the highly resistive state: (1) Resistance decreases by injecting oxygen vacancies from the reactive electrode into a small localized filament, thereby increasing the conductivity of the localized filament. (2) Once the filament conductivity saturates, injected vacancies must go to the edge of the filament where saturation has not yet been reached. This stage is depicted in Fig. 1(a), where the resistance decreases by increasing the radius of the filament. (3) To decrease the conductivity, the polarity of the applied field is reversed and oxygen vacancies are pulled into the reactive electrode. This stage is depicted in Fig. 1(b), where resistance is increased by decreasing the conductivity of the filament. The physical explanation for bipolar resistive switching is the asymmetry of the reactivity of the electrodes. The reactive electrode acts as a source of vacancies that can flood into the filament to increase its radius (Fig. 1(a)). In the opposite polarity, the reactive electrode acts as a sink of vacancies to empty the filament and decrease conductivity (Fig. 1(b)). If however, there were a strong flux of vacancies causing an accumulation of vacancies at the inert electrode, then both polarities would be capable of injecting vacancies into the filament. In that case, the radius would increase (Figs. 1(c) and 1(d)) for both polarities. In terms of electrical behavior, the device would be in a shorting failure mode because both polarities would cause the resistance to decrease.

III. SUB-SATURATION VACANCY INJECTION

Inspecting the I-V ON switching half cycle shown in Fig. 2 there are four clearly distinguishable sections, each

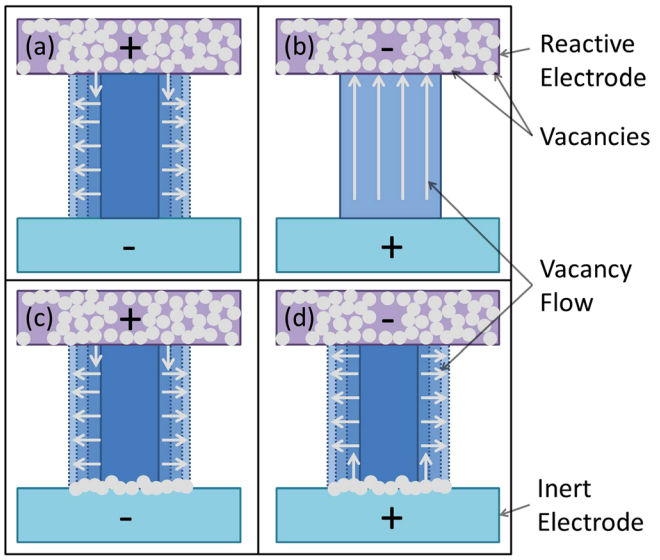

FIG. 1. In normal operation, oxygen vacancies flow from the reactive electrode (vacancy reservoir) to increase the filament radius (a) or into the reactive electrode to decrease the filament conductivity (b). With a reservoir on both sides, radius of the filament increases for both polarities ((c) and (d)) causing the shorting failure.

marked with an arrow indicating the direction of the sweep. No switching occurs in the two sections that intersect the origin since the electrical power is too low to activate switching. The other two sections do involve switching but, notably, the early stage (highlighted in green) has a negative slope and the latter stage has a positive slope. To compare with Fig. 2, we take the first derivative of current with respect to voltage from Eq. (1)

$$\frac{dI_r}{dV_r} = \frac{1}{R_{min}} + \frac{A_r(T_{crit} - T_{RT})}{R_{min}V_r^2}, \tag{2a}$$

$$\frac{dI_\sigma}{dV_\sigma} = \frac{1}{R_{max}} - \frac{A_\sigma(T_{crit} - T_{RT})}{R_{max}V_\sigma^2}. \tag{2b}$$

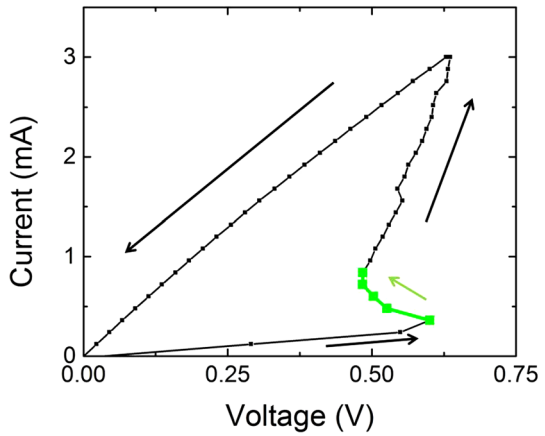

FIG. 2. There is often a negative sloping region at the onset of ON switching that cannot be described by the radius change equations so conductivity change must be used instead. The green highlighted section of the ON switching curve corresponds to the sub-saturation regime, where vacancy injection increases the conductivity of a small filament.

It can be seen that $\frac{dI_r}{dV_r}$ is always positive (Eq. (2a)) since R_{min} and A_r are positive constants, and T_{crit} is always larger than T_{RT}. Therefore, it is not possible to use the radius change equation to describe a negative slope in the I-V relation such as the highlighted section of Fig. 2. However, allowing conductivity to be the changing variable (Eq. (2b)) instead of radius (Eq. (2a)), it is possible to describe either a negative or positive slope in the I-V relation. The conductivity change equation is needed in the early stages of ON switching because this section of the I-V relation corresponds to vacancy injection prior to reaching the saturation concentration. In the early stages of ON switching, conductivity increases until saturation and then radius begins to increase. Once again, the need for using conductivity change instead of radius change can be seen in the first derivatives of current with respect to voltage (Eqs. (2)), showing that radius change cannot account for negative I-V slopes.

IV. SHORTING FAILURE

Converting these equations to back to their original[13] power-resistance (P-R) coordinates, instead of the I-V coordinates described above, provide more insight into the switching mechanisms and failure modes. Taking the data from the I-V relation and multiplying I and V to give P and dividing V by I to give R gives the P-R relation. By taking the first derivative of power with respect to resistance, more limiting behavior becomes apparent

$$\frac{dP_r}{dR_r} = \frac{-A_r(T_{crit} - T_{RT})}{(R - R_{min})^2}, \tag{3a}$$

$$\frac{dP_\sigma}{dR_\sigma} = \frac{A_\sigma(T_{crit} - T_{RT})}{(R_{max} - R)^2}. \tag{3b}$$

According to Eqs. (3a) and (3b), the rate of change of power with respect to resistance can only give negative slope for radius change and can only give positive slope for conductivity change. Data that act otherwise would likely imply that the cylindrical filament approximation used in their derivation was invalidated. The cylindrical filament approximation may not be valid over the entire switching range[4] and so it may be possible, under some conditions, to have a negative P-R slope for changing conductivity with constant radius. However, there is no physical explanation for a decreasing resistance for a negative polarity of applied power within this model, provided that there is an inert electrode that cannot inject vacancies.

Negative polarity is typically used to increase the resistance, but following an electrical shorting failure, both polarities cause resistance reduction. This effect is shown in Fig. 3, where a normally operating device (Fig. 3(a)) became shorted abruptly and attempts were made to use negative polarity power (Fig. 3(b)) to return the device to a high resistance state. Prior to the shorting failure, these negative polarities increased the resistance but, following the shorting failure, both polarities caused decreases in resistance. These decreases in resistance can be easily and quantitatively explained if a source of vacancies accumulates at the inert electrode. In such a case, both polarities would inject vacancies at the saturation concentration which would result in increasing radius. This increasing radius would have a negative slope in P-R space and can be fit by using the radius change equation as shown in Fig. 3(c).

Fitting the data to the radius change equation (Eq. (1a)) in P-R coordinates shows very good agreement (Fig. 3(c)) with the measured data in the failed device. And it is possible to extract quantities such as filament radius (6.6–8.2 nm), saturation conductivity (6500 $\Omega^{-1}\,cm^{-1}$), and activation temperature (1750 K) from the fit. These quantities are in the range of expected values, further supporting the use of the radius change equation for resistance decreases in negative polarity during shorting failure, and supporting the conclusion that the electrical shorting failure mode is caused by the creation of a secondary reservoir at the inert electrode.

Although the device in Fig. 3 was permanently shorted, in some cases it is possible to recover a high resistance state after secondary reservoir creation, provided that the reservoir can be sufficiently depleted. A minimum resistance exists[9] (Eq. (1a)), beyond which sufficient power cannot be supplied to enable switching. Provided that the reservoir can be depleted without reaching the minimum resistance, the radius will increase (decreasing resistance) until the secondary reservoir is depleted. Once the reservoir is depleted, further increases in negative polarity power will cause the conductivity of the filament to decrease (increasing resistance). This

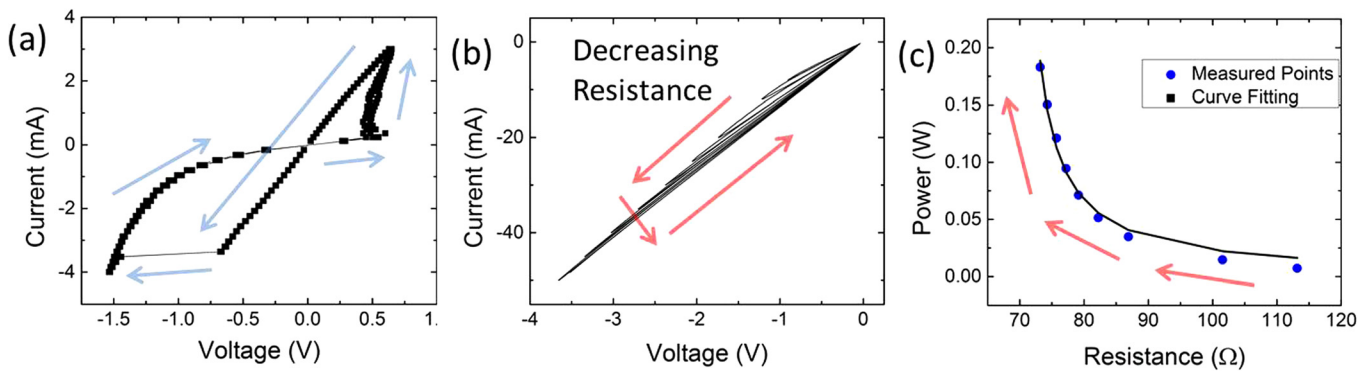

FIG. 3. Hysteresis loops show a working device (a) that became shorted and decreased resistance in both positive and negative polarities (b). Fitting the measured data to the radius change equation (c) in activation power-resistance coordinates shows that radius was increasing for negative polarities.

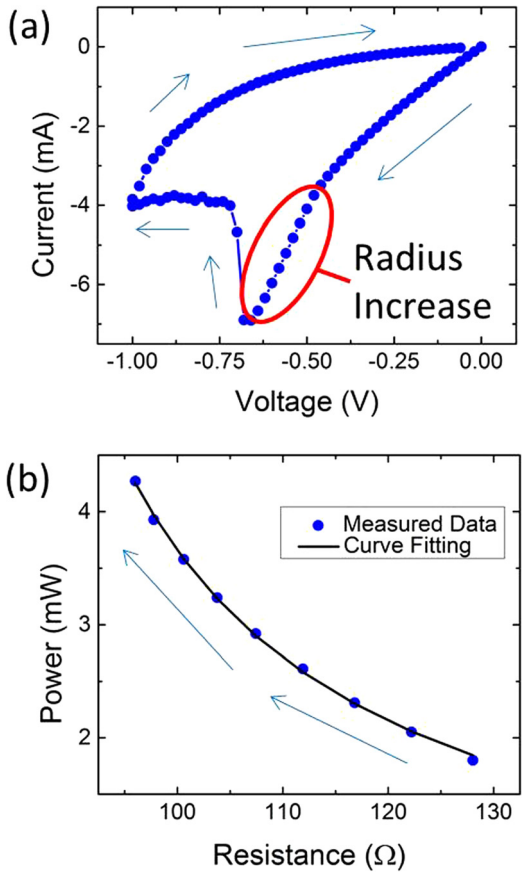

FIG. 4. A shorted device can recover a high resistance state provided that the secondary reservoir can be depleted with less than infinite power. An OFF switching half cycle is shown in (a) where the resistance decreased then increased with negative applied power. (b) The resistance decreasing region is shown in activation power-resistance coordinates and fit to the radius change equation, suggesting that the radius of the filament is increasing while the reservoir is being depleted.

process is shown experimentally in Fig. 4 in I-V coordinates (Fig. 4(a)) and in P-R coordinates (Fig. 4(b)). Once switching is initiated, there is a range where increasing power causes resistance to decrease, then further increases in power cause resistance to increase. The resistance decreasing region is plotted in P-R coordinates (Fig. 4(b)) and fit to the radius change equation. The fit in P-R coordinates for this device is very good and also gives reasonable values for radius (11–12.6 nm), conductivity (2070 $\Omega^{-1}\,cm^{-1}$), and activation temperature (1200 K), again suggesting that vacancy injection from a secondary reservoir at the inert electrode is a likely mechanism for explaining the failure.

V. CONCLUSIONS

Using an analytical approach and electrical characterization of switching and failed devices, we are able to show that the negative slope at the onset of ON switching is due to vacancy injection below saturation and that the common failure mode of electrical shorting is due to secondary reservoir formation at the inert electrode. This failure is characterized by a sign change of the slope in P-R coordinates or generally decreasing resistance in negative polarities. Recovery from this failure mode is possible if the process of depleting the reservoir does not reduce the resistance beyond the minimum resistance which would imply the need for infinite electrical power.

ACKNOWLEDGMENTS

The authors would like to thank James E. Stevens and the support of Sandia's MESA Fab as well as Seth A. Decker for measurement and characterization support. This work was funded by Sandia's Laboratory Directed Research and Development (LDRD) program. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under Contract No. DE-AC04-94AL85000.

[1]A. C. Torrezan, J. P. Strachan, G. Medeiros-Ribeiro, and R. S. Williams, "Sub-nanosecond switching of a tantalum oxide memristor," Nanotechnology **22**, 485203 (2011).
[2]J. P. Strachan, A. C. Torrezan, G. Medeiros-Ribeiro, and R. S. Williams, "Measuring the switching dynamics and energy efficiency of tantalum oxide memristors," Nanotechnology **22**, 505402 (2011).
[3]M.-J. Lee, C. B. Lee, D. Lee, S. R. Lee, M. Chang, J. H. Hur, Y.-B. Kim, C.-J. Kim, D. H. Seo, S. Seo *et al.*, "A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta_2O_{5-x}/TaO_{2-x} bilayer structures," Nat. Mater. **10**, 625–630 (2011).
[4]A. J. Lohn, P. R. Mickel, C. D. James, and M. J. Marinella, "Degenerate resistive switching and ultrahigh density storage in resistive memory," preprint arXiv:1406.4033 (2014).
[5]G.-S. Park, Y. B. Kim, S. Y. Park, X. S. Li, S. Heo, M.-J. Lee, M. Chang, J. H. Kwon, M. Kim, U.-I. Chung, R. Dittmann, R. Waser, and K. Kim, "*In situ* observation of filamentary conducting channels in an asymmetric Ta_2O_{5-x}/TaO_{2-x} bilayer structure," Nat. Commun. **4**, 2382 (2013).
[6]F. Miao, J. P. Strachan, J. J. Yang, M.-X. Zhang, I. Goldfarb, A. C. Torrezan, P. Eschbach, R. D. Kelley, G. Medeiros-Ribeiro, and R. S. Williams, "Anatomy of a nanoscale conduction channel reveals the mechanism of a high-performance memristor," Adv. Mater. **23**, 5633–5640 (2011).
[7]D. Ielmini, "Modeling the universal set/reset characteristics of bipolar RRAM by field- and temperature-driven filament growth," IEEE Trans. Electron Devices **58**, 4309–4317 (2011).
[8]R. Degraeve, A. Fantini, S. Clima, B. Govoreanu, L. Goux, Y. Y. Chen, D. J. Wouters, Ph. Roussel, G. S. Kar, G. Pourtois, S. Cosemans, J. A. Kittl, G. Groeseneken, M. Jurczak, and L. Altimime, "Dynamic 'hour glass' model for SET and RESET in HfO_2 RRAM," Symp. VLSI Technol., Dig. Tech. Pap. **2012**, 75–76.
[9]A. J. Lohn, P. R. Mickel, and M. J. Marinella, "Dynamics of percolative breakdown mechanism in tantalum oxide resistive switching," Appl. Phys. Lett. **103**, 173503 (2013).
[10]S. Long, C. Cagli, D. Ielmini, M. Liu, and J. Sune, "Analysis and modeling of resistive switching statistics," J. Appl. Phys. **111**, 074508 (2012).
[11]P. R. Mickel, A. J. Lohn, B. J. Choi, J. J. Yang, M.-X. Zhang, M. J. Marinella, C. D. James, and R. S. Williams, "A physical model of switching dynamics in tantalum oxide memristive devices," Appl. Phys. Lett. **102**, 223502 (2013).
[12]J. J. Yang, D. B. Strukov, and D. R. Stewart, "Memristive devices for computing," Nat. Nanotechnol. **8**, 13–24 (2013).
[13]P. R. Mickel, A. J. Lohn, C. D. James, and M. J. Marinella, "Isothermal switching and detailed filament evolution in memristive systems," Adv. Mater. **26**, 4486–4490 (2014).
[14]T. Ninomiya, T. Takagi, Z. Wei, S. Muraoka, R. Yasuhara, K. Katayama, Y. Ikeda, K. Kawai, Y. Kato, Y. Kawashima, S. Ito, T. Mikawa, K. Shimakawa, and K. Aono, "Conductive filament scaling of TaOx bipolar ReRAM for long retention with low current operation," Symp. VLSI Technol., Dig. Tech. Pap. **2012**, 73–74.
[15]G. Medeiros-Ribeiro, J. H. Nickel, and J. J. Yang, "Progress in CMOS-memristor integration," in *Proceedings of International Computer-Aided Design (ICCAD)* (2011), pp 246–249.

[16]A. J. Lohn, J. E. Stevens, P. R. Mickel, and M. J. Marinella, "Optimizing TaO_x memristor performance and consistency within the reactive sputtering "forbidden region"," Appl. Phys. Lett. **103**, 063502 (2013).
[17]J. E. Stevens, A. J. Lohn, S. A. Decker, B. L. Doyle, P. R. Mickel, and M. J. Marinella, "Reactive sputtering of substoichiometric Ta_2O_x for resistive memory," J. Vac. Sci. Technol., A **32**, 021501 (2014).
[18]A. J. Lohn, J. E. Stevens, P. R. Mickel, D. R. Hughart, and M. J. Marinella, "A CMOS compatible, forming free TaO_x ReRAM," ECS Trans. **58**, 59–65 (2013).
[19]J. J. Yang, M.-X. Zhang, J. P. Strachan, F. Miao, M. D. Pickett, R. D. Kelley, G. Medeiros-Ribeiro, and R. S. Williams, "High switching endurance in memristive devices," Appl. Phys. Lett. **97**, 232102 (2010).