JOURNAL OF APPLIED PHYSICS **115**, 234507 (2014)

Analytical estimations for thermal crosstalk, retention, and scaling limits in filamentary resistive memory

Andrew J. Lohn,[a] Patrick R. Mickel, and Matthew J. Marinella
Sandia National Laboratories, Albuquerque, New Mexico 87123, USA

(Received 6 May 2014; accepted 13 June 2014; published online 19 June 2014)

We discuss the thermal effects on scaling, retention, and error rate in filamentary resistive memories from a theoretical perspective using an analytical approach. Starting from the heat equation, we derive the temperature profile surrounding a resistive memory device and calculate its effect on neighboring devices. We outline the engineering tradeoffs that are expected with continued scaling, such as retention and power use per device. Based on our calculations, we expect scaling to continue well below 10 nm, but that the effect of heating from neighboring devices needs to be considered for some applications even at current manufacturing capabilities. We discuss possible designs to alleviate some of these effects while further increasing device density. © *2014 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4885045]

I. INTRODUCTION

Shrinking device dimensions have created powerful computers with immense storage capacity. This pace of progress has been the cornerstone of much technological and commercial development but has become increasingly difficult to continue. Particularly in memory, new non-transistor devices[1–3] are being given substantial attention as a means to prolong this progress. Among the leading candidates is resistive random access memory (RRAM)[4] or memristors[5,6] because of their non-volatility,[6] speed,[7] endurance,[8] low power operation,[9] and their ability to store immense numbers of states in a single device.[10] They have also been shown to operate at very small device dimensions (<10 nm),[11] which is promising, but the ultimate scalability of these devices is still not known. Understanding the limits of these filaments and the interactions between neighboring filaments is critical to understanding the ultimate scaling potential, and the performance tradeoffs associated with continued miniaturization. This understanding is difficult to achieve experimentally because the active regions are nanometers in dimensions, switching occurs on sub-nanosecond timescales, and they are necessarily encased in an electrode material. In this work, we adopt an analytical approach to describe the filament and its surrounding region, leading to insights about engineering tradeoffs such as scaling, retention, and power consumption.

Recently, an analytical framework was introduced that provides a detailed description of the properties inside the conducting filament, including its radius, composition, and its thermal properties.[12] Using this analytical framework allows for the design of devices to optimize device dimensions and temperatures within the device. Using a similar approach, it is possible to analytically determine the heat flow and temperatures surrounding the device as well. These temperatures will determine the retention of the device itself, and for dense memory arrays, the temperature surrounding the device may affect the retention of neighboring devices[13] due to thermal cross-talk. The results presented here were derived from bipolar switches but are likely to be applicable to unipolar switches as well due to the purely thermal nature of the approach.

The analytical framework for describing the filament itself is summarized in the following equations that are derived for a cylindrical conducting filament:[12]

$$IV_r = A_r \frac{T_{crit} - T_{RT}}{R - R_{min}}, \tag{1a}$$

$$IV_\sigma = A_\sigma \frac{T_{crit} - T_{RT}}{R_{max} - R}, \tag{1b}$$

where $A_r = \frac{2k_z d_O}{\sigma_{max} d_z}$, $A_\sigma = \frac{8 d_o^2 L_{WF} T_{crit}}{r_{op}^2}$. In these equations, R is resistance, and I and V are current and voltage, respectively, where the subscripts r and σ denote whether resistance changes are by modulated by changing the filament radius (r) or the filament conductivity (σ) with the other held constant. The variable r_{op} is the operating radius (discussed later), σ_{max} is the saturation conductivity, L_{WF} is the Wiedemann-Franz constant, T_{RT} and T_{crit} are room temperature and the critical temperature for activation of ion motion, respectively. The variable d_o is the oxide thickness, and $\frac{d_z}{k_z}$ is the electrode thickness divided by electrode thermal conductivity. The factor $\frac{d_z}{k_z}$ is used to approximate the thermal resistance per unit area for heat traveling from the filament through the electrode.

II. FILAMENT RADIUS AND DEVICE PERFORMANCE

There are at least three different types of radii that must be considered within the device. They are depicted in Fig. 1 and are referred to as filament radius (r_f), operating radius (r_{op}), and maximum radius (r_{max}). The maximum radius (r_{max}) is the most easily understood. It is the largest radius that a device is able to achieve. It corresponds to the minimum resistance which has been observed experimentally[14] and predicted theoretically[12] by the following equations:

[a]Author to whom correspondence should be addressed. Electronic mail: drewlohn@gmail.com

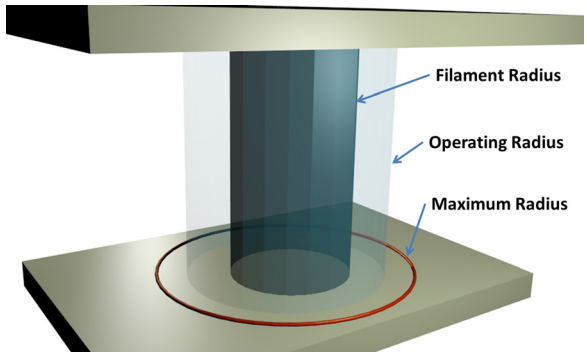

FIG. 1. There are three different radii of interest. The maximum radius (r_{max}) is the largest radius that the conducting filament can be expanded to. The operating radius (r_{op}) is the largest radius that the device has been expanded to. The filament radius (r_f) is the radius that the dominant pathway is currently expanded to. Transparency is used to indicate decreasing conductivity.

$$R_{min} = \frac{k_z}{4\pi\sigma_{max}^2 L_{WF} T_{crit} d_z}, \tag{2a}$$

$$R_{max} = \frac{4d_o^2 L_{WF} T_{crit} d_z}{\pi r_{op}^4 k_z}. \tag{2b}$$

These two values bound the resistance but do not indicate a lower bound on device dimension. They do however lead to an upper bound for filament radius (r_{max}), as can be derived from Eq. (2a)

$$r_{max} = \sqrt{4L_{WF} T_{crit} d_o \sigma_{sat} \frac{d_z}{k_z}}. \tag{3}$$

The device does not need to be operated at its maximum radius though and, for a number of reasons, doing so would lead to sacrifices in device performance as will be discussed later. Instead, we define an operating radius (r_{op}) that can be designed for and used in our expressions to estimate device performance. A simple way to think about the operating radius is that it is the radius at which the conductivity is high enough that, particularly in the high resistance state, current conduction contributes to heat generation in the device. The radius of this region is often determined by the forming step and remains at a slightly elevated conductivity even after OFF switching (resistance increasing), as evidenced by the decreased resistance of the OFF state as compared to the as-deposited film. When the device resistance is decreased, a region of high conductivity forms within the device which can be smaller than the operating radius. The edge of this region is the filament radius (r_f) which we define as the radius of the dominant conduction pathway. In order to keep the radius well below the maximum radius, the forming power and subsequent ON switching (resistance reduction) powers should all be kept low to suppress increases in the radius during operation. It is worth mentioning that this theory was developed using TaO_x-based devices for which the vacancy concentration is believed to be continuously tunable. In devices where changing vacancy concentration involves phase changes between specific metastable states the conductivities may become discretized, but similar definitions for radius should be applicable.

Using the above definitions for radii, we can develop expressions for a large number of device properties. To provide reasonable and approximate values, we use parameters measured from the CMOS-compatible devices fabricated in our laboratory.[15,16] The maximum radius can be calculated from Eq. (3) by combining the minimum resistance value and σ_{sat} which, using the analytical framework derived in Ref. 12, can be measured from a current-voltage sweep of the device. For our devices, the maximum radius is approximately 11 nm. It seems intuitive that one would attempt to decrease the maximum radius to improve scaling but we argue that is not a good approach.

It should be stated clearly that the maximum radius is not a bound on the scaling of the device; it is a point where further scaling may require implementing limits on the operating range (i.e., it is not necessary to operate up to the maximum radius during switching). Altering design parameters to reduce the maximum radius is possible but would lead to sacrifices in performance. The impact of altering design parameters in Eq. (3) (T_{crit}, d_o, σ_{sat}, $\frac{d_z}{k_z}$) on the maximum radius is shown in Fig. 2(a). The parameters are varied from 1/3 to 3 times their initial value. This range of controllable variability may not be achievable for all of the parameters but a single range is shown for ease of comparison. Also, since changes in any one of the four parameters have the same effect on r_{max}, the figure shows the effect of combining parameters (i.e., the effect of changing 1, 2, 3, or all 4 parameters simultaneously). The figure indicates that for increased scaling (i.e., decreased device dimension), each of the design parameters should be decreased but we argue that this is not true. It is not necessary to operate the device at the full maximum possible radius and we would like to state that for improved performance, these design parameters should often be modulated in the opposite way.

From a device performance perspective, a primary concern for RRAM implementation is reduction of power consumption or the ability to have large fan-in which both often require increasing the device resistances. Reducing the operating radius is a straightforward and effective approach to achieving that goal because the maximum resistance (R_{max}) is inversely proportional to the fourth power of the operating radius as can be seen from Eq. (2b). It is important to note that this maximum resistance is measured at the switching voltage. Due to nonlinearity in the current-voltage relation, the maximum resistance at the smaller READ voltage is likely to be much higher. Throughout this document, we will refer to the maximum resistance with the understanding that it is measured at the SET voltage. This maximum resistance is shown in Fig. 2(b) for design parameters in Eq. (2b) varied from 1/3 to 3 times their initial value. It can be seen that increasing the design parameters (T_{crit}, d_o, $\frac{d_z}{k_z}$) gives improved performance in terms of large maximum resistance values and that operating at radii less than r_{max} also increases maximum resistance. This is in contrast to the result in Eq. (3) and Fig. 2(a) where increasing the same parameters resulted in increased $r_{\max}$. This is an example of how designing to decrease r_{max} would sacrifice device performance but this is not an engineering tradeoff because one can simply

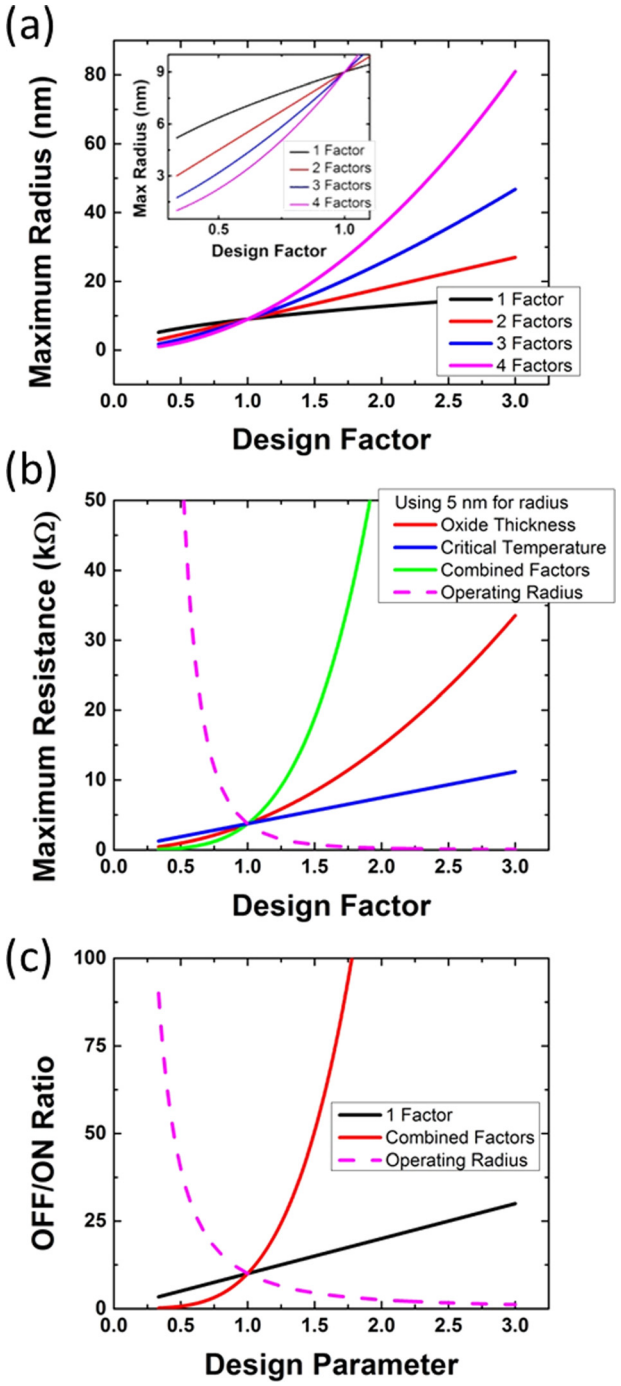

FIG. 2. The effects of varying device design parameters over a range of factors from 1/3 to 3 are shown for (a) the maximum radius, (b) the maximum resistance, and (c) the OFF/ON ratio. Increasing the parameters results in improved performance (b) and (c) but increased maximum radius (a); however, the devices do not need to be operated at their maximum radius. Operating at less than the maximum radius further improves performance (dotted curves in (b) and (c)).

operate the device at radii less than r_{max} and also benefit in performance (i.e., higher resistance).

Aside from increasing device resistance, there are many other performance metrics of interest, including increasing the ratio of OFF state to ON state resistances. The $\frac{R_{OFF}}{R_{ON}}$ ratio is inversely proportional to the square of the operating radius which leads to performance increase for operating at smaller radii. This trend is shown in Eq. (4) and Fig. 2(c), where once again, increasing the design parameters (T_{crit}, d_o, σ_{sat}, $\frac{d_z}{k_z}$) results in improved performance, in contrast to designing for small r_{max}. As a result, it is expected that operating the device at r_{op} much less than r_{max} will offer a means to improved device performance and that the 11 nm maximum filament radius for the device studied here is much larger than the dimensions that are expected in high performance devices

$$\frac{R_{OFF}}{R_{ON}} = \frac{R_{max}}{R_{ON}} = \frac{4L_{WF}T_{crit}d_o\sigma_{sat}\frac{d_z}{k_z}}{r_{op}^2}. \tag{4}$$

III. THERMAL CROSS-TALK

Ultimately, the filament radius is likely to be only one factor in determining the density limits of RRAM. Large amounts of heat are generated within the filaments, and that heat will have the potential to alter the device state itself or the state of its neighbors as it flows radially. The radial temperature profile is determined by the radial flow of heat from the filament and the rate at which heat can escape vertically through the electrodes or other encasing materials. Including this vertical heat loss (q_z) in the steady-state heat equation gives

$$\frac{1}{r}\frac{\partial}{\partial r}k_r r\frac{\partial T}{\partial r} - q_z = 0. \tag{5}$$

This heat loss term can be approximated within a differential area as

$$q_z = \frac{k_z(Area)}{d_z}(T - T_{RT}), \tag{6}$$

where the outer edge of the electrode in the vertical (z-direction) is held at room temperature. This vertical thermal conductivity may be dominated by spreading effects in the electrode or interfacial thermal resistance so it is not usually appropriate to think of k_z as a bulk thermal conductivity value. Its value depends strongly on the specific materials, interfacial properties, and geometry and it should therefore be thought of as an equivalent thermal conductivity.

Then dividing Eq. (6) by the differential volume of the oxide region from which heat is being lost gives

$$\frac{1}{r}\frac{\partial}{\partial r}k_r r\frac{\partial T}{\partial r} - \frac{k_z(T - T_{RT})}{d_o d_z} = 0. \tag{7}$$

The above equation is analytically solved by the modified Bessel function of the second kind

$$T = A\,K_0(\rho) + T_{RT}, \tag{8a}$$

$$A = \frac{(T_{crit} - T_{RT})}{K_0\left(\sqrt{\frac{k_z}{k_o d_o d_z}}r_f\right)}, \tag{8b}$$

$$\rho = \sqrt{\frac{k_z}{k_o d_o d_z}}r, \tag{8c}$$

where K_0 is the modified Bessel function of the second kind, k_o is the thermal conductivity in the radial direction through

the insulating oxide, and r_f is the filament radius which is the radius of the dominant current conduction pathway and may be different from both r_{op} and r_{max}.

Figure 3(a) shows the temperature profiles from Eq. (8) for varied thermal conductivity ratio. The temperature drops quickly from the activation temperature at the edge of the 8 nm filament to room temperature if the vertical thermal conductivity (k_z) is much larger than the lateral thermal conductivity (k_o). This is likely to be the case if the vertical material is an electrode and the surrounding material is an insulating oxide provided that there is negligible interfacial resistance, but the thermal boundary conductance may be significant for such dissimilar materials.[17] The vertical thermal conductance may be comparable to or even less than the lateral thermal conductance because of the interfacial resistance. In that case, the elevated temperatures can remain for several tens of nanometers and could be relevant for scaling considerations. Mathematically, decreases in the prefactor $\sqrt{\frac{k_z}{k_o d_o d_z}}$ from Eq. (8c) result in a wider temperature profile. Physically, increasing vertical thermal conductivity allows more heat to escape that therefore does not contribute to the elevated radial temperatures. These elevated radial temperatures can potentially affect the state of neighboring devices if those devices are in close proximity. By treating the device retention and error rate as having Arrhenius behavior, we can start to describe the stability of the devices in the presence of self-heating and thermal crosstalk.

The Arrhenius equation describes the exponential temperature dependence of thermally activated processes such as ionic motion.[18,19] We compare the retention of a device that is heated by its neighbor to the retention of the same device without heating by dividing the Arrhenius rate equation evaluated at room temperature the Arrhenius rate equation with temperature as a function of radius

$$\text{Retention}_{\text{ratio}} = \text{Error Rate}_{\text{ratio}}^{-1} = e^{\frac{T_{crit}}{T(r)} - \frac{T_{crit}}{T_{RT}}}. \quad (9)$$

Since Eq. (9) expresses the relative rate of ionic motion in neighboring filaments at different locations, the expression can be thought of as the ratio of the retention of a filament that is heated by its neighbor to the retention of that same filament with no heating. The same expression can be thought of as the inverse of the error rate, where heated devices experience state changes more frequently. This ratio is shown in Fig. 3(b) for the same devices described in Fig. 3(a). The figure shows the effect of thermal crosstalk on retention ratio, or equivalently inverse error rate, as a function of distance from the center of a neighboring device that is held constant at the switching temperature. This type of operation condition is not a likely scenario for most applications and acceptable retention times will vary widely between specific applications, but the device design can be easily matched to operational specifications by combining the local temperature (Eq. (8)) with the Arrhenius equation (Eq. (9)). The effect of thermal crosstalk is most likely to increase resistance because the high concentration of vacancies in the filament is unstable without the effect of the driving field that was used to inject them. For that reason, the ON (low resistance state) may be particularly susceptible to thermal crosstalk.

Examples of designing for thermal crosstalk are shown in Figs. 4 and 5 where the retention ratio defined in Eq. (9) is shown for varied design parameters that influence the temperature profile outside the filament. The retention ratio is measured at a location 30 nm from the center of a 5 nm neighboring filament (Fig. 4(a)) and at a location 20 nm from the center of an 8 nm neighboring filament (Fig. 4(b)). The neighboring filaments are held constant at the switching temperature. The parameters used in the figure are $T_{crit} = 1500\,\text{K}$, $T_{RT} = 358\,\text{K}$, $k_o = 1\ \text{W/mK}$, $d_o = 10$ nm, $d_z = 50$ nm, and $k_z = 5\ \text{W/mK}$. The values for k_z, d_o, and r_f were varied between 1/3 and 3 times their initial value in Fig. 4, and the values for T_{crit} and T_{RT} were varied over the same range in Fig. 5. Since the radii were varied to values comparable to and larger than the test location, the retention ratio (or inverse error rate) becomes very low for large radii. The specific values for optimized performance will depend greatly on the application and design constraints but Figs. 4 and 5 are demonstrations of how Eqs. (8) and (9) can be combined to provide design insight regarding thermal crosstalk. In Fig. 5, it is important to remember that the retention ratio is normalized by the retention at room temperature which is dependent on both the critical temperature and the room temperature.

In the context of designing devices to meet scaling specifications, it is worth discussing a few of the engineering trade-offs that are likely to be encountered as scaling continues. Perhaps, the most relevant is between device density and energy consumption per device. As discussed above, when thermal crosstalk becomes a limiting design consideration, increasing the thermal conduction in the vertical direction is a good approach to decreasing radial temperatures

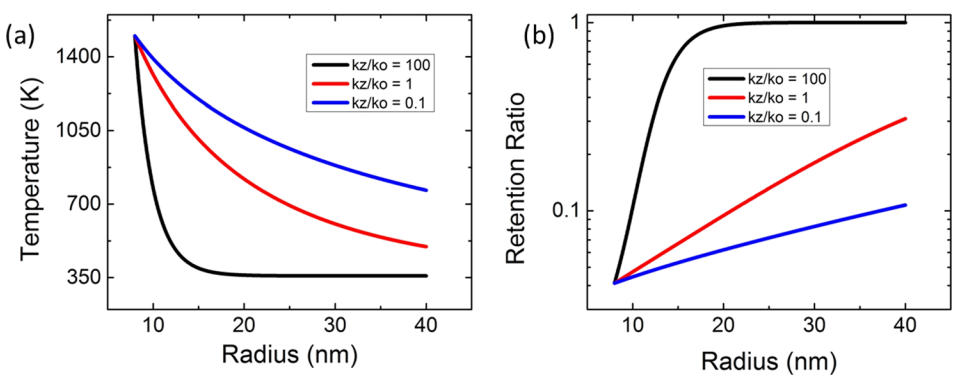

FIG. 3. The temperature surrounding the filament follows a modified Bessel function of the second kind (a) and can be modulated by varying design parameters such as vertical and lateral thermal conductivities. The lateral temperature can decrease the retention of neighboring devices as compared to isolated devices (b).

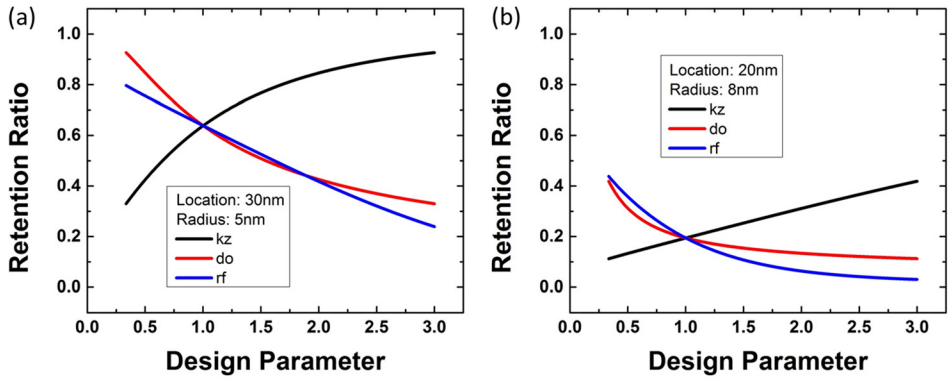

FIG. 4. By varying the design parameters, it is possible to control the thermal cross-talk between devices. The retention at a location 30 nm (a) and 20 nm (b) from the center of a neighboring device held at the switching temperature is shown. The retention is normalized by the retention for an isolated device at room temperature. The effect is shown for a neighboring device with 5 nm (a) and 8 nm (b) filament radii.

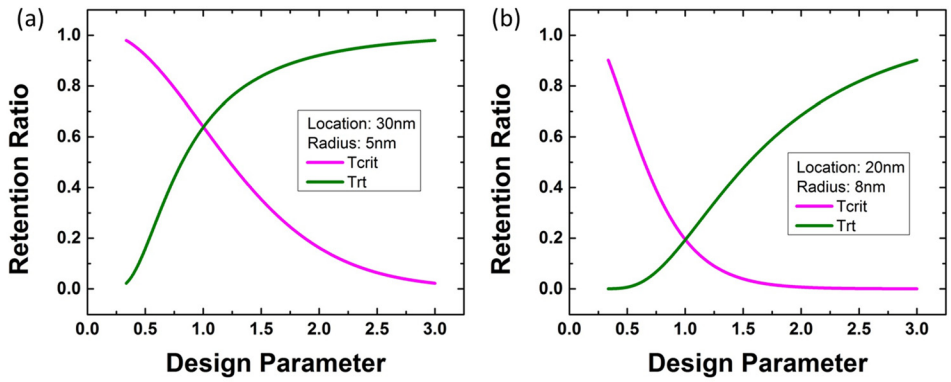

FIG. 5. By varying the design parameters, it is possible to control the thermal cross-talk between devices. The retention is normalized by the retention for an isolated device at room temperature. Changing the critical temperature or room temperature changes the normalization so extra attention is needed for interpreting these results. The retention is shown at a location 30 nm (a) and 20 nm (b) from the center of a neighboring device with 5 nm (a) and 8 nm (b) filament radii held at the switching temperature.

and enabling larger device densities. Unfortunately, that approach leads to increased power required to maintain T_{crit} within the filament because heat escapes more efficiently. This leads to higher switching power and likely higher switching voltage. An alternate approach with the same tradeoff involves designing or selecting materials with increased T_{crit} to increase device density. That change would also lead to increased power requirements as more heat would be required to reach the higher T_{crit} value. In Fig. 5, increasing T_{crit} appears to lead to decreased retention or increased error rates but that figure is normalized by the retention of an isolated device so, although retention increases, normalized retention (i.e., retention ratio) decreases. In most cases, absolute retention is the performance metric and not relative retention, so the device density could be increased as a result of this added stability but at the cost of increased power.

It may be possible to avoid some of these tradeoffs by designing the devices and their surroundings to guide heat flow. By separating the devices laterally with a thermally resistive interlayer dielectric, it may be possible to reduce both energy consumption and the extent of lateral temperatures, provided that the interlayer dielectric can hold heat in the filament and suppress the ratio of radial to vertical heat flow between devices. This effect can be further improved by using thermally conductive material or interfaces in the vertical direction between devices.

IV. SELF-HEATING

Controlling heat flow in the vertical direction will also have an impact on the device itself, not just its neighbors. Careful design of the devices may be able to improve power efficiency and scaling. One possible approach involves the use of temperature dependent thermal conductivities and has the additional advantages of decreasing the risk of perturbations caused during a READ operation or due to the half-select problem[6] in crossbar arrays. During the nonpertubative resistance READ, a device will experience significant Joule heating that will eventually degrade its state. The same effect is particularly prevalent in a device that is "half-selected" in a crossbar architecture. For a device that is constantly at a READ voltage, and where temperature is proportional to the applied power,[12] the earlier definition of retention ratio or inverse error rate yields

$$\text{Retention}_{\text{ratio}} = \text{Error Rate}_{\text{ratio}}^{-1} = e^{\frac{T_{crit}}{|T_{crit}-T_{RT}|\left(\frac{V_{read}}{V_{crit}}\right)^2+T_{RT}}-\frac{T_{crit}}{T_{RT}}}. \quad (10)$$

For a continuously half-selected device with a critical temperature of 1500 °C (Ref. 12) and an ambient temperature of 300 K, the error rate increases by an order of magnitude (12.2×) compared to a device that is never half-selected. For the READ operation of a device with the same critical and ambient temperatures being constantly read at 20% of the turn-on voltage (i.e., 100 mV READ for a 0.5 V turn-ON device), the retention is halved when compared to a device that is never read. This effect is shown in Fig. 6, highlighting retention loss at the half-select and 20% READ voltages.

The above derivation ignores possible electric field effects that may contribute to retention in partially selected devices. Our results and calculations indicate that the thermal dependence dominates the ionic motion and that the

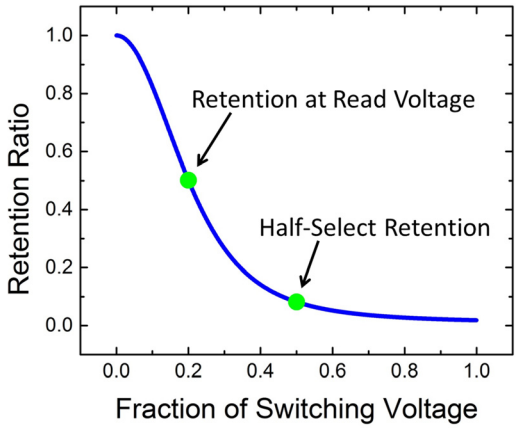

FIG. 6. Applying voltages less than the switching voltage will cause self-heating that will increase the error rate or decrease the retention. The retention is shown for a device that is continuously exposed to applied voltages. The retention is normalized by a device with zero voltage applied.

field effects can be ignored over a wide range of conditions. The field dependence of ionic motion could easily be included by adding a hyperbolic sine factor that is also thermally activated.[13,20] The derivation of Eq. (10) also assumes that all thermal conductivities are temperature independent but that is not a requirement for device design. A vertical thermal conductivity that is inversely related to temperature would allow heat to escape more efficiently at low temperatures than at high temperatures, thereby decreasing the temperature on a device during READ or half-select. The exponential dependence of activation on temperature indicates that even seemingly small changes in the temperature dependence of thermal conductivity can have significant impact on retention.

Designing electrodes with inverse temperature dependence should be possible. Many electrode materials (including copper and tungsten)[21] have this property intrinsically but common RRAM electrodes such as platinum[21] and tantalum[22] exhibit the opposite behavior which exacerbates the effect in Fig. 6. Titanium nitride (TiN), which is a common electrode in CMOS-compatible processing, is even more complicated because the phonon contribution to thermal conductivity increases rapidly with temperature.[23] The total thermal conductivity for TiN decreases over a wide range of temperatures but then increases between 1300 and 1800 °C (Ref. 24) which is in the expected range of transition temperatures. In addition to selecting favorable metals, defect structure can be used to provide thermal conductivity that decreases with increasing temperature[25] for increasing retention. Highly doped silicon also has a favorable dependence of thermal conductivity on temperature[26] and may be an interesting choice for electrode material.

V. CONCLUSIONS

We provided a theoretical justification for device scaling well into the single digit nanometer lateral dimensions without specifying a hard lower limit. We also showed that, as device densities are increased, their thermal interactions will become increasingly important and we derived analytical expressions for the temperature profile outside of a conducting filament. We used that temperature profile to quantify the related effects on retention and error rates of neighboring devices and described, in analytical detail, the retention effects of reading devices and of the potential "half-select" problem. We proposed designs to reduce thermal cross-talk, thereby increasing device density and we suggested a class of electrode materials that not only increase device density but can help to increase retention in devices that are frequently read or half-selected. Our results discuss several engineering tradeoffs that may need to be considered in highly scaled and densely packed devices but we believe that RRAM scaling can continue well beyond the capabilities of current commercial manufacturing technologies.

ACKNOWLEDGMENTS

The authors would like to thank Thomas Beechem for helpful discussion of thermal interfaces. This work was funded by Sandia's Laboratory Directed Research and Development 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]B. Dieny and C. Jagadish, "Emerging non-volatile memories: Magnetic and resistive technologies," J. Phys. D: Appl. Phys. **46**, 070301 (2013).
[2]D. Ielmini, "Reliability issues and modeling of flash and post-flash memory," Microelectron. Eng. **86**, 1870–1875 (2009).
[3]G. W. Burr, "Overview of candidate device technologies for storage-class memory," IBM J. Res. Develop. **52**, 449–464 (2008).
[4]H.-S. P. Wong, H.-Y. Lee, S. Yu, Y.-S. Chen, Y. Wu, P.-S. Chen, B. Lee, F. T. Chen, and M.-J. Tsai, "Metal-Oxide RRAM," Proc. IEEE **100**, 1951–1970 (2012).
[5]P. R. Mickel, A. J. Lohn, and M. J. Marinella, "Memristive switching: Physical mechanisms and applications," Mod. Phys. Lett. B **28**, 1430003 (2014).
[6]J. J. Yang, D. B. Strukov, and D. R. Stewart, "Memristive devices for computing," Nat. Nanotechnol. **8**, 13–24 (2013).
[7]A. C. Torrezan, J. P. Strachan, G. Medeiros-Ribeiro, and R. S. Williams, Nanotechnology **22**, 485203 (2011).
[8]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, U.-I. Chung, I.-K. Yoo, and K. Kim, "A fast, high-endurance and scalable non-volatile memory device made from asymmetric Ta_2O_{5-x}/TaO_{2-x} bilayer structures," Nature Mater. **10**, 625–630 (2011).
[9]J. P. Strachan, A. C. Torrezan, G. Medeiros-Ribeiro, and R. S. Williams, Nanotechnology **22**, 505402 (2011).
[10]A. J. Lohn, P. R. Mickel, C. D. James, and M. J. Marinella, "Degenerate resistive switching and ultrahigh density storage in resistive memory," arXiv preprint arXiv:1406.4033 (2014).
[11]Z. Zhang, Y. Wu, H.-S. P. Wong, and S. S. Wong, "Nanometer-scale HfO_x RRAM," IEEE Electron Device Lett. **34**, 1005–1007 (2013).
[12]P. R. Mickel, A. J. Lohn, C. D. James, and M. J. Marinella, "Isothermal switching and detailed filament evolution in memristive systems," Adv. Mater. (published online).
[13]D. B. Strukov and R. S. Williams, "Exponential ionic drift: Fast switching and low volatility of thin-film memristors," Appl. Phys. A **94**, 515–519 (2009).
[14]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).
[15]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).

[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]P. E. Hopkins, "Thermal transport across solid interfaces with nanoscale imperfections: Effects of roughness, disorder, dislocations, and bonding on thermal boundary conductance," ISRN Mech. Eng. **2013**, 682586.
[18]Z. Wei, T. Takagi, Y. Kanzawa, Y. Katoh, T. Ninomiya, K. Kawai, S. Muraoka, S. Mitani, K. Katayama, S. Fujii, R. Miyanaga, Y. Kawashima, T. Mikawa, K. Shimakawa, and K. Aono, "Retention model for high-density ReRAM," in *Proceedings of the IEEE International Memory Workshop* (IEEE, 2012), pp. 1–4.
[19]S. Larentis, C. Cagli, F. Nardi, and D. Ielmini, "Filament diffusion model for simulating reset and retention processes in RRAM," Microelectron. Eng. **88**, 1119–1123 (2011).
[20]N. Cabrera and N. F. Mott, "Theory of the oxidation of metals," Rep. Prog. Phys. **12**, 163 (1948).
[21]R. W. Powell, C. Y. Ho, and P. E. Liley, Thermal Conductivity of Selected Materials, National Standards Reference Data Series 8 (National Bureau of Standards, 1966), Category 5–Thermodynamic and Transport Properties.
[22]I. V. Savchenko and S. V. Stankus, "Thermal conductivity and thermal diffusivity of tantalum in the temperature range from 293 to 1800 K," Thermophys. Aeromech. **15**(4), 679–682 (2008).
[23]R. E. Taylor and J. Morreale, "Thermal conductivity of titanium carbide, zirconium carbide and titanium nitride at high temperatures," J. Am. Ceram. Soc. **47**(2), 69–73 (2006).
[24]J. F. Shackelford and W. Alexander, *Materials Science and Engineering Handbook*, 3rd ed. (CRC Press, 2001), Table 104.
[25]S.-H. Ju and X.-G. Liang, "Investigation on interfacial thermal resistance and phonon scattering at twist boundary of silicon," J. Appl. Phys. **113**, 053513 (2013).
[26]A. Stranz, J. Kahler, A. Waag, and E. Peiner, "Thermoelectric properties of high-doped silicon from room temperature to 900 K," J. Electron. Mater. **42**, 2381–2387 (2013).