Modeling of the Thermodiffusion-Induced Filament Formation in $TiN/Ta_xO_{1-x}/TiN$ Resistive-Switching Devices

Jingjia Meng,[1,*] Enkui Lian,[1] Jonathan D. Poplawsky,[2] and Marek Skowronski[1,†]

[1] *Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213, USA*

[2] *Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA*

(Received 6 January 2022; revised 4 April 2022; accepted 25 April 2022; published 24 May 2022)

Nonvolatile memory devices have received a lot of interest in both industry and academia in the last decade. Transition metal oxide-based memories offer potential applications as universal memory and artificial synapses. Here we focus on the one-time conditioning of metal/oxide/metal structures leading to the formation of a conducting filament in $TiN/Ta_xO_{1-x}/TiN$ structures and develop a finite-element model of this process. The process considered here consists of two steps. First the thermal runaway increases the temperature of the device and sets up large temperature gradients. In the second step, the lateral temperature gradient drives the ion motion forming a Ta-rich and O-poor filament. The process comes to steady state when the ion flux due to concentration and stress gradients balances the thermophoretic fluxes. The model replicated the structure of the filament including the size of the Ta-rich filament core (20 nm diameter), the surrounding Ta-depleted ring (50 nm), and the compositions of both regions. In addition, the model reproduced characteristic dynamics of the electroformation with slow changes of conductance during the incubation period, rapid increase of conductance during compositional runaway, and saturation. The range of critical material parameters, namely transport heats for Ta and O, is discussed.

DOI: 10.1103/PhysRevApplied.17.054040

I. INTRODUCTION

Oxide-based resistive-switching devices show promise for nonvolatile solid-state memory applications [1] with the first commercial memory product introduced by Panasonic in 2013. More recently, Taiwan Semiconductor Manufacturing Company added oxide resistive random-access-memory devices to its process portfolio [2]. They also can be used in neuromorphic circuits serving as artificial synapses [3]. This potential prompted wide-ranging investigations, including exploration of different materials and device structures, and fundamental studies of processes involved in memory switching. While research efforts have produced a lot of specific data, we still do not have a comprehensive understanding of the switching processes in a form of a quantitative physics-based device model.

A considerable number of resistive-switching device types based on a variety of materials have been fabricated, tested, and their operating mechanism investigated. Here we focus only on one class of such devices referred to as valence change memories (VCM) [4], which encode information by changing the composition of the functional oxide by motion of ions. The initially uniform functional oxide undergoes a one-time conditioning process, referred to as electroformation, which creates a small-diameter conductive filament [4]. This process invariably reduces the resistance of the device with low and high resistance states of the switch having resistance blow the value for as-fabricated structure. The original interpretation of electroformation [5] assumed that the oxygen ions driven by the electric field cross the interface with the anode and locally reduce the functional oxide. The process was suggested to occur at close to room temperature [4–6]. Loss of oxygen from the functional layer to an easily oxidizable metal electrode, such as Ta or Ti, has been suggested by numerous experimental observations [7–9]. However, the process has not been confirmed and quantified by a direct analytical method. It is also not clear at what temperature such a transfer occurs.

An alternative model concerned structures that exhibited unipolar switching by a thermochemical mechanism and involved two steps [10,11]. First, the devices undergo a local electrical breakdown producing a hot spot. This is followed by oxygen diffusion driven by radial temperature gradients (mechanism referred to as thermodiffusion, thermophoresis, or Soret effect). It was assumed that oxygen is thermophobic (negative Soret coefficient) and is depleted within the hot spot. A somewhat similar approach was developed by Lee *et al*. [12]. These authors simulated the evolution of the structure consisting of a relatively thick (30 nm) reduced TaO_x and thin (5 nm)

*jingjiam@andrew.cmu.edu
†mareks@cmu.edu

almost stoichiometric Ta_2O_5 with as low as 1×10^{16} cm^{-3} vacancy concentration sandwiched between metal electrodes. The initial temperature increase was assumed to be due to locally higher oxygen-vacancy concentration in the stoichiometric part of the functional layer. The operating mechanism was a combination of thermodiffusion and electromigration allowing the oxygen vacancies from the reduced oxide to diffuse into the highly resistive Ta_2O_5.

This paper describes the model of the electroformation process in devices with reduced TaO_x functional layer and inert TiN electrodes. This type of structure was extensively characterized by Ma *et al*. [13–17]. These authors documented the two steps in the electroformation. The initial increase of current was identified as due to threshold switching caused by thermal runaway [18] a mechanism that is common in oxides [19]. Threshold switching results in spontaneous current constriction and formation of the hot spot in larger devices [20]. This, in turn, sets up large lateral temperature gradients that drive the motion of ions [21]. Transmission electron microscopy-based methods indicated that both the Ta and O ions move during the electroformation: Ta accumulates in the filament core with a ring of lower Ta content surrounding the core. Oxygen ions move away from the hot spot [13,14]. Both of these motions produce a positive feedback loop between composition and temperature gradients resulting in a compositional runaway [15]. Taken together, all observations summarized above and associated electrical characteristics form an extensive base for a well substantiated model without the necessity of making numerous *ad hoc* assumptions.

This work uses the finite-element method to simulate lateral ion motion resulting in formation of the filament in $TiN/TaO_x/TiN$ structures. We allow for motion of both Ta and O ions as evidenced by the experiments and so far, not incorporated in any model. The large changes of film composition in the filament led us to include the strain energy term in the simulation. This term couples the motion of Ta and O ions and limits the extent of composition changes. Previous models of electroformation and switching neglected this term assuming that the relative concentration changes were below 1%. Lastly, while both Marchewka *et al.* [6] and Lee *et al*. [12] incorporated thermodiffusion in their models, the effects of this term were either absent or not clear. Neither could reproduce the composition maps of Ma *et al*. [14]. Also, our model addresses the dynamics of the electroformation in addition to steady-state composition profiles.

II. RESULTS AND DISCUSSION

A. Device structure and model construction

The structure modeled in this work is a $TiN/Ta_{0.333}O_{0.667}/TiN$ structure with 50-nm-thick uniform functional oxide and a lateral size of 150 nm (Fig. 1). The structure is integrated with an on-chip 37-kΩ load resistor in order to prevent current overshoot due to discharge of parasitic capacitances [22]. See Sec. S1 within the Supplemental Material [35] for a detailed description of the device fabrication.

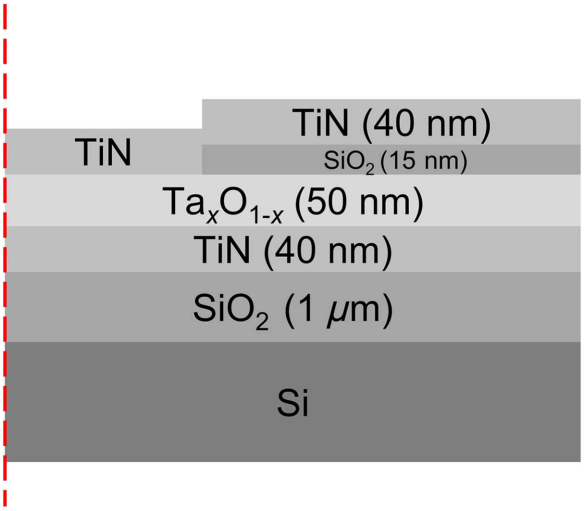

FIG. 1. Geometry of inverted-via memory device used in this work. The device is circular with the red dashed line representing the center axis.

The critical input parameter is the electrical conductivity of the functional oxide as a function of composition and temperature. We measure the conductivity of as-deposited Ta_xO_{1-x} films in the composition range of $0.299 < x < 0.370$ and temperatures between 300 and 600 K [Fig. 2(a)]. The conductivity of all investigated samples increased with increasing temperature in a linear fashion when plotted as $\ln \sigma$ versus $1/T$ as is typical of semiconductors and insulators. The activation energy is decreasing with Ta content eventually reaching zero when material becomes metallic [23].

The experimental conductivity data are fitted with the formula:

$$\sigma(x,T) = \sigma_o(x) e^{-(E_{\mathrm{cond}}(x))/kT}, \tag{1}$$

where $E_{\mathrm{cond}}(x)$ is the composition-dependent activation energy and σ_o is the factor depending on the density of centers responsible for generation of free carriers and their mobility. Both parameters have been extracted directly from the experiment and approximated by linear functions of composition:

$$E_{\mathrm{cond}}(x) = \begin{cases} -3.75 \times x + 1.496\ (\mathrm{eV}) & 0.286 < x < 0.392 \\ -0.025 \times x + 0.025\ (\mathrm{eV}) & 0.392 < x < 1.00 \end{cases}, \tag{2}$$

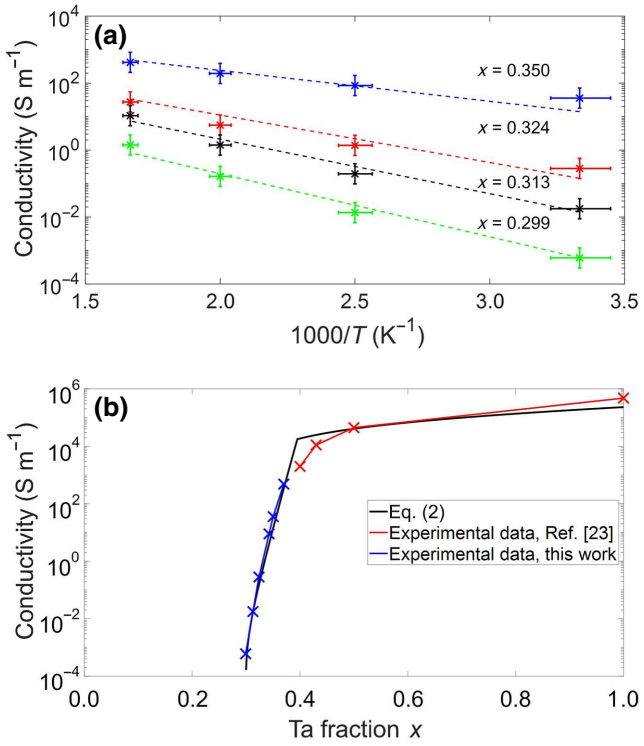

FIG. 2. (a) Experimental values of electrical conductivity as a function of temperature and composition. (b) Electrical conductivity at 300 K in the entire $0.286 < x < 1.00$ composition range including data from Ref. [23].

$$\sigma_o(x) = 3.30 \times 10^5 \times x - 9.97 \times 10^4 \ (\mathrm{S\ m^{-1}}) \quad 0.286 < x < 1.00. \tag{3}$$

The experimental data (crosses) and fitting results (dashed lines) are shown in Fig. 2(a). Since the filament composition range of interest is wider than our deposition system can produce, we adopt the data published by Rosario *et al.* [23]. Figure 2(b) shows the conductivity versus composition at room temperature. The blue crosses represent our experimental data, the red crosses denote the results of Rosario's experiment, and the black curve is the regression fit. It is apparent that the fitting function reproduces experimental data quite well with a slight overestimate of the conductivity at fractional concentration of Ta near 0.4.

We model the charge transport in the device by solving Poisson's equation with $\sigma(x, T)$ as input:

$$\nabla(\sigma(x, T)\nabla\varphi) = 0, \tag{4}$$

where φ is the electrostatic potential. In this part of the model, we neglect the ion contribution to conductivity as it is much smaller than electronic conductivity. Coupled with charge transport is the heat-transfer equation:

$$\rho C_p \frac{\partial T}{\partial t} - \nabla \cdot (k_{\mathrm{th}} \nabla T) = \vec{J} \cdot \vec{E}, \tag{5}$$

where ρ is the mass density of $\mathrm{Ta}_x\mathrm{O}_{1-x}$, C_p is the thermal capacitance of a given material, k_{th} is the thermal conductivity, J is the current density, E is the electric filed, T is the temperature, and t is the time. Material parameters C_p, k_{th}, and ρ are assumed to be constants independent of composition and temperature. The third coupled effect is the mass transport:

$$\frac{\partial c_i}{\partial t} = \nabla \cdot \left(D_i(T) \left(\nabla c_i + \frac{S_o^i}{kT^2} c_i \nabla T - \frac{\Omega_i}{kT} c_i \nabla \sigma_{\mathrm{hydro}} \right) \right) \tag{6}$$

and

$$\sigma_{\mathrm{hydro}} = B \frac{c_{\mathrm{Ta}} + c_{\mathrm{O}}}{c_{\mathrm{Ta}}^o + c_{\mathrm{O}}^o}. \tag{7}$$

The three terms in parenthesis on the right side of Eq. (6) represent fluxes induced by the concentration, temperature, and hydrostatic stress gradients. The subscript i corresponds to the two types of ions in the material: Ta and O. c_i is the number of atoms of type i per unit volume, $D_i(T)$ is the diffusion coefficient, S_o^i represents the transport heat, Ω_i is the ionic volume, and B is the bulk modulus of the film. The parameters D_i, S_o^i, Ω_i, and B are assumed to be independent of composition and, with exception of D_i, of temperature. The equation does not include the term describing the ion motion due to electromigration. Electromigration is almost entirely vertical while here we focus on the mechanism of the filament formation relying on lateral motion. As a consequence, the model does not describe the switching effects, which will be addressed in a separate publication. See Sec. S3 within the Supplemental Material [35] for a detailed description of modeling assumptions.

The data on diffusivities of Ta and O ions in amorphous $\mathrm{Ta}_x\mathrm{O}_{1-x}$ to be used in Eq. (6) are scarce. It is clear that both Ta and O ions are mobile and according to molecular dynamics calculations are of the same order of magnitude in reduced material [24,25]. This is due to similar formation energies on Ta interstitials and oxygen vacancies at low oxygen potentials and similar heights of barriers for diffusion [26]. To keep the number of adjustable material parameters to a minimum, we assume that the diffusion coefficients of both tantalum and oxygen ions are the same. The expression for the diffusivity is

$$D_i(T) = \frac{1}{2}\lambda^2 f\, e^{-(E_{\mathrm{diff}}/kT)}, \tag{8}$$

where λ is the effective hopping distance, f is the attempt frequency, and E_{diff} is the activation energy for ion migration. The values of the attempt frequency and effective

hopping distance are 10^{13} Hz and 0.32 nm, respectively [12,27].

The mass transport due to temperature gradient has been considered only in a few models of resistive switching and electroformation [12,28–30]. In solids, thermodiffusion has been studied mostly in the UO_{2-x} system due to composition changes observed in nuclear fuel rods [31]. In a macroscopic system such as this, the typical temperature gradients are of the order of 10^3 K m^{-1}. In nano-sized switching devices, they can be as high as 10^{10} K m^{-1} and it is essential to have their effect included. The Soret coefficient multiplying the $c_i\nabla T$ term has a form S_o^i/kT^2 with S_o^i representing the transport heat of specie i. At this point, the transport heats are known in very few materials systems and there is no established model predicting their values or even sign. In the zero-order approximation, one can replace transport heat with the enthalpy of diffusion with sign "−" for vacancy- and "+" for interstitial-mediated diffusion [32,33]. This suggests that the signs for Soret coefficients for oxygen and tantalum ions should be opposite. Given the above, we consider several sets of the transport heat with the S_o^{Ta} value being positive or zero and S_o^{O} either zero or negative.

The last term in the Eq. (6) considers the effects of stress during the electroforming process. In general, the thermo- and electromigration terms tend to accumulate ions in certain portions of the structure and deplete them in others. For large composition changes, the corresponding strain energy is comparable to other terms and its effects cannot be ignored. This term prevents unphysically large changes of ion density. The absolute total concentration of ions at zero stress is taken as 6.5×10^{22} atom cm^{-3}. The initial concentration for Ta and O ions is 2.16×10^{22} atom cm^{-3} and 4.33×10^{22} atom cm^{-3}, respectively.

B. Filament evolution under constant current

We simulate two types of experiments described by Ma *et al.* [15]. The first one used the current source applying a rectangular current pulse (200 μA) with the short current ramp up of 1 ms and constant current for 5 s. The ramp time is much longer than the thermal time constant of the device allowing it to reach a quasi-steady-state temperature distribution throughout the process. It is included to ease the process of numerical model conversion. To explore how each parameter influences the process of electroforming, we start with a simplified model assuming that there is no strain energy ($B = 0$), with $S_o^{\mathrm{Ta}} = 0.05$ eV, and $S_o^{\mathrm{O}} = 0$ eV. This eliminates the second and third terms in Eq (6) for oxygen flux. Since the initial oxygen distribution is uniform (c_{O}= const), the first term also vanishes, and the oxygen concentration stays constant everywhere throughout the process.

Figure 3(a) shows the simulated device voltage versus time with the inset showing the expanded view of

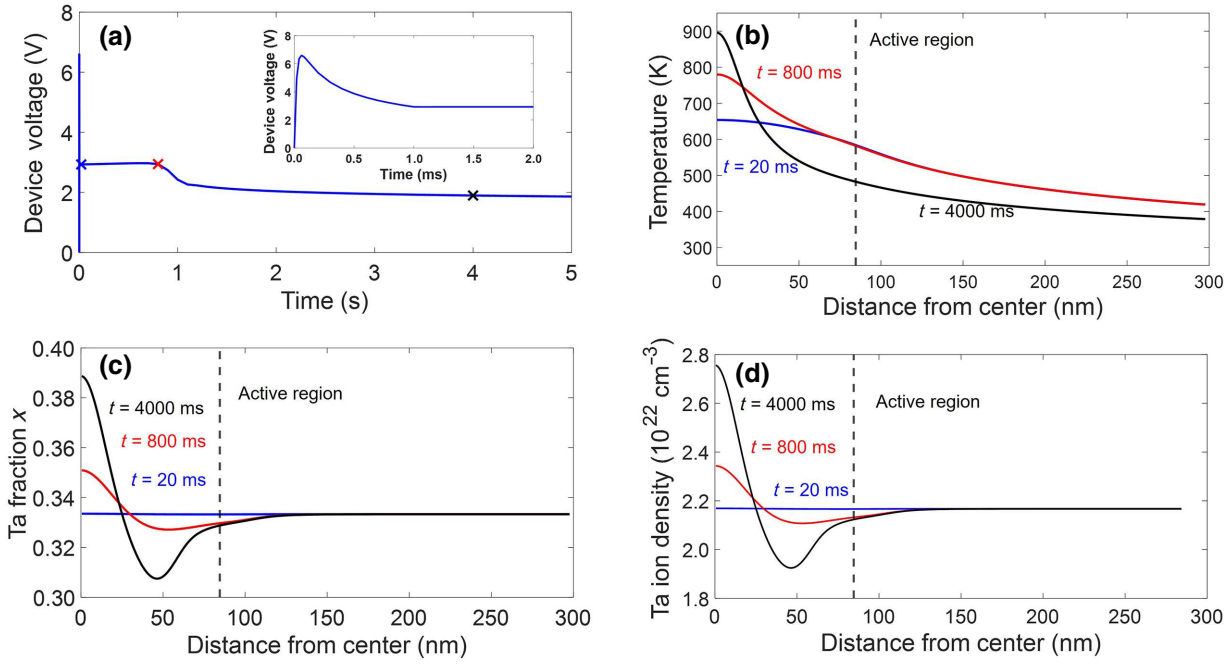

FIG. 3. (a) Voltage across the device versus time for the current pulse of 200 μA with 1-ms rise time for $S_o^{\mathrm{Ta}} = 0.05$ eV, $S_o^{\mathrm{O}} = 0$ eV, and $B = 0$. The inset shows an expanded view of the initial voltage transient with the overshoot caused by threshold switching. (b) Radial temperature distribution in the middle of the functional layer at three different times: after the threshold-switching event, at the beginning of the compositional runaway, and at steady state after electroformation. (c) Corresponding Ta fraction profiles and (d) Ta ion density profiles.

the initial transient. During current ramp up, the device retraces the quasi-static *V-I* characteristic of a threshold switch: it reaches the maximum voltage of $V_{max} = 6.6$ V at 0.06 ms and decreases to a quasi-steady-state value of 2.9 V for current of 200 μA. This behavior is induced by strong dependence of oxide conductivity on temperature and leads to decrease of voltage with increasing current and spontaneous current constriction [18]. At about 0.8 s, the device voltage drops rapidly again and stabilizes at about two thirds of the 2.9 V with a long time constant of approximately 1.1 s. Since this relaxation is much longer than the device thermal time constant, the changes should be thought as due to ion redistribution.

Figure 3(b) shows the radial temperature profiles in the middle of the functional layer modeled at three different times marked in (a) as colored crosses. The vertical dashed black line indicates the active region of the device defined by the size of the top contact. The blue curve is calculated 20 ms after the beginning of the pulse. This time is much longer than the thermal time constant of the device and corresponds to the steady-state temperature distribution after the threshold switch but before any significant ion motion has occurred. This statement is based on the blue trace of composition distribution in Fig. 3(c), which is a flat line corresponding to initial uniform Ta distribution with $x = 0.333$. The highest temperature of 650 K is reached at the center with the point of $\Delta T = T_{max}/2$ located 180 nm away and beyond the active region of the device. The temperature drop from the center to the edge of the active region is only 70 K. In other words, the volatile electronic filament resulting from threshold switching fills the entire active area of the device.

As the time passes, the temperature gradient starts to drive the Ta ions towards the center of the device. Initially the motion is slow due to low temperature (620 K) and low radial gradient ($< 6.5 \times 10^8$ K m^{-1}). As the center of the nascent permanent filament becomes Ta rich [red trace in Fig. 3(c) at 800 ms], it becomes more conductive, the temperature increases [red trace in Fig. 3(b)], and the voltage across the device starts to drop. This process forms a positive feedback loop in which composition changes enhance conductivity and cause T_{max} and T gradient increase, which, in turn, accelerate diffusion leading to a rapid decrease of resistance and voltage. Eventually the runaway slows down at about 1.5 s as the conductivity starts to saturate at high Ta content and the concentration gradient results in significant Fick's diffusion. At the steady state, the maximum temperature within the filament reaches 900 K with the maximum gradient of 1.3×10^{10} K m^{-1}. The FWHM of the temperature distribution is 20 nm. The corresponding composition profile [black trace in Fig. 3(c)] shows a maximum Ta content of 38.8% with the FWHM of 15 nm. In addition, it also shows a pronounced trench around the Ta-rich core with composition $Ta_{0.300}O_{0.700}$. Ta content is significantly lower here than the initial Ta density with material approaching composition of fully oxidized tantalum in $Ta_{0.286}O_{0.714}$. The trench forms mostly due to temperature distribution. If one considers a volume element at some distance from the center in initially uniform Ta_xO_{1-x}, the temperature and the diffusion rate at the face closer to the center is higher than at the outer face. The ion flux out of the element and toward the center is higher than the flux into the cell resulting in depletion of Ta. The rate of this process drops rapidly with distance from the center creating a quasi-steady-state. The diameter of the trench, however, will continuously grow with time at temperature together with corresponding conductance of the device. This could represent the ultimate endurance limit for devices operating on this principle. The composition profiles shown in Fig. 3(c) reflect the main features found in the experimental Ta distribution. Ma *et al.* [14] reported the composition of the filament Ta_xO_{1-x} to be $x = 0.667$–0.777, which is higher than the simulated value of $x = 0.390$. This deviation might be a result of overestimation of electrical conductivity in the simulation or Ta content in the filament in the experiment.

Figure 3(d) shows the Ta ion density distribution at three different times. It has a similar shape as the Ta fraction profile with the maximum Ta ion density at a steady state of 2.78×10^{22} atom cm^{-3}. This value is about 27% higher than the initial Ta density. The corresponding increase of total ion density accounting for Ta and O ions is about 10% compared with the initial value. This implies an unphysical level of stress much above yield strength of all materials in the structure and indicates that the strain energy associated with the formation of the filament must be included in Eq. (6).

The stress term is included in the simulation described below setting $B = 70$ GPa while keeping all other parameters the same as before ($S_o^{Ta} = 0.05$ eV, $S_o^{O} = 0$ eV, initial concentrations, etc.). See Fig. S4 within the Supplemental Material [35] for the result of ion distribution with other values of bulk modulus. It is worth noting that there is no driving force for the segregation of oxygen other than the coupling with the tantalum motion through the stress term. Figure 4(a) shows the Ta density distribution at $t = 20$ ms (before compositional runaway) and at $t = 4000$ ms at quasi-steady-state. The density profile is similar to that without inclusion of the stress term: Ta accumulates in the center of the device while a ring of depleted Ta forms around it. The FWHM of the Ta distribution shrunk to 11 nm while the maximum Ta density is 3% lower than the simulation result without the stress. The depletion region has a diameter of 60 nm. The inner 40 nm of the depleted ring (from 20 to 60 nm) shows a deep depression followed by gradual Ta density increase. The minimum concentration of Ta is 11% lower than its initial value but almost the same as the simulated result in Fig. 3. Figure 4(b) shows the distribution of O ion density, which forms almost a negative image of Ta distribution. Oxygen density in the

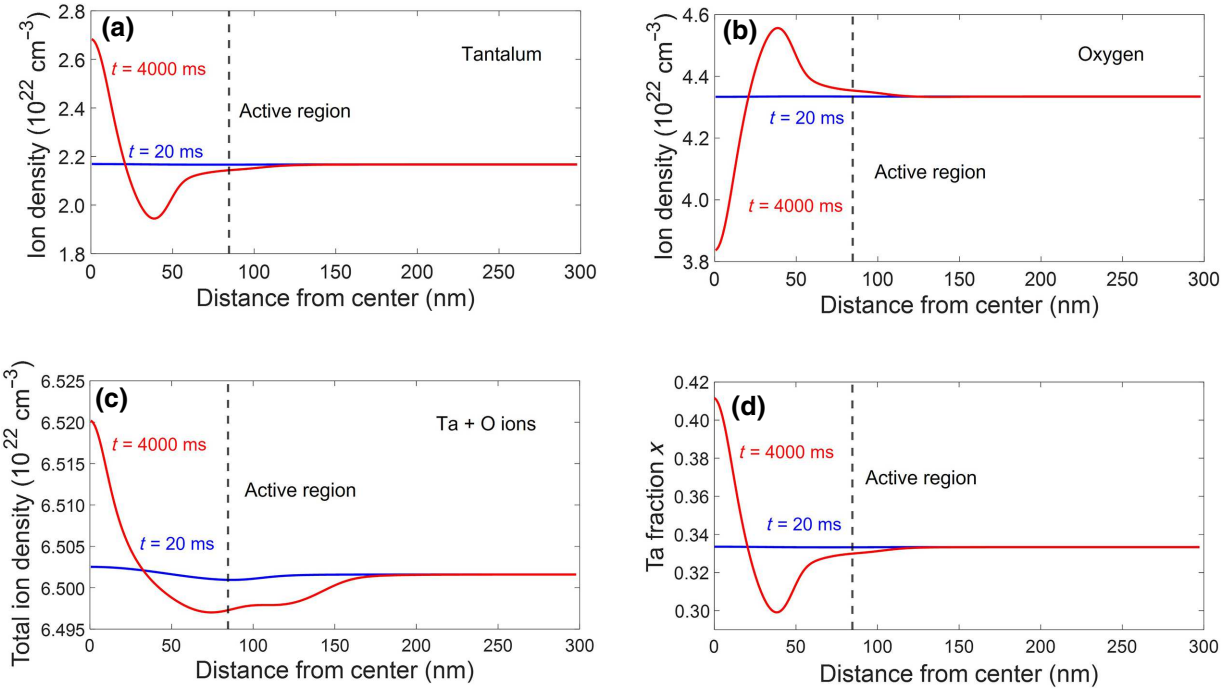

FIG. 4. (a), (b) Distribution of Ta and O ions before (blue line) and after electroforming (red curve) for the current pulse of 200 μA, $S_o^{\mathrm{Ta}} = 0.05$ eV, $S_o^{\mathrm{O}} = 0$ eV, and $B = 70$ GPa. (c) Distribution of total density of ions. (d) Fractional composition of Ta before and after filament formation.

center decreased by 11.5% compared with initial density with the FWHM of depleted region of 12.0 nm. The removal of the oxygen ion from the core region relieves part of the strain energy generated by the tantalum ion accumulation. The oxygen density at the tantalum depleted ring increased by 5% compared to the initial value. It is worth noting that the integrated quantity of either Ta or O in the cylinder with 80 nm radius is approximately the same as the initial value due to the diffusion across the interface to either electrode being neglected. Both types of ions redistribute within this volume and do not diffuse in or out of it. The total local ion density in the filament region remains almost the same with the changes below 0.3% [Fig. 4(c)]. Another noticeable effect of adding the stress term is a somewhat broader redistribution area compared to Fig. 3(d).

Figure 4(d) shows the Ta fraction distribution along radial direction (the Ta fraction affects the electrical conductivity while the absolute density does not). One can see the overall shape of the curve closely following the one calculated while neglecting stress. The fraction of Ta in the center increased from 0.389 (without stress) to 0.412 (with stress). This represents 6% change. The conductivity change associated with such a small change increase is quite significant increasing by a factor of 3. With increased conductivity in the filament region, more current will flow through this region heating up the device to higher temperature and generating higher temperature gradient. The balance between Fick's diffusion, stress-induced drift, and thermodiffusion leads to FWHM of Ta fraction shrinking to 11 nm (by about 27%).

In summary, the effect of including the stress effects in the mass transport equation has not affected Ta distribution significantly for the same value of transport heat. It induced oxygen to move out of the area of Ta accumulation and move into the region where Ta is depleted. Allowing for motion of oxygen increased the composition and conductance changes within the filament and reduced the filament diameter.

One can easily extrapolate these results to cases of different values of S_o^{Ta} and S_o^{O} with high value of bulk modulus. For example, assuming negative oxygen transport heat and $S_o^{\mathrm{Ta}} = 0$ would drive thermophobic O ions out of the center of device and produce an oxygen-enriched ring around it. This would induce tensile and compressive stresses in areas of depletion and accumulation and force the Ta ions to move in the direction opposite to that of oxygen. The resulting distribution quite closely resembles the one shown in Fig. 4. See Fig. S5 within the Supplemental Material [35] for the detailed ion profiles. In essence, the elemental distributions are determined by the difference between transport heats for the two elements. If $S_o^{\mathrm{Ta}} > S_o^{\mathrm{O}} > 0$, Ta still forms excess Ta in the center with oxygen depleted. The experimental elemental distributions, therefore, can be replicated by a wide range of transport heats as long as $S_o^{\mathrm{Ta}} > S_o^{\mathrm{O}}$.

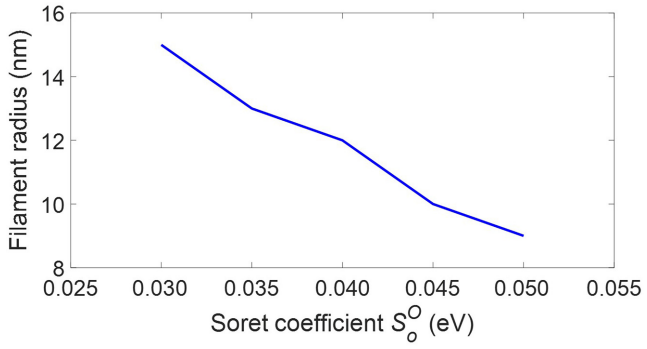

FIG. 5. Filament size as a function of tantalum transport heat S_o^{Ta}.

Figure 5 shows the diameter of the filament (defined as the FWHM of Ta fraction) as a function of S_o^{Ta} with $S_o^{\mathrm{O}} = 0$ and $B = 70$ GPa. The diameter is decreasing for increasing value of transport heat with the filament size at $S_o^{\mathrm{Ta}} = 0.05$ eV being almost half of the value at $S_o^{\mathrm{Ta}} = 0.03$ eV. Also, the maximum density of Ta increases and the incubation time decreases with increasing value of S. For values above 0.1 eV, the size of the filament becomes comparable to interatomic distances and the continuous model used here is no longer valid. For smaller values, the overall effect becomes weaker and the filament does not form.

Ma *et al*. [15] has reported that the electroforming process is thermally activated with the time of electroforming determined by the temperature in the "hot zone." The activation energy is assumed to be equal to the diffusion activation energy of mobile ions.

To test this assertion, we simulate the testing conditions used by Ma *et al*. Rather than using current pulse as done in modeling leading to Figs. 3–5, we use the voltage pulse source connected to the device under test via a load resistor (37 kOhm). Since the current source is equivalent to voltage source with infinite input resistance, this change is minimal as the load resistance is much higher than that of the device. It does not affect the ion distribution to any significant extent. The source voltage pulse has 1 μs ramp up followed by a constant voltage lasting 50 s. The ramp time is comparable to the thermal time constant of the device. The blue, red, and black curves in Fig. 6(a) show the simulated current as a function of time with the source voltage equal to 8, 10, and 14 V, respectively. See Fig. S7 within the Supplemental Material [35] for the direct comparison with experimental data. All three curves show two abrupt current jumps. The first current jump is due to the threshold switching process. With the higher source voltage, the time to reach the threshold becomes shorter. It takes 1 μs for the source voltage of 14 V for the device to reach its first plateau while 1.4 μs for 10 V and 2 μs for 8 V. These times are clearly affected by the ramp-up time, but they have been discussed extensively before [18]. After the

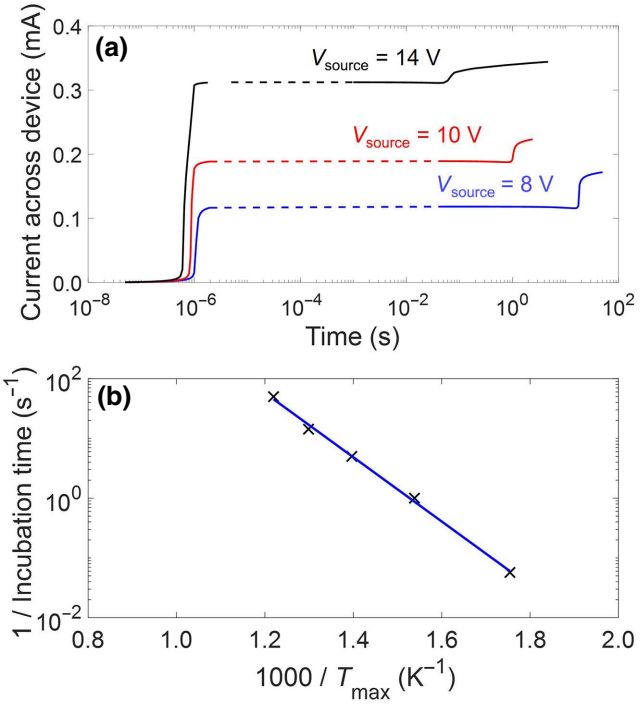

FIG. 6. (a) Current across the device versus time with different applied source voltage. (b) Plot of the inverse of incubation time versus temperature.

threshold-switching event, the temperature stabilizes for a relatively much longer period referred to as incubation time. These times are 17, 2, and 0.02 s for 8, 10, and 14 V source voltage, respectively. The explanation of the change is quite intuitive. Higher applied voltage leads to higher dissipated power density, and higher temperature within the device. This implies faster diffusion and, even for only slightly increased driving force of a temperature gradient, much shorter incubation time. The changes of composition and temperature are slow during incubation as evidenced by Figs. 3–5, but the rate is continuously increasing with time leading to the second rapid increase of current. This overall behavior agrees well with the experimental results (see Fig. S7 within the Supplemental Material [35]).

In Ma *et al*.'s work [15], the change of the incubation time for compositional runaway with applied bias has been approximated by a thermally activated behavior by calculating the maximum temperature within the device after threshold switching for a given value of bias. This assumed that the temperature during the incubation time is constant. However, Fig. 3(b) indicates that the temperature changes during the incubation time by about 100 K, clearly a not a negligible amount. A better estimate is to use the average temperature during incubation. Figure 6(b) shows the plot of the forming rate, defined as 1/simulated incubation time, versus inverse of the maximum temperature in the film at half of the incubation time. The plot shows a linear dependence corresponding to the activation energy of

$E_a = 1.07$ eV, which is close to our assumed value of the activation energy of diffusion $E_{\text{diff}} = 1.1$ eV. This result validates the argument that the electroforming process is an Arrhenius process with the activation energy determined by the activation energy of thermal diffusion.

The model consisting of Eqs. (1)–(8) with parameters used to create Figs. 5 and 6 is also used to simulate the common way of electroforming by applying a quasistatic voltage sweep. The time for voltage sweep is set to 1.2 s, the maximum source voltage to 16 V, and load resistor to 37 kOhm mimicking conditions used by Ma *et al.* [15]. It needs to be noted here that only a fraction of the source voltage appears across the device at any point in time and that Fig. 7(a) is plotted as a function of device voltage. The blue and red parts of one curve illustrate the simulation results while the solid black line shows the experimental *I-V* characteristics collected by Ma *et al.* [15]. In the simulation, the current initially increases superlinearly and reaches 13 μA at the knee at threshold voltage of 6.6 V. This result agrees with the measured threshold voltage of 6.3 V and the current at the threshold of 18 μA. Soon after the knee (marked by letter *B*), the simulated *I-V* experiences a thermal runaway rapidly transitioning along the load line denoted by the dashed part of the *I-V*. This transition is caused by the simulated device *I-V* dipping below the load line. At higher source voltage, the *I-V* continues approximately paralleling the experimental data. This part of *I-V* up to change in slope marked by *C* is fully reversible with no permanent changes taking place. The agreement between simulation and experiment is reasonably good with deviations caused mostly by an oversimplified dependence of conductivity on temperature. In the segment *B-C* of *I-V*, the maximum temperature within the device changes from 600 to 900 K and is much above the temperatures used to measure the conductivity [Fig. 2(a)]. It is likely that actual conductivity is higher than the values extrapolated from lower temperatures.

As the device evolves along the continuous blue line toward *C* with increasing source voltage, the device voltage continues to drop, but the current, dissipated power, maximum temperature, and lateral temperature gradient all increase. In parallel, the driving force for the thermodiffusion and the diffusion rate increases as well. At point *C*, the rate of current change due to the motion of ions exceeds the changes associated with heating and the *I-V* exhibits a second runaway marked by the dashed line in Fig. 7(a). The device at this point evolves with time even without the change of source voltage. The process eventually slows down with the *I-V* turning up and ion distribution approaching the steady state. The simulated beginning of the electroformation process occurs at a lower power compared with the experimental result (0.70 versus 0.93 mW) but the difference is not significant.

On the return sweep of source voltage [from point *D* to origin of coordinates, solid red line in Fig. 7(a)], the device

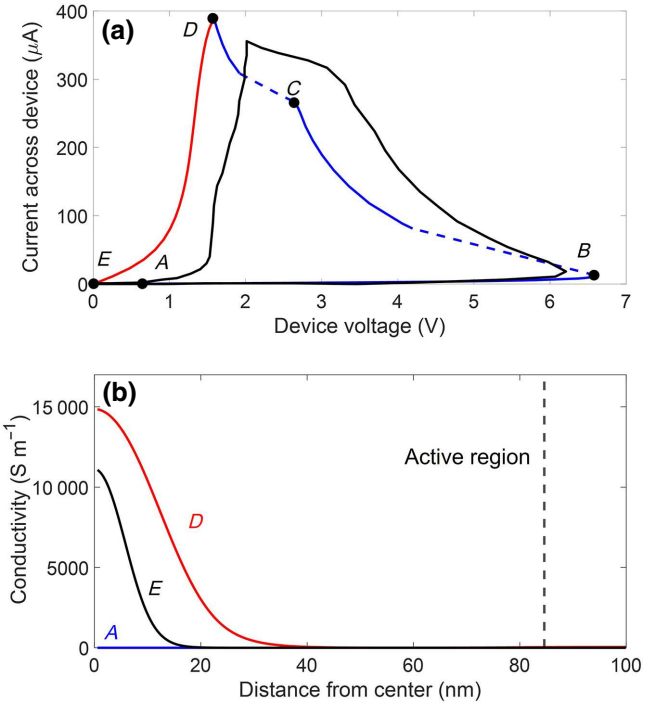

FIG. 7. (a) Simulation of quasi-dc voltage sweep of a memory-switching device (the red and blue curve is the simulation and the black curve shows experimental data from Ma *et al.* [15], reprinted with permission). (b) Electrical conductivity as a function of position at three points (marked by circles) along the *I-V*.

does not retrace its original path as the changes due to ion redistribution are permanent. Somewhat surprisingly, the return path is not linear even though there is a cylindrical conductive filament connecting the two electrodes [see Fig. S8(a)–S8(c) within the Supplemental Material [35] for relevant maps of temperature, composition, and conductivity at point *D*]. It is worthwhile to comment on the origin of this behavior as it should be common in oxide structures with a nanoscale conducting filament. In fact, many, although not all, switching devices exhibit nonlinear low resistance state characteristics. Before going forwards, one should point out that the model presented here is not intended to reproduce the postforming *I-V* characteristics. These are affected by the vertical motion of ions not considered here and can include effects not described by this model.

The insight into the origin of the nonlinearity can be gained considering the radial profiles of electrical conductivity at three different times [Fig. 7(b)]. The blue curve shows the temperature distribution at the beginning of the sweep [letter *A* in Fig. 7(a)]. The device at this point is still at 300 K and highly resistive everywhere. At the maximum source voltage (letter *D*), the device is hot and has reached its quasi-steady-state ion distribution with little

change expected on the way down. The decreasing voltage sweep is plotted as the red part of the curve. At the end of the sweep, the applied voltage is 0 V, the temperature is back to 300 K, and the conductivity everywhere in the device is lower than that at point *D*. This is to be expected as the conductivities in Ta_xO_{1-x} increase with temperature. What is somewhat surprising is the magnitude of the increase between *E* and *D*. It is due to the small diameter of the filament, which leads to high current densities and high temperatures. Moreover, the very significant contribution is due to heating of the surrounding oxide. The radius of the cylinder with conductivity exceeding 5000 S/m is 6.4 nm after electroformation but 16.5 nm at the maximum voltage. A similar effect is expected in other structures of this type with the degree of nonlinearity determined by the diameter of the filament and the thickness of the functional layer.

Finally, one needs to address the question of whether the model presented above applies to other materials and device designs. We believe that one aspect of the model, the existence of two steps with the first one leading to temperature increase and second relying on ion motion, is almost universal. Ions in transition metal oxides such as TaO_x and HfO_x are generally not mobile at room temperature. Before rearrangement of ions can occur, the device needs to be heated to temperatures above 600 K. Since as-fabricated devices have low conductance, generating enough Joule heat would require very high voltages. In the case considered here, the heating is made possible by the threshold-switching event induced by thermal runaway giving highly nonlinear *I-V* of the as-fabricated device. While thermally induced threshold switching is expected in all systems that exhibit superlinear dependence of conductivity on temperature, the threshold voltages are too high for highly resistive oxides such as HfO_2. Highly nonlinear *I-V* would have to originate in a different mechanism such as avalanching. Since virtually all experimental data indicate that the diameter of the filament is small, this mechanism would have to be local. It can be noted here that highly resistive oxides are typically used as very thin layers in switching devices. This results in electric field higher by at least an order of magnitude than the one in structures considered here making local breakdown more likely. The nature of such localized current flow has not yet been established.

The ion motion necessary in the second step of the process can either be driven by lateral temperature gradient, the process that is considered here, and/or by electric field causing vertical redistribution of ions. It is quite apparent that vertical composition changes within the oxide itself cannot produce lower resistances reported after electroformation. The increase of conductance must be induced by lateral redistribution of ions or by a net loss of oxygen to the anode. In the model discussed here, we neglect the second process by assuming negligible diffusion coefficient of O and Ta within TiN. This is, in our opinion justified, as Ma *et al*. argued that the exchange of oxygen with the TiN anode is caused by diffusion and occurs only after the filament is formed [15]. Such an exchange could be the dominant mechanism in other structures, for example, ones incorporating easily oxidizable metal as the anode. The rates and extent of such processes have not yet been quantified by direct analytical microscopy methods and their modeling would need to make *ad hoc* assumptions of exchange rates.

Another difference between resistive-switching structures that could impact the operating mechanisms is the oxygen deficiency in the as-fabricated devices. It is known that the diffusion coefficients of oxygen and tantalum depend on composition and increase with the oxygen deficiency [23]. The ratio D_{Ta}/D_O also appears to increase with oxygen deficiency making the redistribution of Ta ions more likely.

In summary, the electroformation model presented here is fully consistent with a large set of electrical and structural data available for $TiN/TaO_x/TiN$. It potentially applies to many device structures based on different materials. It is, however, not universal. It is likely that other structures operate based on different and yet unconfirmed mechanisms.

III. CONCLUSIONS

We present a finite-element model of 150 nm $TiN/Ta_{0.33}O_{0.67}/TiN$ resistive-switching device responding to applied voltage and current pulses and undergoing electroformation. Model reproduced the accumulation of tantalum in the core of the filament and its depletion in the surrounding cylinder with the oxygen concentration decreasing in the core and increasing at the outer periphery of the filament. This is due to the thermodiffusion term in expression for mass transport with Ta being thermophilic ($S_o^{Ta} > 0$) and/or O being thermophobic ($S_o^{O} < 0$). To replicate the experimental ion distribution, it is enough for one of the transport heats to be different from zero with stress coupling responsible for motion of the other type of ion. The size of the core is sensitively dependent on the value of *S* parameters decreasing with increase of *S*.

IV. METHODS

The Ta_xO_{1-x} films are deposited using reactive sputtering method with chamber pressure set to 3 mTorr and the power at 50 W. A mixture of Ar and O gas flow into the system as the reactive gas. The composition of Ta_xO_{1-x} is controlled by adjusting the oxygen and argon flow rates. The oxygen content with respect to oxygen and argon flow rates is measured by atom-probe tomography (APT). Needle-shaped specimens are fabricated from layered Ta_xO_{1-x} films using standard focused ion-beam liftout and milling techniques described by Thompson *et al*.

[34]. The APT experiments are run in a CAMECA LEAP 4000XHR with a 60-pJ laser power, 100-kHz pulse repetition rate, 50-K base temperature, and a 0.2% detection rate. See Fig. S2 within the Supplemental Material [35] for the experimental results and regression fit. Electrical conductivity is measured using circular transmission-line model structure and the hot stage.

Modeling is performed using a commercial COMSOL software package. See Table S1 within the Supplemental Material [35] for the properties of all material used in the model. Voltage is applied to the upper surface of the top electrode with the lower surface of the bottom electrode grounded. The entire structure is placed on 500-nm-thick SiO_2 with the bottom surface at 300 K and all other surfaces thermally insulated.

ACKNOWLEDGMENTS

This work is supported by the National Science Foundation (NSF) under Grant No. DMR-1905648. The authors acknowledge the use of the Materials Characterization Facility at Carnegie Mellon University supported by Grant No. MCF-677785. APT is conducted at ORNL's Center for Nanophase Materials Sciences (CNMS), which is a U.S. DOE Office of Science User Facility.

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The U.S. Government retains, and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan [36].

[1] 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, Nat. Mater. **10**, 625 (2011).

[2] *TSMC Offers 22 nm RRAM, Taking MRAM on to 16 nm*, https://www.eenewsanalog.com/news/tsmc-offers-22nm-rram-taking-mram-16nm

[3] Q. Xia and J. J. Yang, Memristive crossbar arrays for brain-inspired computing, Nat. Mater. **18**, 309 (2019).

[4] R. Waser, R. Dittmann, G. Staikov, and K. Szot, Redox-based resistive switching memories - nanoionic mechanisms, prospects, and challenges, Adv. Mater. **21**, 2632 (2009).

[5] T. Menke, P. Meuffels, R. Dittmann, K. Szot, and R. Waser, Separation of bulk and interface contributions to electroforming and resistive switching behavior of epitaxial Fe-doped $SrTiO_3$, J. Appl. Phys. **105**, 066104 (2009).

[6] A. Marchewka, R. Waser, and S. Menzel, in *Int. Conf. Simul. Semicond. Process. Devices, SISPAD, IEEE 2017-September* (IEEE, 2017), p. 133.

[7] U. Celano, J. Op De Beeck, S. Clima, M. Luebben, P. M. Koenraad, L. Goux, I. Valov, and W. Vandervorst, Direct probing of the dielectric scavenging-layer interface in oxide filamentary-based valence change memory, ACS Appl. Mater. Interfaces **9**, 10820 (2017).

[8] D. Lee, J. Woo, E. Cha, S. Park, S. Lee, J. Park, and H. Hwang, Defect engineering using bilayer structure in filament-type RRAM, IEEE Electron Device Lett. **34**, 1250 (2013).

[9] Y. Wang, Y. Yan, C. Wang, Y. Chen, J. Li, J. Zhao, and C. S. Hwang, Controlling the thin interfacial buffer layer for improving the reliability of the $Ta/Ta_2O_5/Pt$ resistive switching memory, Appl. Phys. Lett. **113**, 072902 (2018).

[10] D. Ielmini, R. Bruchhaus, and R. Waser, Thermochemical resistive switching: Materials, mechanisms, and scaling projections, Phase Transitions **84**, 570 (2011).

[11] D. B. Strukov, F. Alibart, and R. Stanley Williams, Thermophoresis/diffusion as a plausible mechanism for unipolar resistive switching in metal-oxide-metal memristors, Appl. Phys. A: Mater. Sci. Process. **107**, 509 (2012).

[12] S. H. Lee, J. Moon, Y. Jeong, J. Lee, X. Li, H. Wu, and W. D. Lu, Quantitative, dynamic TaO_x memristor/resistive random access memory model, ACS Appl. Electron. Mater. **2**, 701 (2020).

[13] Y. Ma, D. Li, A. A. Herzing, D. A. Cullen, B. T. Sneed, K. L. More, N. T. Nuhfer, J. A. Bain, and M. Skowronski, Formation of the conducting filament in TaO_x-resistive switching devices by thermal-gradient-induced cation accumulation, ACS Appl. Mater. Interfaces **10**, 23187 (2018).

[14] Y. Ma, J. M. Goodwill, D. Li, D. A. Cullen, J. D. Poplawsky, K. L. More, J. A. Bain, and M. Skowronski, Stable metallic enrichment in conductive filaments in TaO_x-based resistive switches arising from competing diffusive fluxes, Adv. Electron. Mater. 1800954, 1 (2019).

[15] Y. Ma, J. M. Goodwill, and M. Skowronski, in *2019 IEEE 11th International Memory Workshop, IMW 2019* (IEEE, 2020).

[16] Y. Ma, D. A. Cullen, J. M. Goodwill, Q. Xu, K. L. More, and M. Skowronski, Exchange of ions across the TiN/TaO_x interface during electroformation of TaO_x based resistive switching devices, ACS Appl. Mater. Interfaces **12**, 27378 (2020).

[17] Y. Ma, P. P. Yeoh, L. Shen, J. M. Goodwill, J. A. Bain, and M. Skowronski, Evolution of the conductive filament with cycling in TaO_x-based resistive switching devices, J. Appl. Phys. **128**, 194501 (2020).

[18] J. M. Goodwill, A. A. Sharma, D. Li, J. A. Bain, and M. Skowronski, Electro-thermal model of threshold switching in TaO_x-based devices, ACS Appl. Mater. Interfaces **9**, 11704 (2017).

[19] G. A. Gibson and G. A. Gibson, Designing negative differential resistance devices based on self-heating, Adv. Funct. Mater. **28**, 1704175 (2018).

[20] J. M. Goodwill, G. Ramer, D. Li, B. D. Hoskins, G. Pavlidis, J. J. McClelland, A. Centrone, J. A. Bain, and M.

Skowronski, Spontaneous current constriction in threshold switching devices, Nat. Commun. **10**, 1 (2019).

[21] A. R. Allnatt and A. V. Chadwick, *Thermal Diffusion In Crystalline Solids*, (n.d.).

[22] J. Meng, B. Zhao, Q. Xu, J. M. Goodwill, J. A. Bain, and M. Skowronski, Temperature overshoot as the cause of physical changes in resistive switching devices during electro-formation, J. Appl. Phys. **127**, 235107 (2020).

[23] C. M. M. Rosario, B. Thoner, A. Schonhals, S. Menzel, A. Meledin, N. P. Barradas, E. Alves, J. Mayer, M. Wuttig, R. Waser, N. A. Sobolev, and D. J. Wouters, Metallic filamentary conduction in valence change-based resistive switching devices: The case of TaO_x thin film with $X \sim 1$, Nanoscale **11**, 16978 (2019).

[24] B. Xiao and S. Watanabe, in *2014 International Conference on Solid State Devices and Materials* (The Japan Society of Applied Physics, 2014), pp. 1–2.

[25] B. Xiao, X. Yu, and S. Watanabe, A comparative study on the diffusion behaviors of metal and oxygen ions in metal-oxide-based resistance switches via ab initio molecular dynamics simulations, ACS Appl. Electron. Mater. **1**, 585 (2019).

[26] L. Zhu, J. Zhou, Z. Guo, and Z. Sun, Synergistic resistive switching mechanism of oxygen vacancies and metal interstitials in Ta_2O_5, J. Phys. Chem. C **120**, 2456 (2016).

[27] S. Kim, S. J. Kim, K. M. Kim, S. R. Lee, M. Chang, E. Cho, Y. B. Kim, C. J. Kim, U. In Chung, and I. K. Yoo, Physical electro-thermal model of resistive switching in Bi-layered resistance-change memory, Sci. Rep. **3**, 1 (2013).

[28] A. Marchewka, B. Roesgen, K. Skaja, H. Du, C. L. Jia, J. Mayer, V. Rana, R. Waser, and S. Menzel, Nanoionic resistive switching memories: on the physical nature of the dynamic reset process, Adv. Electron. Mater. **2**, 1500213 (2016).

[29] S. Kim, S. Choi, and W. Lu, Comprehensive physical model of dynamic resistive switching in an oxide memristor, ACS Nano **8**, 2369 (2014).

[30] 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).

[31] J. Janek and H. Timm, Thermal diffusion and Soret effect in $(U, Me)O_{2+\delta}$: The heat of transport of oxygen, J. Nucl. Mater. **255**, 116 (1998).

[32] J. A. Brinkman, The effect of temperature gradients on diffusion in crystals, Phys. Rev. **93**, 345 (1954).

[33] A. D. LeClaire, Some predicted effects of temperature gradients on diffusion in crystals, Phys. Rev. **93**, 344 (1954).

[34] K. Thompson, D. Lawrence, D. J. Larson, J. D. Olson, T. F. Kelly, and B. Gorman, In situ site-specific specimen preparation for atom probe tomography, Ultramicroscopy **107**, 131 (2007).

[35] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevApplied.17.054040 for the details of fabrication and simulation.

[36] http://energy.gov/downloads/doe-public-access-plan.