Journal of Applied Physics

RESEARCH ARTICLE | MARCH 19 2025

Parametric study of "filament and gap" models of resistive switching in TaO_x-based devices.

Rongchen Li ; Yang Bai ; Marek Skowronski ☐

J. Appl. Phys. 137, 114501 (2025)
https://doi.org/10.1063/5.0246985

Journal of Applied Physics ARTICLE pubs.aip.org/aip/jap

Parametric study of "filament and gap" models of resistive switching in TaO_x-based devices.

Cite as: J. Appl. Phys. 137, 114501 (2025); doi: 10.1063/5.0246985
Submitted: 5 November 2024 · Accepted: 22 February 2025 ·
Published Online: 19 March 2025

Rongchen Li, Yang Bai, and Marek Skowronski [a)]

AFFILIATIONS

Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

[a)]Author to whom correspondence should be addressed: ms3s@andrew.cmu.edu

ABSTRACT

A finite element model consisting of a conducting filament with or without a gap was used to reproduce the behavior of TaO_x-based resistive switching devices. The specific goal was to explore the range of possible filament parameters such as filament diameter, composition, gap width, and composition to reproduce the conductance and shape of I–V while keeping the maximum temperature within the acceptable range allowing for ion motion and preventing melting. The model solving heat and charge transport produced a good agreement with experimental data for the oxygen content in the filament below $TaO_{1.3}$, the filament diameter range between 6 and 22 nm, and the gap oxygen content between $TaO_{1.7}$ and $TaO_{1.85}$. Gap width was not limited to either low or high sides according to the criteria considered in this report. The obtained filament composition corresponds to oxygen deficiency an order of magnitude higher than one estimated by other modeling efforts. This was in large part due to the use of recent experimental values of conductivity as a function of composition and temperature. Our modeling results imply that a large fraction of atoms leaves and/or accumulates within the filament to produce a large relative concentration change. This, in turn, necessitates the inclusion of strain energy in the filament formation modeling. In addition, the results reproduce non-linear I–V without the necessity of assuming the Poole–Frenkel type of electrical conduction or the presence of a barrier at the oxide/metal interface.

I. INTRODUCTION

Resistive switching devices hold great promise for the next generation non-volatile solid state computer memories. The fabrication of these devices has been optimized in terms of functional material type and composition, electrode materials, deposition techniques, device design, and many other parameters to produce highly reliable devices. The most promising resistive switching devices appear to be valence change memories (VCMs) based on TaO_x.[1] In recent years, they entered the preproduction stage in major foundries.[2,3] However, the mechanism of the device operation and, in particular, the parameters of the small diameter filament responsible for the device conductance and switching are still a matter of dispute.

The conductance of VCMs is determined by the electroformation process which creates a small diameter conducting filament in an initially uniform resistive functional oxide layer.[4] The filament is made of point defects comprising either oxygen vacancies and/or metal interstitials with both acting as donors in many functional oxides. Switching occurs by the ion motion within the filament opening a gap with low donor density in the filament. Given the variety of device designs (type of functional oxide and its oxygen content, electrode metals, thicknesses, and device sizes) and formation procedures, it is surprising that the I–V characteristics of formed devices are quite similar implying similar filament characteristics. Typically, the switching voltages range from 1 to 2 V while the switching currents are between 0.1 and 1 mA.[4–11] The simulations of the device I–V characteristics, however, arrived at the values of filament diameter and defect densities that vary widely. For example, the diameter of the filament was assumed to be as little as 5 nm[12] or as large as 600 nm[13] (this is four orders of magnitude difference of the filament cross-sectional area). The width of the gap was between 1[14] and 10 nm[15,16] while the donor concentration was typically in the single atomic percent range of oxygen deficiency[17,18] with few studies considering the range of 10%–20%.[14] Such a wide range of parameters was caused, in part, by ad hoc assumptions about many unknown material parameters. Another

important consequence of such assumptions is the value of temperature simulated within the filament which varied from 500[14] to 3000 K.[18] It is apparent that the temperatures at the high end of the range allow for very different phenomena to take place during switching compared to the low end of estimates.

This report attempts to narrow the range of acceptable filament parameters (diameter, composition, and width of the gap) in both the low resistance and high resistance states (LRS and HRS, respectively) using the fewest possible ad hoc assumptions. We limit the materials discussed to amorphous TaO_x with Ti or Pt electrodes forming ohmic contacts. We specifically exclude from discussion structures with crystalline functional layers such as $SrTiO_3$[19] and ultrathin structures in which direct tunneling can play a major role.[20] The range of filament parameters was arrived at by using the values of current as a function of voltage and by the highest and lowest acceptable temperatures in the HRS and LRS. The implications of the estimated filament composition are discussed in terms of the energies of filament formation.

A. Setting up the model and its evaluation criteria

The filament in LRS is frequently approximated in finite element models of resistive switching devices by a cylinder with a diameter d and uniform composition x.[8–11,21–24] The composition here and the remainder of this report is defined as x in TaO_x. This is equivalent to the frequently used concentration of point defects (α_D equal to the sum of oxygen vacancies and metal interstitials) as

$$\alpha_D(TaO_x) = (2.5 - x) \times 2.2 \times 10^{22}\ cm^{-3}. \quad (1)$$

HRS is created by moving donor ions in the direction of the applied field forming a gap with the composition y and width w close to the anode. The remainder of the filament constitutes a "trunk," which we assumed has the same donor density as one in LRS. This is due to the length of the trunk being typically much larger than the width of the gap.[21,24,25] The diameter of the filament was assumed to be the same in LRS and HRS.

Electrical transport in amorphous films of TaO_x has been assessed by multiple groups. Most of the work has been focused on films with compositions close to stoichiometric and film thicknesses below 5 nm. The most relevant study was that of Perevalov et al. who studied films with $1.9 < x < 2.5$ and a thickness of 50 nm.[26] These authors considered four different models of conductivity proposed by Frenkel,[27] Hill,[28] Makram-Ebeid and Lannoo,[29] as well as Nasyrov and Gritsenko.[30] The best model was that of Nasyrov and Gritsenko. The Poole–Frenkel model could be used to reproduce experimental data but only for unphysical values of dielectric constant. Other models such as direct tunneling and trap-assisted tunneling were rejected as producing significant conductivities only in films below 10 nm. In this report, we have used conductivity as a function of composition and temperature extracted directly from experimental data.[31] The data are in good agreement with the partial data of Rosário et al.[32] and Bondi et al.[33] While the exact mechanism of conductivity is not firmly established, the fit to experimental data offers a well-founded starting point for device modeling.

An important parameter of the model is the contact resistivity as it can control the temperature at the switching interface in the LRS. Our estimation listed below is an order of magnitude higher than that of Ascoli et al.[34] to account for the difference in contact material used: TiN vs Ta. The thermal boundary resistivity at the TiN/TaO_x interface was estimated as 2×10^9 K m^2 W^{-1} based on the data of Lyeo and Cahill.[35] We have not included potential barriers at the electrode/oxide interfaces as there is no direct measurement of the barrier height as a function of composition or annealing and the shape of I–V can be reproduced without them. The diameter of the filament d, its composition x, as well as the width of the gap w and its composition y were treated as adjustable parameters of our model with the veracity of the model assessed by the fit to selected experimental data summarized below.

The primary result of the simulations was the functional dependencies of current vs voltage for LRS and HRS. Experimental data show the characteristics in LRS as almost linear while the current in the HRS depends super-linearly on voltage.[12,16,25,36–39] The two experimental studies we use for direct comparison with simulation results here are reported by Ma et al. and Heisig et al.[25,36] These were selected as the authors have attempted to experimentally determine the parameters of the filament.

The other important characteristic of switching devices is the temperature distribution as it determines the rate of ion motion and switching speed. There is no direct measurement of the temperature within the filament. The most reliable estimates are based on the injection of electrons from the filament into a p-type semiconductor by Yalon et al.[40] ($T > 1300$ K), crystallization temperature of HfO_x (Kwon et al.,[41] $T > 800$ K), and I–V analysis (Sharma et al.,[42] $T = 750$ K). Based on these results, we postulate that the lowest temperature allowing for switching is 800 K.

The activation energies for diffusion used in most simulations range from $1 < E_{diff} < 1.8$ eV[13,16,43,44] and are in general agreement with the direct measurements of diffusion rates.[45] Some simulations assumed lower activation energies and/or field-induced acceleration of diffusion.[12,14] Such low values were obtained as a fitting parameter and authors have not provided any independent justification. The consequences of $E_{diff} > 1$ eV are that to produce appreciable ion diffusion rates and switching dynamics, the temperatures within the filament must exceed 800 K. At the same time, the highest temperatures must stay below the melting temperatures of the electrodes and decomposition of the functional oxide. As the Ta–O phase diagram shows the liquid phase appearing at 1850 K, the estimated upper temperature limit was set to be at 1600 K.[46]

The maximum temperature within the filament depends on the dissipated power. It is worth noting that although the switching voltages can be quite different in different device structures, the power dissipated at the point of switching ranges from about 60 to 180 μW with an average of 110 μW.[6,25,36,47–51] The observation implies that the filaments should have a similar structure in different devices switching occurs when the temperature reaches a particular value. The temperature distribution and maximum values discussed below were obtained at the same dissipated power of 110 μW.

II. SIMULATION PROCEDURE

The TiN/TaO_x/TiN resistive switching devices were modeled using the commercial COMSOL Multiphysics software package. The schematic diagram of the active volume of the device in the LRS is shown in Fig. 1(a) (the entire device structure is shown in

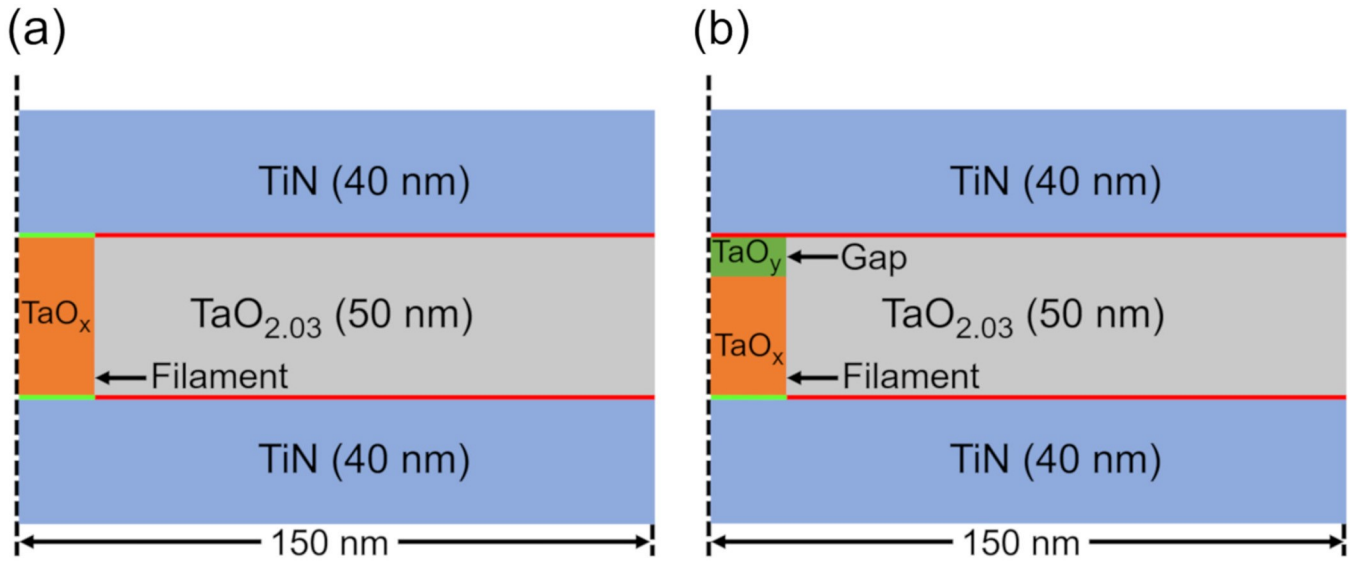

FIG. 1. Axisymmetric cross-sectional view of the device, the black dashed line representing the rotation axis. The green lines represent lower contact resistance while the red lines are the position of higher contact resistance. (a) LRS state. (b) HRS state.

the supplementary material). The as fabricated device has a 50 nm thick $TaO_{2.03}$ functional oxide layer and 40 nm thick TiN top and bottom electrodes. The electroformation process creates a cylindrical filament (marked by orange color in the figure) with composition TaO_x and diameter d, both of which are adjustable parameters in the simulation. The values of contact resistivities are marked with green lines for the contact between reduced oxide in the filament TaO_x and TiN ($6 \times 10^{-13}\ \Omega\ m^2$) and red lines for contact $TiN/TaO_{2.03}$ and TiN/TaO_y ($2 \times 10^{12}\ \Omega\ m^2$). Figure 1(b) shows the device in the HRS with the gap in the filament next to the top electrode. The diameter and the composition of the filament in HRS are assumed to be the same as in LRS with the gap width w and gap composition y serving as parameters in the model. The circuit used for all simulation included a 12 kΩ load resistor. The I–Vs were simulated using a source voltage sweep from 0 to 14 V in 1 ms. All I–V plots show current vs device voltage rather than source voltage. All temperature distributions were compared for a constant dissipated power of 110 μW.

The model solved two coupled differential equations. The heat transport equation was

$$\rho C_p \frac{\partial T}{\partial t} - \nabla \cdot (k_{th} \nabla T) = \vec{J} \cdot \vec{E}, \tag{2}$$

where ρ denotes the mass density, C_p the thermal capacity, k_{th} the thermal conductivity, T the temperature, t the time, $\vec{J}$ the current density, and $\vec{E}$ the electric field. All material parameter values and boundary conditions are listed in the supplementary material. The charge transport equation was

$$\nabla \cdot (\sigma(x, T) \nabla w) = 0, \tag{3}$$

where σ(x, T) is the experimentally determined electrical conductivity and w is the electrostatic potential.

III. RESULTS AND DISCUSSION

A. Low resistance state

The results of the I–V and temperature distribution simulations in the LRS along with two examples of experimental I–Vs are shown in Fig. 2. Figure 2(a) shows the experimental I–V curve from Ma et al.[25] (dashed line) and Heisig et al.[36] (dotted line). Both show almost linear dependence characteristic of LRS. The resistance of Ma et al. device is larger than the one of Heisig et al. by an order of one magnitude. Figure 2(b) shows the simulated I–V curves for the filament diameter d = 16 nm and the filament composition x varying from 0.4 to 1.6. Besides the obvious increase in resistance, the curves show a change in functional dependence. For higher conductivity values with x values 1.3 and below, the I–V curves are linear and the device resistance is between the values of Ma et al. and Heisig et al. In case of filament with higher oxygen content, the I–Vs become superlinear due to the higher activation energy of electrical conductivity [Eq. S2 in the supplementary material]. This limits the composition of the filament to below 1.3. Figure 2(c) shows the plot of the temperature distribution along the centerline of the filament. The distance scale starts at the interface with the top electrode. For the filament composition x = 0.4, the maximum temperature is 998 K and occurs close to the interfaces with electrodes. This is due to filament conductance being higher than that of the contacts. With increasing oxygen content and decreasing conductivity of the filament, the maximum temperature location shifts to the middle of the oxide layer and its value reaches 1065 K at x = 1.6. Since the total dissipated power is constant, the

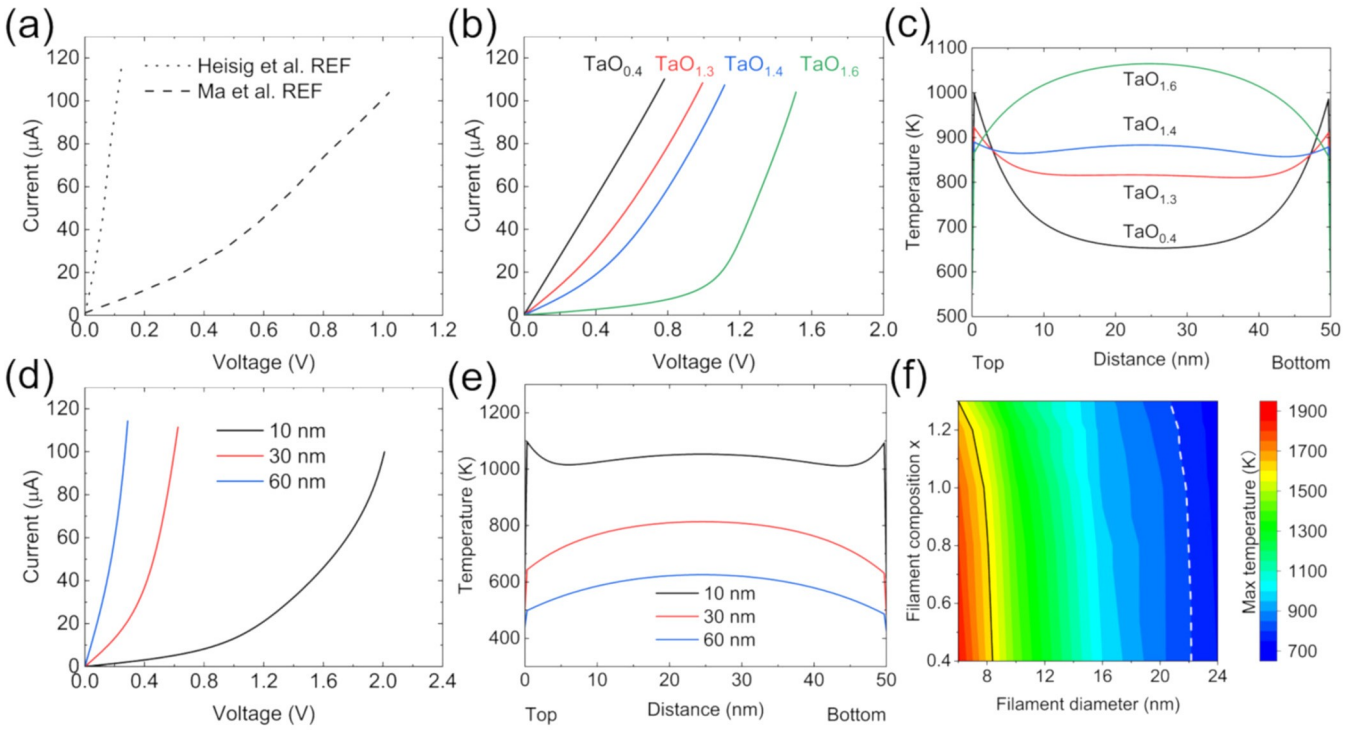

FIG. 2. Characteristics of LRS. (a) Experimental I–V curves after Ma et al.[25] and Heisig et al.[36] (b) Simulated I–V curves and (c) the temperature line profile along the center of the filament as a function of filament composition for d ¼ 16 nm. (d) I–V curves for x ¼ 1:5 as a function of filament diameter d. (e) Temperature along the line in the center of the filament (x ¼ 1:5) as a function of filament diameter d. (f) Maximum temperature with the device in LRS as a function of composition x and diameter d.

rising oxygen content lowers the current, the dissipated power at contacts, and the temperature close to the interfaces. Composition changes between 0.4 and 1.6 results in the temperature at the interfaces gradually dropping from 1065 to 866 K. Concomitantly, the temperature in the middle of the filament increased from 653 to 1065 K even though the current dropped by over 30%. This decrease was compensated by the conductivity falling by almost 90%. Throughout all changes, the maximum temperatures stayed within the acceptable 800–1600 K range. The maximum temperature decreases with an increase in oxygen context in the range of 0:4 x 1:3 and increase for 1:3 x 1:6. However, the position of the maximum temperature presents another limitation. Previous studies have shown that the gap in the filament was formed at the interface with the anode during the electroformation process.[25] This indicates that the temperature and ion diffusivity were the highest at this location or at least that both were more or less uniform throughout the filament. This is true for x 1:4 but is not the case in filaments with the higher oxygen content. At x ¼ 1:6, the difference of temperature in the middle of the filament and at the interface is about 200 K, indicating that ions are mobile and the gap should form in the center. The limits arrived at by analysis of Figs. 2(b) and 2(c) agree with the filament composition estimated by Ma et al.[25] ($TaO_{0.4}$) but not that of Heisig et al.[36] ($TaO_{1.9}$–$TaO_{2.1}$).

Figures 2(d) and 2(e) present the simulation results for the filament with fixed composition x ¼ 1:5 and varying diameter. The I–V characteristics shown in Fig. 2(d) indicate that quite obviously the conductance is increasing with the diameter and is accompanied by I–Vs becoming more linear. This softens somewhat the previous result limiting the composition of the filament to below x 1:3 allowing for somewhat higher oxygen content. However, increasing diameter rapidly lowers the maximum temperature within the filament [Fig. 2(e)]. The maximum temperature in the device with 10, 30, and 60 nm filament diameters are 1098, 814, and 626 K, respectively. This decrease is mostly due to a lower dissipated power density and a larger area of the filament surface. At the same time, the position of maximum temperature moves from the interface to the middle. Since the maximum temperature is lower than the lowest acceptable value and its position is wrong, the filament composition x ¼ 1:5 should be rejected even with larger filament diameters. The upper limit of filament composition should still be x ¼ 1:3.

Figure 2(f) is the map of the maximum temperature within the LRS filament with different diameters and compositions. The black solid line on the left side of the figure is an isotherm corresponding to the upper limit of acceptable temperatures while the dashed white line represents the lower temperature limit. The dominant trend is the decrease in temperature with increasing diameter

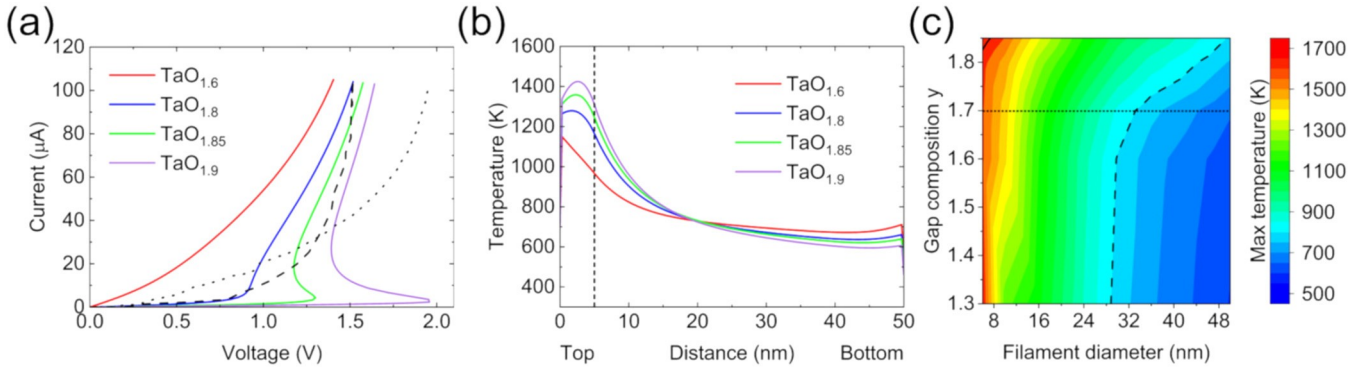

FIG. 3. Dependence of HRS I–V characteristics on the composition of the gap for width w ¼ 5 nm. (a) I–V curve (dashed line depicts experimental data of Ma et al.,[25] dotted line is based on Heisig et al.[26]). (b) Temperatures along the center line of the device (gap extends from 0 nm to the dashed line). (c) Map of the maximum temperature in the device.

due to lower power density and more efficient lateral heat extraction. The dependence on composition is much less pronounced for d 12 nm with the isotherms being almost vertical. The map indicates that the diameter of the filament should be between 6 and 22 nm.

The slight dependence of maximum temperature on composition can be understood in the following way: lowering the composition, while keeping the diameter and dissipated power constant, results in slightly higher electrical conductance and higher current density. This increases the dissipated power density and temperature. The effect is mostly due to increasing relative importance of contact resistance. The dependence of maximum temperature on composition is stronger for small filament diameters with a difference of 264 K for x ¼ 1:3 and x ¼ 0:4 for d ¼ 6 nm.

B. Modeling of high resistance state

Transmission electron microscopy (TEM) results from Ma et al. show that the filament diameter does not appreciably change during the set process.[25] Accordingly, we have assumed that the diameter remains the same in LRS and HRS with its size limits determined above for LRS need to be met in HRS as well. Figure 3 summarizes the results for the HRS as a function of the gap composition denoted here by y and constant diameter d ¼ 16 nm, gap width w ¼ 5 nm, and filament composition x ¼ 1:3. Panel (a) shows the simulated I–V curves as solid lines with the dashed and dotted lines corresponding to experimental HRS I–Vs from reports by Ma et al.[25] and Heisig et al.[26] The plots are limited to a voltage range that did not allow for ion motion and permanent changes of the I–V. Both experimental curves exhibit strongly superlinear dependence of current on voltage with the I–V slope becoming almost vertical at higher voltages.

The simulations show an increasing degree of nonlinearity with increasing oxygen content of the gap. For y 1:8, the curves showed a region of negative differential resistance (NDR), which becomes more pronounced for higher y. NDR regions are expected in the I–V characteristics of devices using functional material with conductivity increasing with temperature[52] and are well documented and understood in structures based on TaO_x.[52–54] It is due to the thermally activated conductivity of TaO_x and is more pronounced at higher oxygen content. With increasing composition, the overall resistance of the device is increasing moving the "knee" of the characteristics to higher voltages. When the y increases to 1.9, the critical voltage increases to near 2.0 V while the voltage at maximum current is only 1.6 V. The NDR region has never been reported in the LRS and HRS of resistive switches. This is a very strong indication that the composition x of the filament in LRS and the gap in the HRS is below 1.9. The acceptable gap composition y then should be between 1.7 and 1.85. The effects of the filament diameter on the HRS are minimal as discussed in the supplementary material.

Figure 3(b) shows the temperature distribution along the centerline of the filament as a function of the gap composition with other filament parameters the same as in Fig. 3(a). The dissipated power is the same for all curves. The gap is located at the interface with the top electrode starting at 0 and extending to the dashed vertical line. The maximum temperature in the device always occurs within the gap and moves from the top interface to the middle of the gap with increasing y. It is increasing with oxygen content and remains in the acceptable temperature range.

The maximum temperature in HRS as a function of filament diameter and gap composition is shown in Fig. 3(c) for constant width w ¼ 5 nm. The upper and lower limits of acceptable temperatures are marked by black solid and dashed lines (the upper limit only shown in the upper left corner of the figure). The dotted horizontal line at y ¼ 1:7 is the lower limit of the gap composition arrived at in the discussion of Fig. 3(a). Both filament diameter d and gap composition y have a significant influence on the maximum temperature. The diameter scale starts at 6 nm, which is the lowest value allowed by the LRS. For the filament diameter range 6–8 nm and the constant diameter, the maximum temperature initially decreases and then increases with the gap composition. This is the same effect as encountered in the discussion of LRS. At low gap composition y, the maximum temperature is

dominated by the Joule heating of the contact resistance. Accordingly, the maximum appears very close to the interface. At higher gap compositions, relatively more power is dissipated in the oxide with maximum temperature moving to the middle of the gap. The map does not impose any additional restrictions on the diameter of the filament beyond the limits obtained for the LRS.

In addition to the composition of the gap and filament diameter of the gap, the I–V characteristics could also be affected by gap width w and the composition of the trunk. The composition of the trunk does not change the results much and can be neglected (estimates are discussed in the supplementary material). The influence of gap width w is summarized in Fig. 4. The remaining parameters included the composition of the trunk x ¼ 1:3, gap composition y ¼ 1:8, and filament diameter d ¼ 16 nm.

Figure 4(a) shows the I–V characteristics for different gap widths. The investigated range was from 2 to 20 nm. It is possible that the gap is narrower than 2 nm but the finite element models can be used only for sizes much larger than the interatomic distance. On the large side, 20 nm appeared as a reasonable limit as the diffusion times would be higher than that of interest for memory applications. For small values of w, the voltage as a function of current increases monotonically while for w . 5 nm it shows a local minimum corresponding to the appearance of the NDR region. It should be noted that the I–V of the gap itself

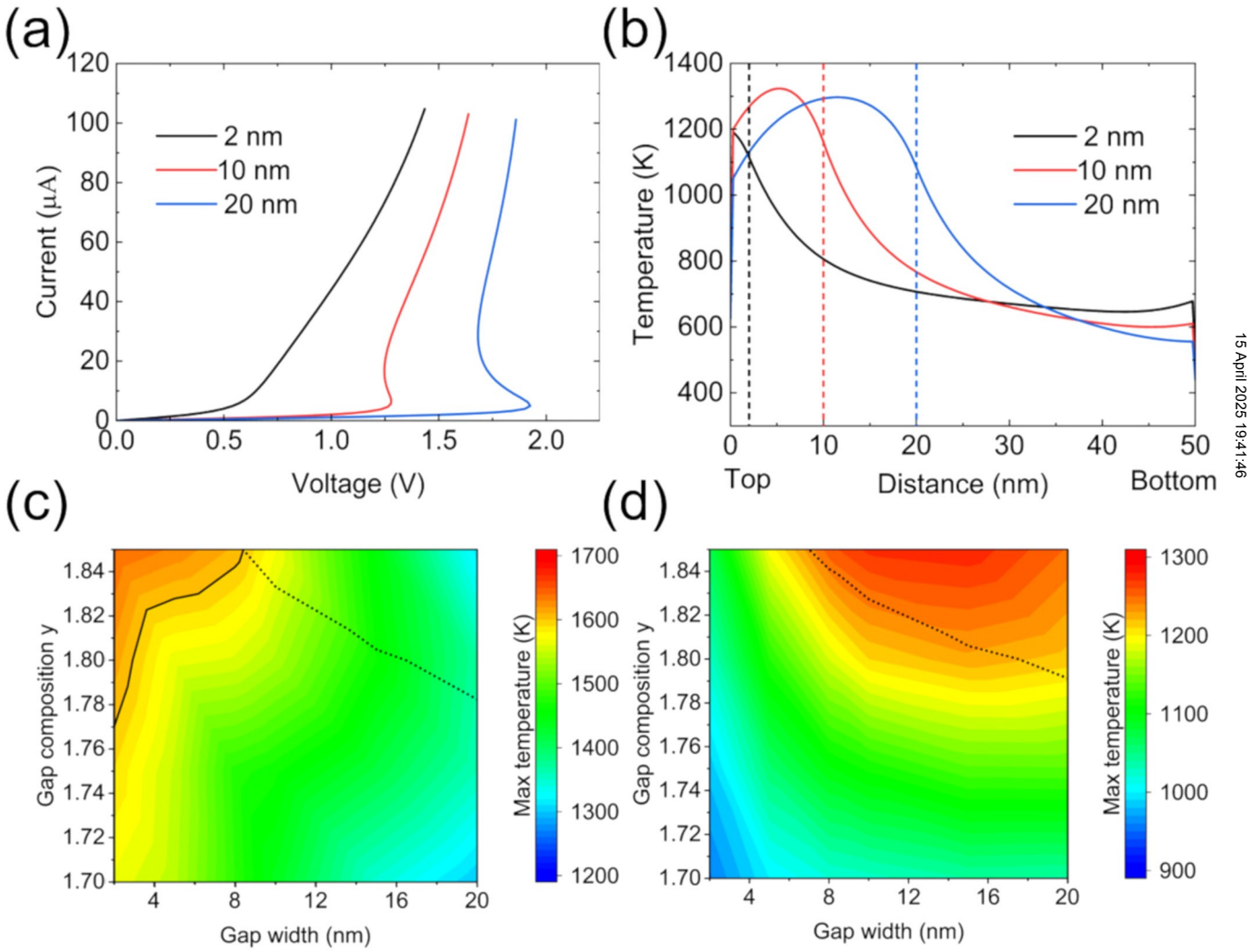

FIG. 4. (a) and (b) Dependence of LRS on gap width and composition for filament diameter d ¼ 16 nm, trunk composition x ¼ 1:3, and gap composition y ¼ 1:8. (a) I–V curves. (b) Temperature line profile in the center of the gap extends from 0 nm to the dashed line with the corresponding color. (c) and (d) Maps of maximum temperature as a function of gap width and composition for filament diameter of (c) 6 and (d) 22 nm. Note the different temperature ranges in (c) and (d).

always exhibits NDR with the threshold voltage increasing with an increase in its width. As the gap is connected in series with the fixed resistance of the trunk, the NDR becomes invisible for small values of the gap resistance.

Figure 4(b) shows the temperature distribution along the centerline of the filament with the edge of the gap shown as vertical dashed lines. The maximum temperature is reached close to the middle of gap except for small values of w when the heat generated at the contact dominated the temperature distribution. The changes of the gap width have a relatively minor effect on the maximum temperature with the temperature changing from 1200 to 1300 K for width changing from 2 to 20 nm. The maps of the maximum temperature as a function of gap composition y and filament diameter d are shown in Figs. 4(c) and 4(d) for the filament diameter 6 and 22 nm. This covers the widest acceptable diameter range. The solid and dotted black lines correspond to 1600 K isotherm and the NDR region boundary, respectively, with NDR appearing for gap compositions above the dotted line.

For the small filament diameter [Fig. 4(c)], the maximum temperature decreases with increasing gap width and decreasing gap composition (except for large w). The first effect is attributed to the decrease in current density, whereas the second is due to the change in the temperature distribution. At lower y, the conductance of the gap increases and the power dissipated within it decreases. This lowers the temperature in the gap while increasing it elsewhere in the filament. For gap widths below 8.5 nm, the high gap composition is excluded due to exceeding the upper temperature limit. Also, for gap widths above 8.5 nm, devices with a high oxygen content in the gap are excluded due to the appearance of the NRD. For the gap composition 1:7 $\lesssim$ y $\lesssim$ 1:77, there is no specific limitation on the gap width.

For the filament diameter at the upper limit of 22 nm [Fig. 4(d)], the temperatures within the device are lower than for small diameter and do not limit the gap composition or width. Similar as for d = 6 nm, the maximum temperature falls with the decreasing composition but first rises and then falls with a gap width consistent with the previously discussed results. The dotted line marking the appearance of NDR is very similar to the one in panel (c).

The model presented above assumed abrupt changes in composition both at the boundaries of the filament. It is an oversimplification that can somewhat affect the values of the extracted parameters but the estimated effect of gradual changes is small.

IV. DISCUSSION

The composition of the filament in resistive switching devices is typically discussed in terms of the concentration of oxygen vacancies with the typical concentrations assumed in the modeling reports in the 2×10^{20}–$1:2\times10^{21}$ cm^{3} range.[12,16,17,55] This corresponds to $TaO_{2.49}$–$TaO_{2.45}$ range in the notation used here. Only two modeling studies considered significantly lower oxygen content of TaO_2[14] and $TaO_{1.5}$.[43] The composition of the trunk of the filament estimated here is x = 1:3 or lower. In terms of relative oxygen concentration changes, the numbers above correspond to 0.3%, 2%, 20%, 40%, and 50% of missing oxygen compared to fully oxidized $TaO_{2.5}$. Such large differences originate from very different parameters in commonly assumed dependence of electrical conductivity on composition and temperature,

$$\sigma(x, T) = \sigma_0(x) e^{-\frac{E_{act}(x)}{kT}}. \quad (4)$$

In this report, we have used values of σ_0 and E_{act} obtained by fitting experimental values of conductivity measured in wide composition and temperature range.[31] Our activation energies ranged from 0.34 eV for oxygen deficiency equal 3×10^{21} cm^{3} to about 0 eV for oxygen deficiency 2.5×10^{22} cm^{3}. Most authors assumed activation energies corresponding to those of n-type silicon ($E_{act} = 0.05$ eV for donor densities n_D 0 and 0 for $n_D = 2\times10^{20}$ cm^{3}) or lower.

In parallel, most authors overestimated carrier mobility to compensate for using low carrier densities assumed to be the same as oxygen deficiency. The most direct test of the oxygen deficiency is spectroscopic methods. Their use in the case of nanometer size filament is difficult but recently published data are in reasonable agreement with our assumed values. Analysis of high angle annular dark field imaging estimated the composition of the filament at $x = 0.4$.[25] The x-ray photo-electron emission microscopy indicated $x = 2.0$,[36] and scanning nearfield optical microscopy gave $x = 1.2$.[56] The low temperature electrical conductivity measurements estimated x ffi 1.[32] While the above values appear to be scattered, they differ only by a factor of 4 and all are in the tens of percent range with an average of 50% or $TaO_{1.3}$.

Our composition estimate has very important consequences. The commonly assumed mechanism of the electroformation process is the diffusion of oxygen out of the filament volume either laterally or vertically. To get x ffi 1:3 requires the removal of a large fraction of oxygen atoms from the volume of the filament and is bound to produce large tensile stresses and associated large elastic strain energy. The strain contribution to the total energy of formation can be neglected for oxygen deficiency of 1% but not for 50% deficiency. In addition to elastic strain, one can expect plastic one as well resulting in the formation of a depression on the surface of the top electrode. Since such depressions have not been reported, it is likely that the stress is compensated by the in-diffusion of tantalum. This could be induced either by stress itself or by an independent driving force such as the Soret effect.[43,57]

One consequence of the TaO_x composition in the filament estimated above is the language we have used in most of this report. Specifically, we have used oxygen content or the oxide composition rather than vacancy concentration as frequently done in the literature. The point defect concentrations or oxygen deficiency given above are mentioned only to compare our estimates with that in other reports. Generally, it is inappropriate to talk about oxygen vacancies in a material with composition much different from stoichiometric. The $TaO_{1.3}$ structure is not the same as that of Ta_2O_5 with some oxygen sites left unoccupied.

An additional consequence of using electrical conductivity with large activation energies is the shape of I–V. Most experimental studies reported nonlinear I–V characteristics in HRS. Such nonlinearities have been modeled by assuming the presence of the Schottky barrier at the oxide–metal interfaces or by field-dependent conductivity of TaO$_x$. While it is certainly possible that such effects exist, the thermal effects offer a simpler explanation. The linearity

of I–V in LRS is due to low oxygen concentration throughout the filament and associated metallic conduction. An increase in temperature in the filament does not cause much change in conductivity. In HRS, the conductivity in the gap region exhibits significant activation energy and associated non-linear I–V.

Lastly, recent experimental results provide evidence for the phase separation in the filament.[57–59] The result is the appearance of Ta-rich and Ta-poor regions (conversely, oxygen poor and oxygen rich) within the filament and likely changes in the device resistance. At this point, we cannot offer a predictive model of resistance as a function of phase separation.

The last comment concerns the applicability of the results presented here to other material systems. While the limits of filament size and composition discussed here are not universal for resistive switches, it is likely that similar limits apply to another important material system, i.e., metal/HfO_2/metal structures. This is due to a similar dependence of HfO_2 conductivity on composition and temperature.[60]

V. CONCLUSIONS

A simple model of resistive switching devices consisting of a cylinder of reduced TaO_x with or without a gap depleted of donor-type defects was used to reproduce I–V characteristics of high and low resistance states and the temperature distribution therein. The effect of four parameters was analyzed including filament composition x, filament diameter d, gap composition y, and gap width w. Matching the shape of I–V with representative experimental data and constraining the maximum temperature to between 800 and 1600 K, we have limited the filament diameter to between 6 and 22 nm. The composition of the trunk of the filament needs to be $x \geq 1.3$ to give linear LRS I–V. The same constraints restricted the composition of the gap to $1.7 \leq x \leq 1.85$. The filament composition agrees with the most recent experimental estimates. The model also demonstrates that the nonlinear I–Vs in HRS can be explained by Joule heating without the necessity to invoke field-dependent conductivity or the presence of Schottky barriers at the interfaces.

SUPPLEMENTARY MATERIAL

See the supplementary material for details on COMSOL simulations and additional simulations.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (NSF) under Grant No. DMR-2208488.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Rongchen Li: Conceptualization (equal); Formal analysis (lead); Investigation (lead); Methodology (lead); Visualization (lead); Writing – original draft (lead); Writing – review & editing (equal). Yang Bai: Formal analysis (supporting); Methodology (supporting); Visualization (supporting); Writing – review & editing (supporting). Marek Skowronski: Conceptualization (equal); Funding acquisition (lead); Methodology (supporting); Project administration (lead); Supervision (lead); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] R. Waser, R. Dittmann, G. Staikov, and K. Szot, Adv. Mater. 21, 2632 (2009).
[2] D. Schor, "Intel expands 22FFL with production-ready RRAM and MRAM on FinFET," WikiChip Fuse (2019), available at https://fuse.wikichip.org/news/2801/intel-expands-22ffl-with-production-ready-rram-and-mram-on-finfet/.
[3] P. Clarke, "eeNewsEurope" (2020); available at https://www.eenewseurope.com/en/tsmc-offers-22nm-rram-taking-mram-on-to-16nm/.
[4] D. S. Jeong, H. Schroeder, and R. Waser, Phys. Rev. B 79, 195317 (2009).
[5] 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, Nat. Mater. 10, 625 (2011).
[6] T. Tan, T. Guo, X. Chen, X. Li, and Z. Liu, Appl. Surf. Sci. 317, 982 (2014).
[7] D. Carta, I. Salaoru, A. Khiat, A. Regoutz, C. Mitterbauer, N. M. Harrison, and T. Prodromakis, ACS Appl. Mater. Interfaces 8, 19605 (2016).
[8] S. Yu, Y. Wu, and H.-S. P. Wong, Appl. Phys. Lett. 98, 103514 (2011).
[9] Z. Jiang, Y. Wu, S. Yu, L. Yang, K. Song, Z. Karim, and H.-S. P. Wong, IEEE Trans. Electron Devices 63, 1884 (2016).
[10] M. Ismail, E. Ahmed, A. M. Rana, I. Talib, and M. Y. Nadeem, J. Alloys Compd. 646, 662 (2015).
[11] Y. S. Lin, F. Zeng, S. G. Tang, H. Y. Liu, C. Chen, S. Gao, Y. G. Wang, and F. Pan, J. Appl. Phys. 113, 064510 (2013).
[12] S. Kim, S. Choi, and W. Lu, ACS Nano 8, 2369 (2014).
[13] S. Menzel, M. Waters, A. Marchewka, U. Böttger, R. Dittmann, and R. Waser, Adv. Funct. Mater. 21, 4487 (2011).
[14] S. H. Lee, J. Moon, Y. Jeong, J. Lee, X. Li, H. Wu, and W. D. Lu, ACS Appl. Electron. Mater. 2, 701 (2020).
[15] F. Nardi, S. Larentis, S. Balatti, D. C. Gilmer, and D. Ielmini, IEEE Trans. Electron Devices 59, 2461 (2012).
[16] S. Larentis, F. Nardi, S. Balatti, D. C. Gilmer, and D. Ielmini, IEEE Trans. Electron Devices 59, 2468 (2012).
[17] A. Marchewka, B. Roesgen, K. Skaja, H. Du, C. Jia, J. Mayer, V. Rana, R. Waser, and S. Menzel, Adv. Electron. Mater. 2, 1500233 (2016).
[18] K. Zhang, Y. Ren, P. Ganesh, and Y. Cao, npj Comput. Mater. 8, 76 (2022).
[19] C. Baeumer, C. Funck, A. Locatelli, T. O. Menteş, F. Genuzio, T. Heisig, F. Hensling, N. Raab, C. M. Schneider, S. Menzel, R. Waser, and R. Dittmann, Nano Lett. 19, 54 (2019).
[20] C. Funck and S. Menzel, ACS Appl. Electron. Mater. 3, 3674 (2021).
[21] P. Huang, X. Y. Liu, B. Chen, H. T. Li, Y. J. Wang, Y. X. Deng, K. L. Wei, L. Zeng, B. Gao, G. Du, X. Zhang, and J. F. Kang, IEEE Trans. Electron Devices 60, 4090 (2013).
[22] K. M. Kim and C. S. Hwang, Appl. Phys. Lett. 94, 122109 (2009).
[23] D.-H. Kwon, K. M. Kim, J. H. Jang, J. M. Jeon, M. H. Lee, G. H. Kim, X.-S. Li, G.-S. Park, B. Lee, S. Han, M. Kim, and C. S. Hwang, Nat. Nanotechnol. 5, 148 (2010).
[24] C.-T. Tung, C. Kumar Dabhi, S. Salahuddin, and C. Hu, Solid-State Electron. 220, 108989 (2024).
[25] Y. Ma, J. M. Goodwill, D. Li, D. A. Cullen, J. D. Poplawsky, K. L. More, J. A. Bain, and M. Skowronski, Adv. Electron. Mater. 5, 1800954 (2019).

[26]T. V. Perevalov, V. A. Gritsenko, A. A. Gismatulin, V. A. Voronkovskii, A. K. Gerasimova, V. S. Aliev, and I. A. Prosvirin, Nanotechnology 29, 264001 (2018).
[27]J. Frenkel, Phys. Rev. 54, 647 (1938).
[28]R. M. Hill, Philos. Mag. 23, 59 (1971).
[29]S. Makram-Ebeid and M. Lannoo, Phys. Rev. B 25, 6406 (1982).
[30]K. A. Nasyrov and V. A. Gritsenko, J. Appl. Phys. 109, 093705 (2011).
[31]K. Bao, J. Meng, J. D. Poplawsky, and M. Skowronski, J. Non-Cryst. Solids 617, 122495 (2023).
[32]C. M. M. Rosário, B. Thöner, A. Schönhals, S. Menzel, A. Meledin, N. P. Barradas, E. Alves, J. Mayer, M. Wuttig, R. Waser, N. A. Sobolev, and D. J. Wouters, Nanoscale 11, 16978 (2019).
[33]R. J. Bondi, M. P. Desjarlais, A. P. Thompson, G. L. Brennecka, and M. J. Marinella, J. Appl. Phys. 114, 203701 (2013).
[34]A. Ascoli, S. Menzel, V. Rana, T. Kempen, I. Messaris, A. S. Demirkol, M. Schulten, A. Siemon, and R. Tetzlaff, Adv. Electron. Mater. 8, 2200182 (2022).
[35]H.-K. Lyeo and D. G. Cahill, Phys. Rev. B 73, 144301 (2006).
[36]T. Heisig, K. Lange, A. Gutsche, K. T. Goß, S. Hambsch, A. Locatelli, T. O. Menteş, F. Genuzio, S. Menzel, and R. Dittmann, Adv. Electron. Mater. 8, 2100936 (2022).
[37]P. R. Mickel, A. J. Lohn, B. Joon Choi, J. Joshua Yang, M.-X. Zhang, M. J. Marinella, C. D. James, and R. Stanley Williams, Appl. Phys. Lett. 102, 223502 (2013).
[38]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, Adv. Mater. 23, 5633 (2011).
[39]F. Miao, W. Yi, I. Goldfarb, J. J. Yang, M.-X. Zhang, M. D. Pickett, J. P. Strachan, G. Medeiros-Ribeiro, and R. S. Williams, ACS Nano 6, 2312 (2012).
[40]E. Yalon, S. Cohen, A. Gavrilov, and D. Ritter, Nanotechnology 23, 465201 (2012).
[41]J. Kwon, A. A. Sharma, C.-Y. Chen, A. Fantini, M. Jurczak, A. A. Herzing, J. A. Bain, Y. N. Picard, and M. Skowronski, ACS Appl. Mater. Interfaces 8, 20176 (2016).
[42]A. A. Sharma, M. Noman, M. Skowronski, and J. A. Bain, in Proceedings of Technical Program - 2014 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA), Hsinchu, Taiwan (IEEE, 2014), pp. 1–2.
[43]J. Meng, E. Lian, J. D. Poplawsky, and M. Skowronski, Phys. Rev. Appl. 17, 054040 (2022).
[44]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, Sci. Rep. 3, 1680 (2013).
[45]U. N. Gries, H. Schraknepper, K. Skaja, F. Gunkel, S. Hoffmann-Eifert, R. Waser, and R. A. De Souza, Phys. Chem. Chem. Phys. 20, 989 (2018).
[46]S. P. Garg, N. Krishnamurthy, A. Awasthi, and M. Venkatraman, J. Phase Equilib. 17, 63 (1996).
[47]Z. Wei, Y. Kanzawa, K. Arita, Y. Katoh, K. Kawai, S. Muraoka, S. Mitani, S. Fujii, K. Katayama, M. Iijima, T. Mikawa, T. Ninomiya, R. Miyanaga, Y. Kawashima, K. Tsuji, A. Himeno, T. Okada, R. Azuma, K. Shimakawa, H. Sugaya, T. Takagi, R. Yasuhara, K. Horiba, H. Kumigashira, and M. Oshima, in 2008 IEEE International Electron Devices Meeting, San Francisco, CA, USA (IEEE, 2008), pp. 1–4.
[48]A. Prakash, S. Maikap, C. S. Lai, H. Y. Lee, W. S. Chen, F. T. Chen, M. J. Kao, and M. J. Tsai, Jpn. J. Appl. Phys. 51, 04DD06 (2012).
[49]J. J. Yang, M.-X. Zhang, J. P. Strachan, F. Miao, M. D. Pickett, R. D. Kelley, G. Medeiros-Ribeiro, and R. S. Williams, Appl. Phys. Lett. 97, 232102 (2010).
[50]A. L. Jagath, T. N. Kumar, and H. A. F. Almurib, in 2019 IEEE 9th International Nanoelectronics Conference (INEC), Kuching, Malaysia (IEEE, 2019), pp. 1–5.
[51]S. Sihn, W. L. Chambers, M. Abedin, K. Beckmann, N. Cady, S. Ganguli, and A. K. Roy, Small 2310542 (2024).
[52]G. A. Gibson, Adv Funct. Mater. 28, 1704175 (2018).
[53]J. M. Goodwill, A. A. Sharma, D. Li, J. A. Bain, and M. Skowronski, ACS Appl. Mater. Interfaces 9, 11704 (2017).
[54]J. M. Goodwill, D. K. Gala, J. A. Bain, and M. Skowronski, J. Appl. Phys. 123, 115105 (2018).
[55]K. Zhang, J. Wang, Y. Huang, L.-Q. Chen, P. Ganesh, and Y. Cao, npj Comput. Mater. 6, 198 (2020).
[56]K. Kim, G. Wirth, K. Goss, T. Heisig, C. Bauerschmidt, A. Hessler, H. Li, L. Waldecker, R. Dittmann, and T. Taubner, Adv. Funct. Mater. 34, 2312980 (2024).
[57]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, ACS Appl. Mater. Interfaces 10, 23187 (2018).
[58]M. Skowronski, in 2023 IEEE International Reliability Physics Symposium (IRPS), Monterey, CA, USA (IEEE, 2023), pp. 1–5.
[59]J. Li, A. Appachar, S. L. Peczonczyk, E. T. Harrison, A. V. Ievlev, R. Hood, D. Shin, S. Yoo, B. Roest, K. Sun, K. Beckmann, O. Popova, T. Chiang, W. S. Wahby, R. B. Jacobs-Godrim, M. J. Marinella, P. Maksymovych, J. T. Heron, N. Cady, W. DLu, S. Kumar, A. A. Talin, W. Sun, and Y. Li, Matter 7, 3970–3993 (2024).
[60]N. M. Tallan, W. C. Tripp, and R. W. Vest, J. Am. Ceram. Soc. 50, 279 (1967).