Kinetic Monte Carlo Simulation of Interface-Controlled Hafnia-Based Resistive Memory

Xu Xu, Bipin Rajendran, and M. P. Anantram

***Abstract*— Kinetic Monte Carlo simulations of resistive memory devices have been performed by paying attention to the vacancy–interstitial generation near the Hafnia-metal electrode interface. In our model, an oxygen vacancy is generated in Hafnia near the interface, with the corresponding oxygen atom residing in the metal electrode. These oxygen atoms form a thin insulating oxide layer at the Hafnia-active electrode interface. This interfacial layer is essential to thicken the filament, even after the filament bridges the two metal electrodes at low current levels. This thickening of the conducting filament is captured by the model and it naturally explains the trend of resistance decrease with an increase in compliance current found in some experiments. Simulations results as a function of the bonding energy between vacancies show a large increase in retention time with an increase in bonding energy. We also find that as the compliance current increases, the morphology of the filament transitions from conical to dumbbell-shaped. Finally, using a single set of values for various energies, our simulations capture the SET, RESET, and retention processes.**

***Index Terms*— Kinetic Monte Carlo (KMC), resistive random access memory (RRAM), resistive switching.**

I. Introduction

RESISTIVE random access memory (RRAM) devices are promising candidates to replace Flash memory, and their memory switching characteristics have been demonstrated in a variety of material systems [1]–[12]. RRAM devices can be integrated in a crossbar architecture, enabling the realization of large and dense memory arrays. The switching speeds of these devices are high (less than 10-ns switching time [11]), and they exhibit long retention times (over 10^5 s at 100 °C [6]). Recent work also points to the potential of using RRAM devices for neuromorphic computing [2], [9], [13], and for efficiently realizing logic operations [1], [14]–[18]. All these applications require a comprehensive physics-based understanding and a physical model that captures the dynamics of the filament formation process, including its stability and switching times between resistance levels. The model should have the ability to simulate statistically relevant device parameters such as resistance values and switching times because of the stochastic nature of filament formation.

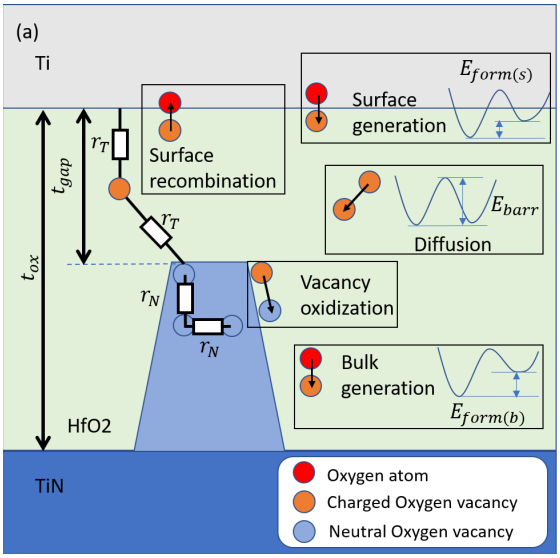


Fig. 1. Device structure with a partially formed filament. The various formation and diffusion processes are also shown. The top electrode is made of an active metal, whereas the bottom electrode is made of a passive metal.

Although prior Monte Carlo modeling work shows filament formation, it has been difficult to capture the decrease in resistance with increase in compliance current that is observed in some experiments [16], [19], [20], [21]. We will now discuss the possible reasons for this. Before the filament bridges the two electrodes, the electric field in the region of maximum filament growth (tip of the filament) is much larger than $V_{\text{bias}}/t_{\text{ox}}$ (a closer approximation for the electric field is $V_{\text{bias}}/t_{\text{gap}}$). Here, V_{bias} is the applied voltage, t_{ox} is the oxide thickness, and t_{gap} is the distance between the partially formed filament and the electrode (Fig. 1). Once the filament bridges the two electrodes, the electric field in the oxide becomes more uniform, is smaller in magnitude ($\sim V_{\text{bias}}/t_{\text{ox}}$), and is almost perpendicular to the Hafnia-electrode interface (figure shown in Section III). As the electric field in the lateral (along the oxide-electrode interface) direction is small, vacancies are unlikely to move in the lateral direction. Therefore, the growth of the filament slows down once the filament bridges the two electrodes even when current levels are smaller than the compliance current. As a result, the device resistance does not decrease, if the magnitude of the

Manuscript received September 27, 2019; revised November 8, 2019; accepted November 13, 2019. Date of publication December 16, 2019; date of current version December 30, 2019. The work of X. Xu and M. P. Anantram was supported by Winbond Inc. The review of this article was arranged by Editor S. Hong. *(Corresponding author: M. P. Anantram.)*

X. Xu and M. P. Anantram are with the ECE Department, University of Washington, Seattle, WA 98195-2500 USA (e-mail: axuuw@uw.edu; anant@uw.edu).

B. Rajendran is with the ECE Department, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: rajendran.bipin@gmail.com).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2019.2953917

0018-9383 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: University of Washington Libraries. Downloaded on November 19,2020 at 20:07:28 UTC from IEEE Xplore. Restrictions apply.

compliance current is increased (figure shown in Section III). In this article, we discuss a novel physical mechanism to explain the relationship between current compliance and resistance. The new mechanism includes the surface generation of vacancy and the corresponding oxidization of the electrode.

The Kinetic Monte Carlo (KMC) method has been used to simulate resistive switching before and the models implemented in prior work are varied [22]. Aldana *et al.* [19] and Menzel *et al.* [23] showed that the migration of metallic atoms was responsible for resistive switching in Al_2O_3/Cu and HfO_2/Ni systems, respectively. Bersuker *et al.* [10], an early work that the community draws upon, used a vacancy generation rate that is dependent on both the electric field and local temperature. The generation primarily occurred at the intersection of grain boundaries in Hafnia; the role of the oxygen atom generated was neglected. Huang *et al.* [24] considered the generation of a vacancy in the bulk oxide in the context of a 2-D model. Guan *et al.* [25] and Yu *et al.* [26] modeled the current using a tunneling formalism while the vacancy generation and recombination rates are determined by the concentration of oxygen ions. Berco and Tseng [27] used a free energy formulation to model filament growth and used a tunneling formalism to calculate the current. Although these works all provide insights, the relationship between the resistive memory cell's resistance and current compliance remains unclear; this provides motivation for our work. A second motivation for our work stems from the fact that it is not clear from the experimental work if oxygen-interstitial pairs are generated in the bulk of the oxide or at the oxide–metal interface. *Ab initio* calculations [28]–[30] show that the formation energy for the process where the oxygen atom resides in the Ti electrode and the vacancy in the oxide (generation at the interface) is much lower than the formation energy for a Frenkel pair in the bulk oxide. A recent calculation shows that the formation energy in the bulk can be lowered when the oxide is charged [31]. This is an interesting proposal, which requires the evaluation of the population of electrons in the large bandgap oxide. In experiments where this is the case, the bulk generation event hardly occurs, and surface generation should be responsible for filament formation.

Our work in this article provides an analysis of the possibility of surface generation explaining the characteristics of Hafnia-based RRAM. We propose a mechanism for the experimentally observed decrease in resistance of the memory cell with an increase in compliance current that has been observed in some experiments. The oxygen atoms generated with the vacancies near the metal–oxide interface can partially oxidize the metal electrode and form a thin layer with high resistivity. As a result, the electric field points into the filament instead of perpendicular to the metal-oxide interface. This electric field enhancement helps the rapid growth of a thicker filament. Concomitantly, our model where surface generation plays a dominant role also provides a unified description of the set, reset, and retention processes. This article is organized as follows: in Section II of the article, we describe the model in detail. The results are presented in Section III. We conclude the article in Section IV.

II. Model

We now describe a 3-D model for HfO_2-based RRAM devices, based on the physics underlying oxygen vacancy–interstitial generation, recombination, and diffusion. The KMC method is used [2]–[4], where the oxygen interstitial (O) and vacancy (Vo) diffusion rates are determined by Arrhenius law. The local temperature and electric field that determines the diffusion rates are obtained by solving the time-independent heat and Poisson's equations. The current is calculated using the resistive network model that is outlined in [3].

The metal–Hafnia–metal structure model is shown in Fig. 1. The bottom electrode is an inert electrode, which is usually made of TiN or Pt. The top electrode is called an active electrode because it is made with a metal such as Ti which can react with oxygen. Oxygen vacancies can be generated both in the interior of the oxide and at the oxide–active electrode interface. The oxide thickness used in our study is 8 nm. We are using the cubic lattice where only one oxygen vacancy per unit cell can exist. The oxygen vacancy can diffuse to its six nearest neighbors. The Hafnium atoms are not explicitly included. The grid spacing is 0.5 nm in the x-, y-, and z-directions. The vacancy generation in the interior of the oxide forms a Frenkel pair, where one vacancy is generated together with its paired interstitial oxygen. On the other hand, the generation process at the HfO_2/Ti interface involves an oxygen atom being placed directly in the Ti electrode and the vacancy residing in the oxide (Fig. 1). The rate of vacancy-interstitial generation at the Hafnia–Ti interface (called surface generation) is given by

$$R_{G(s)} = R_0 \exp\left(-\frac{E_{\text{form}(s)}(V_{\text{bias}}) - qFd/2}{kT}\right).$$

The bulk generation rate has a similar form and is given by

$$R_{G(b)} = R_0 \exp\left(-\frac{E_{\text{form}(b)}(V_{\text{bias}}) - qFd/2}{kT}\right).$$

R_0 is the attempt frequency, $E_{\text{form}(s)}(V_{\text{bias}})(E_{\text{form}(b)}(V_{\text{bias}}))$ is the surface (bulk) formation energy, k is Boltzmann's constant, and T is the local temperature. Here, q is the charge of the vacancy, F is the local electric field, and d is the lattice constant of HfO_2. The formation energy depends on the applied voltage, $E_{\text{form}(s/b)}(V_{\text{bias}}) = E_{\text{form}(s/b)}(0) - \alpha V_{\text{bias}}$, because of the field enhancement factor in McPherson's theory [32]. Since the Fermi level on the Ti electrode is lower than the defect energy level of the oxygen vacancy, the generated vacancies are likely to be positively charged, and therefore, their diffusion will be guided by the electric field. The diffusion rate for vacancies is given by [3]–[5]

$$R_D = R_0 \exp\left(-\frac{E_{\text{barr}}(q) - qFd/2}{kT}\right)$$

and the diffusion rate for interstitials is given by the same equation with the same parameters. Here, the migration barrier for vacancies $E_{\text{barr}}(q)$ depends on the vacancy's charge q (which is $+2$ for the isolated vacancies [28]). After the vacancies attach to the inert anode, electrons can be injected into the vacancy energy levels, making them neutral. As the

Authorized licensed use limited to: University of Washington Libraries. Downloaded on November 19,2020 at 20:07:28 UTC from IEEE Xplore. Restrictions apply.

interaction between the vacancy charge and the surrounding ions is Coulombic in origin, the ability of the vacancy to nudge the ions in the crystal away and then move depends on the vacancy charge. Guy *et al.* [6] and Padovani *et al.* [7] showed that the positively charged vacancy has a lower migration barrier than the neutral one.

As mentioned before, our work is also motivated by *ab initio* calculations [6], [29] which show that the formation energy for surface generation at Hafnia–Ti interface is much lower (2.2 eV) than the formation energy of a vacancy–interstitial pair in the bulk oxide (6.8 eV). The factor α in (1) and (2) is taken to be 0.4e. In experiments where this is the case, the bulk generation rate is many orders of magnitude smaller than the surface generation rate. A valid question then is if a surface generation-based model can explain the RRAM characteristics. Our goal here is to study the possibility of surface generation explaining the characteristics of Hafnia-based RRAM. We hypothesize that the metal in the active electrode can be oxidized by the oxygen atoms generated at the Hafina–Ti interface, and therefore, a thin insulating layer of TiO_x is formed near the Hafnia–Ti interface [9], [10].

In this model, the center of the simulation box is taken to be a region with small formation energy near the Ti–Hafnia interface. The smaller formation energy may be due to crystal structure variation, such as grain boundary intersection [2] or mechanical stress [4]. The low formation energy helps in filament growth in the region of the grain boundary. The charge on an isolated vacancy is taken to be +2e since the vacancy energy levels are higher than the Fermi level on the electrode. When the vacancy is attached to the inert electrode or other neutral vacancies which are electrically connected to the TiN electrode (this electrode is at a lower voltage during SET), electrons can fill the vacancy energy levels, and the vacancy charge is taken to be zero. Since the electrons are injected from the anode, vacancies near the negative electrode have a very small net charge and their activation energy for diffusion is high. On the other hand, because of the voltage drop in the filament, vacancies near the Ti have a higher probability to be charged and hence have a smaller activation barrier for diffusion. The parameters in our simulation are guided by prior calculations and are taken to be $E_{\text{form}(b)} = 6.8\,\text{eV}$ (formation energy in the bulk oxide) [6], [13] and $E_{\text{form}(s)} = 2.2\,\text{eV}$ (formation energy at the interface) [6], [29]. In the resistive network [3], the resistance between vacancies on the same island is $r_N = 1\,\text{k}\Omega$ and tunneling resistance is modeled as $r_T = r_{T0}\exp(\alpha L/d)$, where L is the minimum distance between islands and $\alpha = 1.5$, $r_{T0} = 1\,\text{k}\Omega$ (Fig. 1 shows these resistances). For Poisson's equation, Dirichlet's condition is applied to the electrode. The filament is represented by a large imaginary dielectric constant (1000i) following the classic work of [33] to represent its conductive nature. The thermal conductivity of HfO_2 is taken to be $1.1\,\text{W/m}\cdot\text{K}$, and periodic boundary conditions are applied in the lateral direction (left and right boundaries). On the oxide–electrode interfaces, mixed boundary conditions, $T + L(\partial T)/(\partial x) = 300\,\text{K}$ are applied to solve for the temperature distribution, where $L$is 2.5 nm [10]. The dielectric constant is 25 for HfO_2. The calculation shows that the

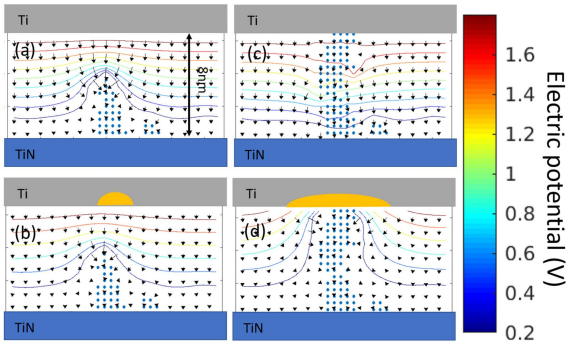


Fig. 2. (a) and (c) Contour plot of electric potential at two instants of time during the evolution of the filament without the TiO_x layer. (b) and (d) Electric potential for the identical filaments as in (a) and (c), respectively, with the TiO_x layer (yellow). The electric field is enhanced near the tip of filament only in (d). The arrows represent the electric field.

vacancy cluster has lower formation energy than the isolated ones [8], [14], which suggests a positive bonding interaction between vacancies. The bonding energy (E_{bond}) is the energy needed to separate two vacancies after clustering together. We show results for a high and a low value of E_{bond} because there are no clear experimental estimates. $E_{\text{barr}} = 1.1\,\text{eV}$ is the energy barrier for vacancy diffusion [29].

III. Results

We first describe the effect of the thin TiO_x layer on the electric field in the device, starting from the high resistance state. We perform Monte Carlo simulations without the TiO_x layer and plot the electric field distribution at two different time instants in Fig. 2(a) and (c), along with the corresponding equipotential lines. For the same filaments in Fig. 2(a) and (c), we show the electric field distribution with a TiO_x layer (yellow region) in Fig. 2(b) and (d). The calculations show that when the filament is small, the electric field in the device is not affected significantly by the TiO_x layer. This feature is clearly seen by comparing Fig. 2(a) and (b). However, after the filament has bridged the two electrodes, the electric field with the TiO_x layer is larger at the tip of the filament because the TiO_x layer is less conductive than the filament. In the Poisson solver, the TiO_x layer is modeled as a 2-nm thick insulating layer. The electric field without the TiO_x layer is more uniform as shown in Fig. 2(c). Note that the electric field enhancement near the filament tip implies that vacancies are formed more easily there. These vacancies diffuse to thicken the filament as shown in Fig. 2(d). Therefore, the resistance of the switched device, which is roughly inversely proportional to filament thickness, decreases with an increase in compliance current and enabling multilevel programmability in the RRAM cell.

The typical switching characteristics (current and voltage versus time) for a device with an 8-nm thick oxide are shown in Fig. 3 when the applied bias is 2 V. The current through the memory cell abruptly rises at ~200 μs, and the voltage across the oxide drops once current compliance is reached. At low current compliance (100 μA), we find that the shape of the filament is conical as seen in the insets of Fig. 3(a). In contrast, at higher current compliance (200 μA), the filament develops a dumbbell shape, with an overall thickening of the filament

Authorized licensed use limited to: University of Washington Libraries. Downloaded on November 19,2020 at 20:07:28 UTC from IEEE Xplore. Restrictions apply.

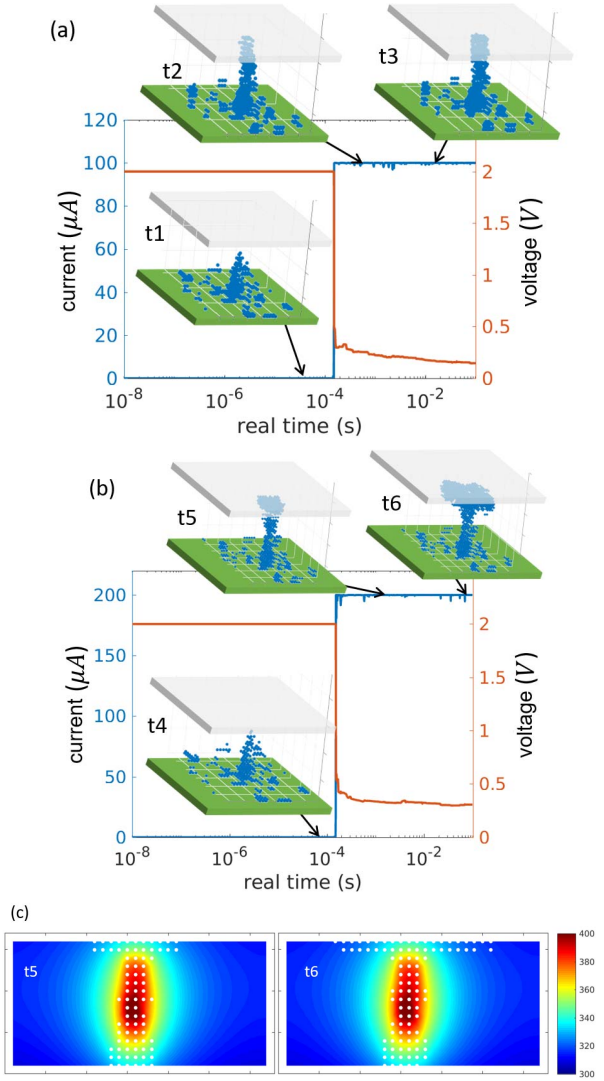


Fig. 3. Current (blue) and voltage (orange) versus time for (a) 100 μA and (b) 200 μA compliance currents. Insets show the shape of filament at different stages of formation. (c) Temperature distribution at the central cross section of the filament. Time is as marked in the insets of (a) and (b), respectively.

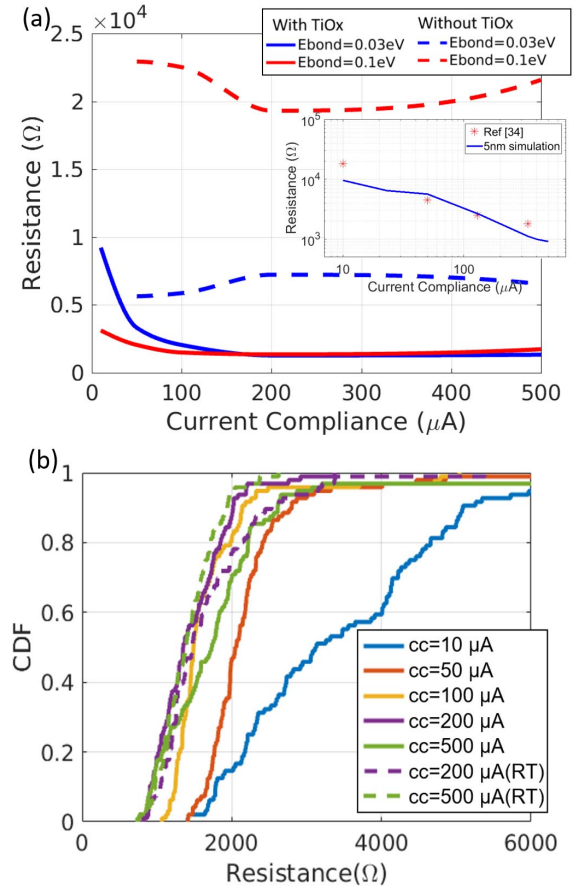


Fig. 4. (a) Filament resistance versus compliance current. Inset: The filament resistance versus compliance current when oxide thickness is 5 nm is in overall agreement with the experimental work of [34]. (b) CDF of $E_{\text{bond}} = 0.1$ eV.

as shown in Fig. 3(b). The temperature distribution is shown in Fig. 3(c). The hottest region is located at the most resistive part of the filament. The maximum temperature is less than 400 K after switching has occurred for current compliance of 200 μA.

An important behavior of RRAM is that it can be programed to intermediate resistance levels in a single cell. Usually, intermediate resistance levels can be controlled by applying different current compliances, which results in the switching layer having a different resistance. The mean value for resistance versus current compliance is shown in Fig. 4(a), for different values of E_{bond}. We observe that for both large (0.1 eV) and small (0.03 eV) E_{bond} values, the mean resistance decreases as the compliance current increases. However, there are two important differences. First, for large E_{bond}, the resistance values are much higher for a given compliance current (see from 100 to 500 μA). Second, for large E_{bond}, the decrease in resistance when the compliance current changes from 100 to 500 μA is negligible. On the other hand, for small $E_{\text{bond}}(0.03\,\text{eV})$, the resistance continues to decrease with the magnitude of compliance current. To demonstrate the importance of the TiO_x region, we compare our results using the same parameters without a TiO_x layer on top of the filament. These results are shown by the dashed lines of Fig. 4(a). The inverse scaling of resistance seen with the TiO_x layer is not seen in these simulations. The reason for this is that once the filament bridges the electrodes, the electric field is no longer much larger than $V_{\text{bias}}/t_{\text{ox}}$. We now discuss the reason for the resistance saturation with compliance current seen in the case of the large bonding energy in Fig. 4(a). When the bonding energy is large, the electric field is not effective in pushing the vacancies because it takes large energy to break the bond between vacancies. Therefore, the resistance stops decreasing in the high current compliance regime. On the other hand, with smaller bond energy, the vacancies diffuse more freely, which helps to thicken the filament resulting in a corresponding decrease in resistance. In calculations, we compute the current through the filament at a given voltage across the device. If the computed current is found to be higher than the compliance current, the voltage across the device is decreased in the next few computational steps so as to not exceed the compliance current value. Our assumption mimics a memory cell in series with a MOSFET where the source–drain

Authorized licensed use limited to: University of Washington Libraries. Downloaded on November 19,2020 at 20:07:28 UTC from IEEE Xplore. Restrictions apply.

voltage drop across the MOSFET is zero before the current is over compliance.

Our calculations compare favorably to the experiments [34] for a 5-nm thick oxide as shown in the inset of Fig. 4(a). We note that in the inset of Fig. 4(a), the resistance continues to drop significantly for larger current compliances in comparison to the thicker oxide devices modeled. The underlying reason is that the local temperature is dependent on the oxide thickness.

Our calculations of formation time as a function of applied bias also compare well with [34] and [35] (results not shown).

Another important device metric is the distribution of resistance values. We plot the cumulative distribution function (CDF) of the device resistance at different current compliances. In the simulations, the randomly distributed vacancies in the device vary, and the seeds of the random number generator are changed from sample to sample. We find that at compliance currents of 100 and 200 μA, the lower 90% of resistance values cluster over a narrow range, and the higher 10% of resistance values have a wider distribution. When the compliance current is high (500 μA), the distribution of resistance values is wider. The wider distribution of resistance stems from the larger local Joule heating in the filament at high current compliance, which leads to a variation in the precise geometry of the narrow spot. We contrast this result to the resistance distribution without Joule heating in Fig. 4(b). When Joule heating is artificially turned off, the resistance decreases with an increase in compliance current and the resistance distribution is narrower at large current compliance [dashed line in Fig. 4(b)]. We note that the CDF remains qualitatively similar at the smaller binding energies considered (not shown).

The increase in resistance at the largest compliance current of 500 uA considered is due to the high temperature. The high temperature causes the resistance spread to be larger because vacancies are more mobile. An interesting question is what should happen in an experiment (That is, "Should the tail at high compliance current increase?"). Our claim in this article is that: 1) as the compliance current increases, the resistance spread will initially decrease as seen in Fig. 4(b); but 2) as the current compliance becomes very high (perhaps close to device breakdown), there should be an increase in resistance spread (that is, a larger tail in resistance). The precise value of compliance current when this would occur would depend on the device details. We now model the bipolar RESET process which occurs when a negative voltage is applied to the Ti electrode. RESET is modeled with an applied bias of -2 V, as shown in Fig. 5. The device is RESET in a short time and the current decreases smoothly with time. The TiO_x layer serves as a reservoir of oxygen, and when an oxygen vacancy is close to the interface, the recombination of vacancy and oxygen occurs: oxygen in the TiO_x reservoir diffuses back to the oxide and fills the oxygen vacancy. The diffusion process of oxygen going back to fill the vacancy is fast as the recombination process is energy favorable and the electric field direction aids recombination. Moreover, the insulating gap helps to enhance the electric field between the filament and metal electrode, and therefore, oxygen atoms in the insulating layer can overcome the barrier from the metal [8], [15].

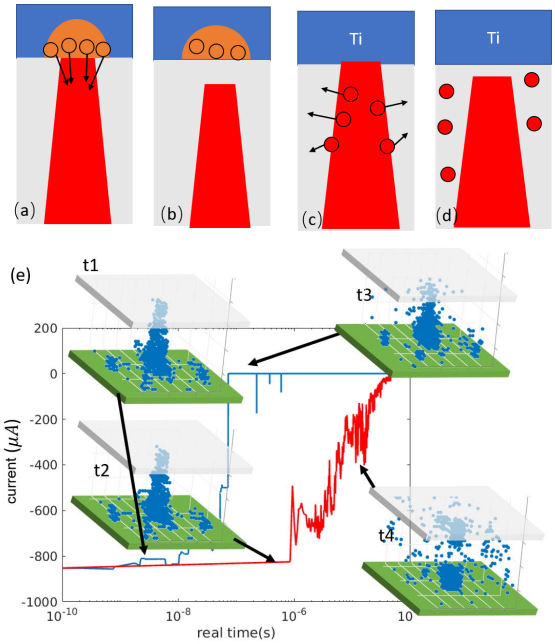


Fig. 5. RESET, with and without recombination. (a) and (b) are for the model where the oxygen atom can migrate back to the oxide and form a high resistance state. (c) and (d) are for the model where vacancies move away from the filament due to Joule heating (oxygen does not diffuse back from Ti here). (e) Blue curve shows the current in the RESET process where oxygen in Ti diffuses back to the oxide and recombines with the vacancy, which leads to a RESET time of $\sim 3 \times 10^{-8}$ s. If the oxygen in Ti cannot diffuse back, the memory cell can still be reset with Joule heating as shown in the red curve but the RESET time is longer.

These oxygen atoms can recombine with the vacancy to form the insulating oxide at the filament tip. Therefore, the resistance increases [Fig. 5(a) and (b)]. Our simulation shows that the current can jump to several intermediate values, and then the cell is RESET to a high resistance state. The RESET process occurs in relatively short time scales, as few oxygen atoms are sufficient to fill the vacancy in the filament. To analyze the RESET process further, we artificially removed the process of oxygen-vacancy recombination at the Ti–Hafnia interface in the simulations to test if RESET can occur without this recombination. We find that resistance can increase but the underlying mechanism involves the filament breaking due to Joule heating. The temperature gradient established due to Joule heating causes the vacancies to diffuse away from the filament. Our simulations show that the RESET time is longer when surface recombination is turned off as shown in Fig. 5(e). The temperature here can reach 1800 K. The simulation shows that the current fluctuates for a longer period of time, and finally settles to a small value. The current fluctuations arise as the filament growth process is controlled by thermal hopping at high temperatures. The vacancies can disconnect and connect to the filament, and therefore, the current fluctuates. The high temperature drives the vacancy away from the central region, and finally, the memory cell is RESET to a high resistance state.

Retention of the memory state is an important device metric for RRAM, which we study next using our model. Since the resistance of the filament is determined by the configuration of vacancies, small changes in the vacancy configuration can lead

Authorized licensed use limited to: University of Washington Libraries. Downloaded on November 19,2020 at 20:07:28 UTC from IEEE Xplore. Restrictions apply.

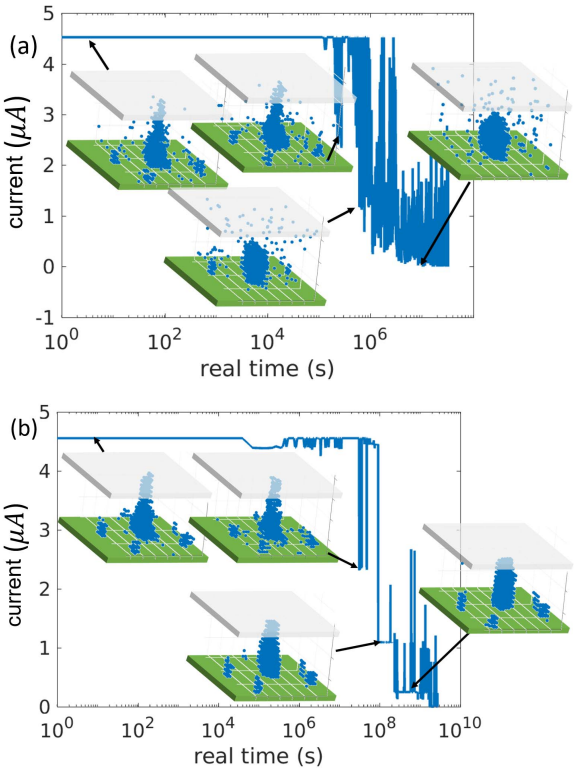


Fig. 6. Retention of filament cell for (a) low and (b) high bonding energy. The ambient temperature is 300 K. The migration barrier is $E_{barr} = 1.1$ eV. Current is computed from a test voltage of 0.3 V after the KMC simulation. For low bond energy (a), $E_{bond} = 0.03$ eV, the vacancies diffuse away from the central region and as a result, the current drops. In the high bond energy (b), $E_{bond} = 0.1$ eV, the vacancies are retained in the central region, and the current is stable for a longer duration.

to significant changes in device resistance. For the retention simulations, a voltage is not applied to the memory cell and the vacancies are neutral. Therefore, the migration barrier is higher than that for a charged vacancy. In the simulation, the migration barrier is taken as $E_{barr} = 1.1$ eV [29]. To calculate the resistance as a function of time, the KMC simulations are interrupted and the current is calculated by applying a small test voltage of 0.3 V. As shown in Fig. 6(a), when $E_{bond} = 0.03$ eV, the low resistance is maintained for about 10^6 s (~10 days) at 300 K. The vacancies diffuse away from the filament region, and the resistance increases (current decreases). The evolution of the filament is shown in Fig. 6(a). We observe that the filament dissolves at the tip, which is the thinnest region. On the other hand, in the case of high bonding energy, $E_{bond} = 0.1$ eV, the filament lasts much longer. As shown in Fig. 6 (b), the low resistance state can last for 10^8s (~3 years). The large bonding energy helps to keep the filament intact as the time scale for vacancy diffusion is much higher. We also see that with time, the thickness of the filament smoothens to minimize the surface free energy.

IV. Conclusion

We have developed a comprehensive 3-D model to simulate the formation, SET, and RESET processes in HfO_2-based resistive memory devices with Ti reservoir layers. Our studies suggest that most vacancies are generated at the metal/oxide interface, as the formation energy of vacancies on the surface is much lower than the formation energy of vacancy–interstitial pairs in the bulk oxide. The charged vacancy can migrate under the electric field and get neutralized near the cathode. As the neutralized vacancy has a higher migration barrier, the conductive filament which consists of these neutralized vacancies can retain the resistance for a long time. Since the generation location is different from the filament location, the thickness of the oxide can affect the filament shape. For sufficiently thick oxides, the filament shape is found to change from conical to a dumbbell with increasing compliance current. The oxygen generated in the forming of the filament can partially oxidize the Ti electrode. This partially oxidized metal forms an insulating cap near the filament tip, which helps to enhance the electric field near the tip. As most of the vacancy generation occurs near the Hafnia–metal interface, this enhancement increases the generation rate of vacancies. Also, the partially oxidized layer changes the electric field from a direction perpendicular to the Hafnia–metal interface to having a large component that points toward the tip of the filament. The change of direction guides the generated vacancy to diffuse toward the filament tip. As a result, the filament can continue thickening after it bridges the two electrodes until current compliance is reached. Therefore, the resistance of the switched cell can be controlled by current compliance. We find that the most dominant process in RESET is recombination at the filament-active electrode interface. When a reverse voltage is applied to the memory cell, the partially oxidized Ti layer acts as an oxygen reservoir. Oxygen atoms can migrate back to the oxide and fill the vacancies. Then, an insulating gap forms in the oxide and the memory cell is switched back to a high resistance state. Our simulation of the retention process suggests that the vacancies diffuse away from the filament region and an insulating gap is formed on top of the filament at sufficiently long time intervals. It is important to note that while the filament may be dumbbell-shaped in the short time scale, in the long time scale, the vacancies redistribute to form a more uniform filament before an insulating gap finally forms. We also find that higher bonding energy helps vacancy clustering which helps in improving the retention time.

Acknowledgment

The authors would like to thank Dr. F. Chen for extensive discussions of experimental results and the importance of a model to capture all three phenomena, set, reset, and retention, J. J. Hsu for the useful discussion, Dr. G. Bersuker for providing critical feedback on results from *ab initio* calculations, and Dr. D. Garbin and Dr. M. Alayan for pointing out the experimental data to test our models.

References

[1] R. Waser and M. Aono, "Nanoionics-based resistive switching memories," *Nature Mater.*, vol. 6, p. 833, Nov. 2007.
[2] T. Hasegawa, K. Terabe, T. Tsuruoka, and M. Aono, "Atomic switch: Atom/ion movement controlled devices for beyond Von-Neumann computers," *Adv. Mater.*, vol. 24, p. 252, Jan. 2012.
[3] T. Tsuruoka, K. Terabe, T. Hasegawa, and M. Aono, "Forming and switching mechanisms of a cation-migration-based oxide resistive memory," *Nanotechnology*, vol. 21, no. 42, 2010, Art. no. 425205.

Authorized licensed use limited to: University of Washington Libraries. Downloaded on November 19,2020 at 20:07:28 UTC from IEEE Xplore. Restrictions apply.

[4] Y. C. Yang, F. Pan, Q. Liu, M. Liu, and F. Zeng, "Fully room-temperature-fabricated nonvolatile resistive memory for ultrafast and high-density memory application," *Nano Lett.*, vol. 9, p. 1636, Mar. 2009.

[5] Z. Wei *et al.*, "Highly reliable TaO_x ReRAM and direct evidence of redox reaction mechanism," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2008, pp. 1–4.

[6] J. Guy *et al.*, "Investigation of forming, SET, and data retention of conductive-bridge random-access memory for stack optimization," *IEEE Trans. Electron Devices*, vol. 62, no. 11, pp. 3482–3489, Nov. 2015.

[7] A. Padovani, L. Larcher, O. Pirrotta, L. Vandelli, and G. Bersuker, "Microscopic modeling of HfO_x RRAM operations: From forming to switching," *IEEE Trans. Electron Devices*, vol. 62, no. 6, pp. 1998–2006, Jun. 2015.

[8] M. Lanza, G. Bersuker, M. Porti, E. Miranda, M. Nafría, and X. Aymerich, "Resistive switching in hafnium dioxide layers: Local phenomenon at grain boundaries," *Appl. Phys. Lett.*, vol. 101, no. 19, 2012, Art. no. 193502.

[9] S. Ambrogio *et al.*, "Neuromorphic learning and recognition with one-transistor-one-resistor synapses and bistable metal oxide RRAM," *IEEE Trans. Electron Devices*, vol. 63, no. 4, pp. 1508–1515, Apr. 2016.

[10] G. Bersuker *et al.*, "Metal oxide resistive memory switching mechanism based on conductive filament properties," *J. Appl. Phys.*, vol. 110, no. 12, 2011, Art. no. 124518.

[11] M. Lanza *et al.*, "Recommended methods to study resistive switching devices," *Adv. Electron. Mater.*, vol. 5, Jan. 2019, Art. no. 1800143.

[12] J. M. Coll *et al.*, "Towards oxide electronics: A roadmap," *Appl. Surf. Sci.*, vol. 482, p. 1–93, Jul. 2019.

[13] T. Tan, Y. Du, A. Cao, Y. Sun, H. Zhang, and G. Zha, "Resistive switching of the HfO_x/HfO_2 bilayer heterostructure and its transmission characteristics as a synapse," *Roy. Soc. Chem. Adv.*, vol. 8, no. 73, pp. 41884–41891, 2018.

[14] Y. Bai, H. Wu, R. Wu, Y. Zhang, N. Deng, Z. Yu, and H. Qian, "Study of multi-level characteristics for 3D vertical resistive switching memory," *Sci. Rep.*, vol. 4, p. 5780, Jul. 2014.

[15] H. W. Pan *et al.*, "1 K bit FinFET dielectric (FIND) RRAM in pure 16 nm FinFET CMOS logic process," in *IEDM Tech. Dig.*, Washington, DC, USA, Dec. 2015, pp. 10.5.1–10.5.4.

[16] D. Ielmini, "Modeling the universal set/reset characteristics of bipolar RRAM by field-and temperature-driven filament growth," *IEEE Trans. Electron Devices*, vol. 58, no. 12, pp. 4309–4317, Dec. 2011.

[17] W. He, H. Sun, Y. Zhou, K. Lu, K. Xue, and X. Miao, "Customized binary and multi-level HfO_{2-x}-based memristors tuned by oxidation conditions," *Sci. Rep.*, vol. 7, Aug. 2017, Art. no. 10070.

[18] A. K. Singh, S. Blonkowski, and M. Kogelschatz, "Resistive switching study in HfO_2 based resistive memories by conductive atomic force microscopy in vacuum," *J. Appl. Phys.*, vol. 124, no. 1, 2018, Art. no. 014501.

[19] S. Aldana *et al.*, "A 3D kinetic Monte Carlo simulation study of resistive switching processes in Ni/HfO_2/Si-n^+-based RRAMs," *J. Phys. D, Appl. Phys.*, vol. 50, no. 33, 2017, Art. no. 335103.

[20] H. Y. Lee *et al.*, "Low power and high speed bipolar switching with a thin reactive Ti buffer layer in robust HfO_2 based RRAM," in *IEDM Tech. Dig.*, Dec. 2008, pp. 1–4.

[21] J. Park *et al.*, "Resistive switching characteristics of ultra-thin TiO_x," *Microelectron. Eng.*, vol. 88, no. 7, p. 1136, 2011.

[22] D. Panda, P. P. Sahu, and T. Y. Tseng, "A collective study on modeling and simulation of resistive random access memory," *Nanoscale Res. Lett.*, vol. 13, Jan. 2018, Art. no. 8.

[23] S. Menzel, P. Kaupmann, and R. Waser, "Understanding filamentary growth in electrochemical metallization memory cells using kinetic Monte Carlo simulations," *Nanoscale*, vol. 7, pp. 12673–12681, Jun. 2015.

[24] P. Huang *et al.*, "A physical based analytic model of RRAM operation for circuit simulation," in *IEDM Tech. Dig.*, Dec. 2012, pp. 26.6.1–26.6.4.

[25] X. Guan, S. Yu, and H.-S. P. Wong, "On the switching parameter variation of metal-oxide RRAM—Part I: Physical modeling and simulation methodology," *IEEE Trans. Electron Devices*, vol. 59, no. 4, p. 1172, Apr. 2012.

[26] S. Yu, X. Guan, and H.-S. P. Wong, "On the switching parameter variation of metal oxide RRAM—Part II: Model corroboration and device design strategy," *IEEE Trans. Electron Devices*, vol. 59, no. 4, pp. 1183–1188, Apr. 2012.

[27] D. Berco and T.-Y. Tseng, "A comprehensive study of bipolar operation in resistive switching memory devices," *J. Comput. Electron.*, vol. 15, no. 2, p. 577, 2016.

[28] B. Traoré *et al.*, "Impact of electrode nature on the filament formation and variability in HfO_2 RRAM," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Waikoloa, HI, USA, Jun. 2014, pp. 5E.2.1–5E.2.5.

[29] S. M. Aspera, H. Kasai, H. Kishi, N. Awaya, S. Ohnishi, and Y. Tamai, "Realization of the switching mechanism in resistance random access memory devices: Structural and electronic properties affecting electron conductivity in a hafnium oxide–electrode system through first-principles calculations," *J. Electron. Mater.*, vol. 42, no. 1, p. 143, 2013.

[30] B. Traore, P. Blaise, B. Sklénard, E. Vianello, B. Magyari-Köpe, and Y. Nishi, "HfO_2/Ti interface mediated conductive filament formation in RRAM: An *ab initio* study," *IEEE Trans. Electron Devices*, vol. 65, no. 2, pp. 507–513, Feb. 2017.

[31] S. R. Bradley, A. L. Shluger, and G. Bersuker, "Electron-injection-assisted generation of oxygen vacancies in monoclinic HfO_2," *Phys. Rev. Appl.*, vol. 4, Dec. 2015, Art. no. 064008.

[32] J. McPherson, J.-Y. Kim, A. Shanware, and H. Mogul, "Thermochemical description of dielectric breakdown in high dielectric constant materials," *Appl. Phys. Lett.*, vol. 82, no. 13, p. 2121, 2003.

[33] D. M. Pozar, *Microwave Engineering*, 3rd ed. Hoboken, NJ, USA: Wiley, 2005, p. 10.

[34] D. Garbin *et al.*, "HfO_2-based O_xRAM devices as synapses for convolutional neural networks," *IEEE Trans. Electron Devices*, vol. 62, no. 8, pp. 2494–2501, Aug. 2015.

[35] M. Alayan, "Investigation of HfO_2 based resistive random access memory (RRAM): Characterization and modeling of cell reliability and novel access device," Ph.D. dissertation, Dept. Micro Nanotechnol./Microelectron., Univ. Grenoble Alpes, Grenoble, France, 2018.

[36] P. Gonon *et al.*, "Resistance switching in HfO_2 metal-insulator-metal devices," *J. Appl. Phys.*, vol. 107, no. 7, 2010, Art. no. 074507.

[37] H. L. Barros, G. V. Chandrashekhar, T. C. Chi, J. M. Honig, and R. J. Sladek, "Specific heat of single-crystal undoped and V-doped Ti_2O_3," *Phys. Rev. B, Condens. Matter*, vol. 7, p. 5147, Jun. 1973.

[38] Y. Guo, S. J. Clark, and J. Robertson, "Electronic and magnetic properties of Ti_2O_3, Cr_2O_3, and Fe_2O_3 calculated by the screened exchange hybrid density functional," *J. Phys., Condens. Matter*, vol. 24, no. 32, 2012, Art. no. 325504.

[39] C. Tang and R. Ramprasad, "Point defect chemistry in amorphous HfO_2: Density functional theory calculations," *Phys. Rev. B, Condens. Matter*, vol. 81, Apr. 2010, Art. no. 161201.

[40] S. R. Bradley, G. Bersuker, and A. L. Shluger, "Modelling of oxygen vacancy aggregates in monoclinic HfO_2: Can they contribute to conductive filament formation?" *J. Phys., Condens. Matter*, vol. 27, no. 41, 2015, Art. no. 415401.

[41] A. O'Hara, G. Bersuker, and A. A. Demkov, "Assessing hafnium on hafnia as an oxygen getter," *J. Appl. Phys*, vol. 115, no. 18, 2014, Art. no. 183703.

[42] D. Lee, Y. Sung, I. Lee, J. Kim, H. Sohn, and D.-H. Ko, "Enhanced bipolar resistive switching of HfO_2 with a Ti interlayer," *Appl. Phys. A, Solids Surf.*, vol. 102, no. 4, p. 997, 2011.

[43] D. S. Jeong, H. Schroeder, and R. Waser, "Coexistence of bipolar and unipolar resistive switching behaviors in a Pt/TiO_2/Pt stack," *Electrochem. Solid-State Lett.*, vol. 10, no. 8, p. G51, 2007.

[44] E. Linn, R. Rosezin, C. Kügeler, and R. Waser, "Complementary resistive switches for passive nanocrossbar memories," *Nature Mater.*, vol. 9, p. 403, Apr. 2010.

[45] G. Niu *et al.*, "Geometric conductive filament confinement by nanotips for resistive switching of HfO_2-RRAM devices with high performance," *Sci. Rep.*, vol. 6, May 2016, Art. no. 25757.

[46] L. Goux *et al.*, "On the gradual unipolar and bipolar resistive switching of TiN\HfO_2\Pt memory systems," *Electrochem. Solid-State Lett.*, vol. 13, no. 6, p. G54, 2010.

[47] D. Duncan, B. Magyari-Köpe, and Y. Nishi, "Filament-induced anisotropic oxygen vacancy diffusion and charge trapping effects in hafnium oxide RRAM," *IEEE Electron Device Lett.*, vol. 37, no. 4, pp. 400–403, Apr. 2016.

[48] J. Y. Shin and Y.-H. Son, "Direct observation of conducting filaments on resistive switching of NiO thin films," *Appl. Phys. Lett.*, vol. 92, no. 22, 2008, Art. no. 222106.

[49] S. Larentis, C. Cagli, F. Nardi, and D. Ielmini, "Filament diffusion model for simulating reset and retention processes in RRAM," *Microelectron. Eng.*, vol. 88, p. 1119, Jul. 2011.

[50] F. T. Chen *et al.*, "Resistance instabilities in a filament-based resistive memory," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Apr. 2013, pp. 5E.1.1–5E.1.7.

Authorized licensed use limited to: University of Washington Libraries. Downloaded on November 19,2020 at 20:07:28 UTC from IEEE Xplore. Restrictions apply.