# Effect of Electrode Roughness on Electroforming in $HfO_2$ and Defect-Induced Moderation of Electric-Field Enhancement

Sanjoy Kumar Nandi,[1,2,3,*] Xinjun Liu,[1] Dinesh Kumar Venkatachalam,[1] and Robert Glen Elliman[1]
[1]*Department of Electronic Materials Engineering, The Australian National University, Canberra, Australian Capital Territory 0200, Australia*
[2]*Research School of Astronomy and Astrophysics, The Australian National University, Canberra, Australian Capital Territory 0200, Australia*
[3]*Department of Physics, University of Chittagong, Chittagong 4331, Bangladesh*

(Received 5 July 2015; revised manuscript received 10 September 2015; published 29 December 2015)

The roughness of Pt electrodes is shown to have a direct impact on the electroforming characteristics of $Pt/Ti/HfO_2/Pt$ device structures. Specifically, an increase in roughness leads to a reduction in the electroforming voltage of $HfO_2$, an increase in the failure rate of devices, and a corresponding reduction in resistive switching reliability. A finite-element model is used to investigate the significance of local electric-field enhancement on the breakdown process. This simulation shows that high-aspect-ratio asperities can produce field enhancements of more than an order of magnitude but that the generation and redistribution of defects moderate this effect prior to dielectric breakdown. As a consequence, the effect of field enhancement is less than anticipated from the initial electric-field distribution alone. Finally, it is argued that the increase in the device failure rate with increasing electrode roughness derives partly from an increase in the film defect density and effective device area and that these effects contribute to the reduction in breakdown voltage.

DOI: 10.1103/PhysRevApplied.4.064010

## I. INTRODUCTION

Dielectric breakdown has been extensively studied in a wide range of metal-oxide-semiconductor [1,2] and metal-insulator-metal (*MIM*) structures [3–10]. More recently, it has received attention in relation to resistive random-access memory devices (RRAM), where an electroforming step is usually required to initiate resistive switching [11–15]. This step is analogous to soft dielectric breakdown and results in a filamentary conductive path through the dielectric. According to the electrochemical model [16,17] of dielectric breakdown, defects are produced by a field-enhanced activation process in which the applied electric field acts to distort polar metal-oxygen bonds and thereby reduce the activation energy for bond breaking. Conduction initially occurs along a filamentary percolation path. This increases the local temperature via Joule heating and further accelerates defect formation to provide the positive feedback that leads to dielectric breakdown [18–20]. Irreversible hard breakdown is generally prevented by limiting the maximum current in the device [21]. Conduction is then limited to a filamentary path comprised of a network of defects.

In its simplest form, the electrochemical model of breakdown considers a dielectric film subjected to a uniform electric field and predicts the random and homogeneous generation of defects within the dielectric layer [16,17]. In this model, the concentration of defects gradually increases to a percolation threshold for continuous conduction at which point there is a sudden increase in current. In reality, breakdown is usually a heterogeneous process dominated by intrinsic defects in the dielectric film and defect generation at the metal or dielectric interface, a process that is accelerated by electric-field enhancement at electrode asperities [22–26]. For example, rough electrodes are known to increase the leakage current [27] and reduce the yield of thin-film oxide capacitors [28–31]. The physical understanding of these processes is based on models of the electric-field distribution at the metal-dielectric interface. However, these models generally assume a static field distribution and ignore the effect of defect generation and migration (drift and diffusion) on the local electric-field distribution.

In this study, we investigate the forming process in $Pt/Ti/HfO_2/Pt$ structures as a function of bottom electrode roughness and introduce a finite-element model of dielectric breakdown that includes the effect of defect generation and migration on the local electric field. Although the model is somewhat idealized, it serves to demonstrate the important role of defects in moderating the effects of geometric field enhancement.

## II. EXPERIMENTS

Simple *MIM* test structures, as shown in Fig. 1(a), are fabricated on thermally oxidized (100) silicon wafers with an oxide thickness of 300 nm. The bottom electrodes comprise a 10-nm Ti wetting layer and a Pt layer with a

---

[*] Sanjoy.Nandi@anu.edu.au

thickness in the range of 25–150 nm. These films are deposited by electron-beam evaporation without breaking vacuum and deposited at an average rate of 2 Å/s for both Ti and Pt. The film thicknesses are determined by a crystal oscillator during deposition and subsequently confirmed by cross-sectional transmission electron microscopy (TEM). A 20-nm-thick $HfO_2$ layer is deposited on the Pt bottom electrode by atomic-layer deposition (ALD). This is achieved by heating the sample to 200 °C and exposing it to alternate pulses of tetrakis-(dymethylamido)-hafnium [TEMAH, $Hf(NMe_2)_4$] and $H_2O$ vapor with an $N_2$ purge between pulses. Top electrodes of 150-$\mu$m diameter comprising 10-nm Ti and 50-nm Pt are then deposited through a shadow mask by electron-beam evaporation to complete the asymmetric *MIM* test structures.

As-deposited $HfO_2$ films are analyzed by grazing-incident-angle x-ray diffraction and TEM electron diffraction (not shown) and found to be amorphous. The stoichiometry of the films is confirmed by electron Rutherford backscattering and is reported elsewhere [32,33]. The roughness of the bottom Pt electrode and the deposited $HfO_2$ film are measured using atomic-force microscopy (AFM) in tapping mode. AFM scans are taken over $1 \times 1\ \mu m^2$, $2 \times 2\ \mu m^2$ and $5 \times 5\ \mu m^2$ areas, and the roughness is calculated for each scan.

Current-voltage ($I$-$V$) characteristics of the devices are measured using an Agilent B1500A semiconductor parametric analyzer attached to a Signatone probe station. Forming measurements are performed by measuring the current between the top and bottom electrodes while scanning the applied voltage from 0 to 10 V. The bottom electrode is grounded, and positive voltage is applied to the top electrode. An instrumental compliance of 10 $\mu$A is applied to avoid permanent damage to the test device. Note that the test devices employ different top and bottom electrode materials, so their $I$-$V$ characteristics are asymmetric under positive and negative bias due to the different Schottky barrier heights (see Appendix A). However, the effect of the electrode roughness is the same in both cases.

## III. RESULTS AND DISCUSSION

### A. Surface roughness

A TEM image of a 150-nm-thick Pt bottom-contact layer is shown in Fig. 1(b). The film is observed to be polycrystalline with individual grains having lateral dimensions in the range of 50–100 nm. The roughness of the bottom electrode is clearly evident, but the details are difficult to interpret due to the finite thickness of the sample cross section. More quantitative analysis is provided by AFM imaging of the surface, as shown in Figs. 2(a)–2(d). These compare images from the thinnest (25 nm) and thickest (150 nm) Pt electrodes prior to [Figs. 2(a) and 2(b)] and following [Figs. 2(c) and 2(d)]$HfO_2$ deposition and clearly show an increase in film roughness with increasing film thickness. This dependence is shown more explicitly in Fig. 3(a), which plots the root-mean-square (rms) roughness of the bottom electrode before and after $HfO_2$ deposition as a function of the Pt film thickness. The data show that the rms roughness varies by a factor of 6–7 between the thinnest ($R_{rms} = 0.44$ nm) and thickest ($R_{rms} = 2.62$ nm) films. The dependence of roughness on film thickness is typically described by a power-law equation of the form $R_{rms} \sim t^{\delta}$, where $R_{rms}$ is the root-mean-square roughness, $t$ is the film thickness, and $\delta$ is a coefficient that depends on the mode of film growth. In the current case, $\delta$ is in the range 0.7–1.0, which is consistent with the roughness origination from facetted grain growth [34]. By way of comparison, the $SiO_2$ substrate is found to have a rms roughness of 0.16 nm.

Because dielectric breakdown is a weakest-link process it is expected to depend on extreme features rather than the average or rms roughness of the electrode. Consequently, it is important to consider the full surface-height distribution of the electrodes. Figure 3(b) shows the measured height distribution for each film, where the height is measured relative to the lowest height in the scanned area. It is clear that the thickest bottom electrode has the largest spread in

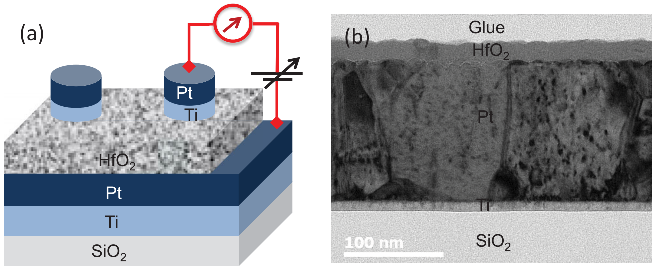


FIG. 1. (a) Schematic of the device structure and (b) XTEM of a 150-nm Pt with 20-nm $HfO_2$ showing both bottom and top interfaces.

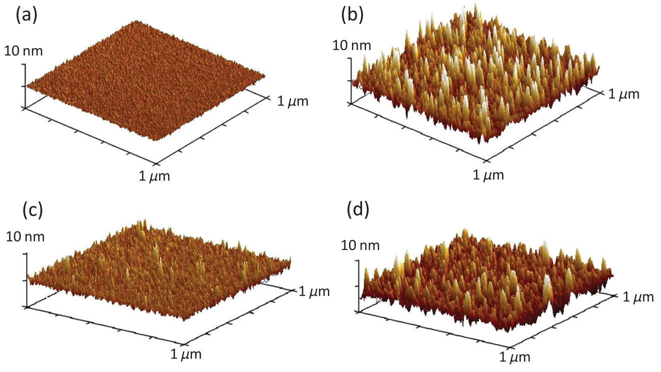
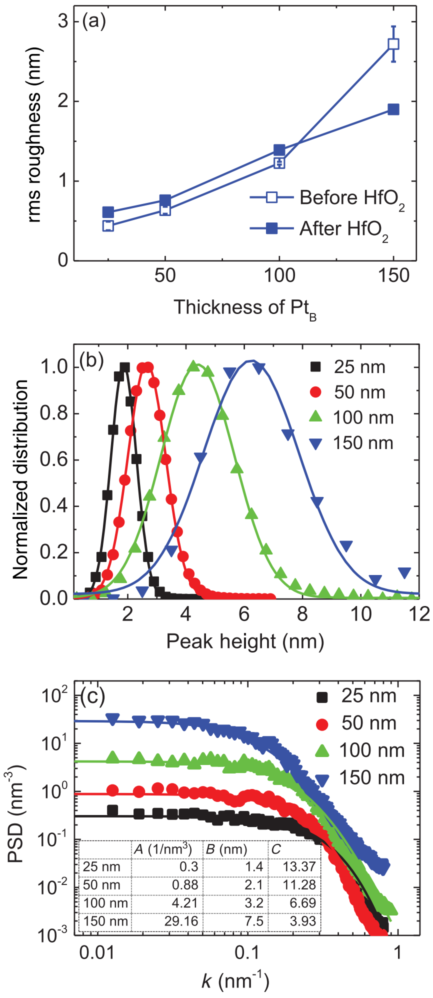


FIG. 2. 3D topography AFM images of (a) 25-nm and (b) 150-nm as-grown Pt films. (c) 25-nm and (d) 150-nm Pt after 20-nm $HfO_2$ deposition, which correspond to the top interface. (Note that: AFM measurements can underestimate the aspect ratio of features due to convolution with the measurement probe.)



| | $A$ (1/nm$^3$) | $B$ (nm) | $C$ |
|---|---|---|---|
| 25 nm | 0.3 | 1.4 | 13.37 |
| 50 nm | 0.88 | 2.1 | 11.28 |
| 100 nm | 4.21 | 3.2 | 6.69 |
| 150 nm | 29.16 | 7.5 | 3.93 |

FIG. 3. (a) rms roughness of bottom Pt electrode before and after HfO$_2$ deposition as a function of film thickness. (b) Height distribution of the surface roughness for different thickness of Pt and (c) PSD profiles of the same Pt layers calculated from AFM measurements. The fit to the *ABC* model is shown by solid lines. Inset represents fitting coefficients of PSD spectra.

peak heights. Indeed, the maximum peak-to-valley distance measured for the 150-nm Pt electrode is 25 nm, whereas that for the 25-nm Pt film is only 5 nm.

Finally, we consider the lateral scale of the surface topography. This is achieved by considering the power spectral density (PSD) of the surface-height distributions, which decomposes the surface profile into its spatial wavelengths [35]. For a condensed description of the Pt films, the spectra are fitted using the $k$-correlation model (also called the *ABC* model) assuming perfectly isotropic surfaces [36,37]. This model allows a quantitative comparison between samples over large length scales.

According to this model, PSD is defined as

$$\mathrm{PSD} = \frac{A}{(1 + B^2 k^2)^{C/2}}, \qquad (1)$$

where $A$, $B$, and $C$ are model parameters, and $k$ is the web vector. The amplitude of the PSD is a measure of the roughness component with particular spatial frequency, and the distributions show two regions with different slopes: a relatively flat region at low spatial frequencies and a steeper region at higher spatial frequencies. The parameter $A$ describes the low-frequency ($k \ll B$) limit of the spectrum, where the PSD is constant and equal to $A$. In this region, there is no significant variation of height in real space. $B$ determines the knee of the PSD and defines a correlation length beyond which the surface-height fluctuations are uncorrelated. Beyond this length, the surface is fractal, and the PSD is determined by $1/k^C$. Such analysis is shown in Fig. 3(c) for the bottom Pt electrodes, with model parameters shown in the inset. The results confirm the correlation between surface topography and film thickness. For the thinnest layer, the correlation length $B$ is approximately 1.5 nm, which suggests that the roughness is dominated by lateral features with dimensions less than this value. In contrast, for the 150-nm film, $B$ is approximately 8 nm. A comparison of these curves suggests that the roughness of the thicker films is dominated by large features, consistent with an increase in the Pt grain size with increasing film thickness [38,39].

### B. Dielectric breakdown

A series of dielectric-breakdown measurements is performed on each sample and representative results are shown in Fig. 4. Figure 4(a) shows the cumulative probability distributions depicting the fraction of devices that undergo breakdown at a given voltage, and Fig. 4(b) plots the mean breakdown voltage as a function of rms roughness, highlighting a reduction in voltage with increasing electrode thickness.

An analysis of these distributions is undertaken assuming a Weibull distribution [40] which has a cumulative distribution function $P(V_f)$ defined as

$$P(V_f) = 1 - \exp\left[-\left(\frac{V_f}{\alpha}\right)^{\beta}\right], \qquad (2)$$

where $\alpha$ is the scale parameter, which is closely related to the mean breakdown voltage, and $\beta$ is the shape parameter, which is a measurement of the width of the distribution [40]. A small value of $\beta$ indicates widely dispersed data; i.e., the data tend to distribute over a relatively wide range of breakdown voltages, whereas for large values of $\beta$, the majority of breakdown tends to fall around the mean voltage. Thus, the reliability of devices can be represented by the Weibull parameters. The solid lines in Fig. 4(a) represent least squares fits of Eq. (2), and the resulting values of $\alpha$ and $\beta$ are included in the legend. A comparison of the various distributions and fitting parameters confirms the reduction in the mean breakdown voltage and a significant increase in the dispersion of the breakdown

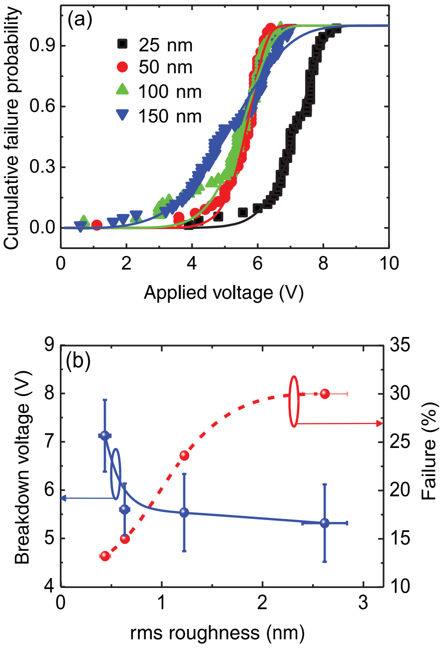


FIG. 4. (a) Weibull distribution of forming voltage of Pt/Ti/$HfO_2$/Pt as a function of bottom electrode thickness. (b) Variation of breakdown voltage and failure with rms roughness.

voltage with increasing bottom electrode thickness (roughness).

The effects of electrode roughness on the breakdown voltage of *MIM* devices, such as that shown in Fig. 4(b)), can arise from roughness-induced variations in the thickness [41,42], variations in the structure and quality of the deposited dielectric film [43], an increase in the effective surface area of the device [44], or local electric-field enhancement in the vicinity of high-aspect-ratio asperities [45]. The first of these effects is particularly evident for films deposited from directional sources, such as evaporation or sputter deposition, where shadowing by surface features is significant [41,42]. The shadowing effect is not a significant problem for films deposited by ALD, as these grow conformally on surfaces with complex morphologies, including nanowires, nanoparticles, and high-aspect-ratio trenches. However, an increase in the density of intrinsic defects [46–51] with increasing substrate roughness can still be expected for ALD films and is consistent with the increase in the device failure rate with increasing roughness shown in Fig. 4(b)). Here the failure rate is the fraction of test devices found to be in a low-resistance state prior to testing (i.e., device yield). These data show that the average forming voltage $V_f$ and device yield both decrease with increasing mean surface roughness.

The reliability and endurance of resistive switching in these structures is also investigated as a function of roughness and is discussed in Appendix B. These measurements show that the reliability (i.e., variations in the set or reset voltage and resistance during repetitive dc switching cycles) and endurance are significantly worse for rough electrodes than for smooth electrodes, despite the reduction in forming voltage. This is attributed to an increased probability for random filament generation during the set process and is consistent with the conclusions of related studies [52,53].

### C. Simulations

To gain insight into the significance of field-enhancement effects, a finite-element model is developed to simulate local field enhancement and defect dynamics, including the role of defect generation and migration on the electric-field distribution. The model is a relatively simple representation of the physical system and does not include detailed chemical (e.g., reduction of $HfO_2$ by Ti) or electrical effects (i.e., Schottky barrier heights, depletion layers, etc.). However, it does aim to capture the essential physics of the electroforming process, and because defects are generated in the vicinity of electrodes (due to the local field enhancement in this region), it is consistent with the current understanding of oxygen-vacancy generation at the metal-insulator interface. Specifically, it comprises a 2D axisymmetric model of a 20-nm $HfO_2$ layer sandwiched between two Pt electrodes, one of which has a Gaussian-shaped asperity. The model assumes that (i) charged defects are generated at a rate given by Eq. (C1), (ii) that these defects can drift and diffuse in the dielectric film, and (iii) that the presence of the defects modifies the local conductivity of the dielectric.

The simulation, therefore, involves the self-consistent solution of the drift-diffusion equation for defect transport with a source term given by $G$ [Eq. (C1)], the continuity equation for electrical conduction [Eq. (C5)], and the Fourier heat equation [Eq. (C6)] including Joule heating (Appendix C). The defects of interest are assumed to be oxygen vacancies in the plus two-charge state (i.e., $V_O{}^{2+}$). No account is taken of the oxygen ions produced by the dissociation reaction, as these do not contribute to electronic conduction. The electrical and thermal conductivity of $HfO_2$ are assumed to be modified by the defect density according to the resistive switching model of Larentis *et al.* [54] (Appendix C).

The effect of surface roughness is modeled by adding a simple Gaussian-shaped asperity to the bottom electrode, centered on the symmetry axis [Fig. 5(a)]. This shape helps to avoid singularities caused by sharp edges and gives numerical simplicity. The height and width of the asperity are varied to study the effects of the asperity aspect ratio on breakdown. Note, however, that the model differs from the experimental situation in that the upper electrode is assumed to be flat rather than reflect the conformal growth of the dielectric layer on the bottom electrode. Nonetheless, the simulation captures the essential features of breakdown process as is evident from Fig. 5(b). This figure compares a calculated breakdown characteristic with experimental data

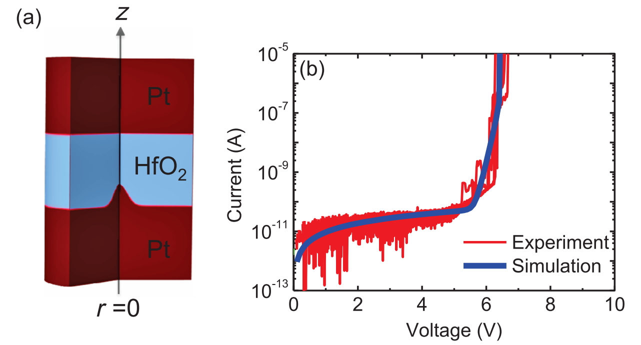


FIG. 5. (a) Schematic of the simulated device structure with a Gaussian asperity at the bottom interface. (b) Comparison of experimental and simulated current-voltage characteristics obtained by applying ramp bias.

and clearly reproduces the most important and obvious features of the measurements. In both cases, the current is observed to increase slowly with increasing voltage before rising rapidly at the point of breakdown, which represents a soft breakdown process prior to the irreversible thermal runaway characteristic of hard breakdown.

Since electrode topography alters the local electric field, it will also alter the defect generation rate and thereby modify the breakdown behavior. This is illustrated in Fig. 6(a), which shows the initial electric-field distribution in the vicinity of a Gaussian asperity of height 2 nm and FWHM (hereafter, referred as *width*) of approximately 20 nm. Both the average electric field and the peak field increase in the vicinity of the asperity, as expected. Consequently, as the applied voltage increases, defects are preferentially generated in the vicinity of the asperity [Fig. 6(b)]. Once generated, they drift towards the anode to create a filamentary conduction path through the dielectric layer [Fig. 6(c)]. The presence of the defects modifies the conductivity of the dielectric layer and, hence, the local electric-field distribution, a fact that must be included in the simulation. But as the conductive filament forms, it confines the electric current and causes an increase in the local temperature due to Joule heating [Fig. 6(d)] and soft dielectric breakdown, as shown previously in Fig. 5.

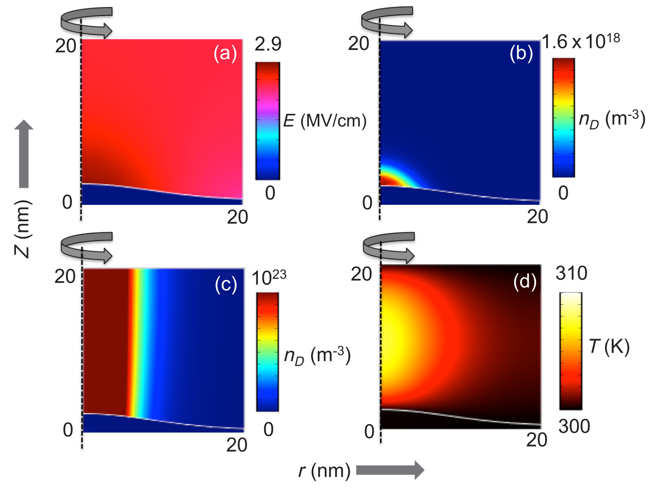


FIG. 6. (a) 2D plot of electric field ($E = \nabla\psi$) and (b) defect density $n_D$ at critical field with an asperity height of 2 nm. Both electric field and defect density are concentrated on the top of asperity. (c) Defect density $n_D$ and (d) temperature $T$ distributions at forming voltage in the same device structure.

Figure 7(a) summarizes the effect of the asperity shape on the forming voltage by plotting the calculated voltage as a function of the peak height and width of the bottom electrode asperity. An increase in peak height is shown to reduce the forming voltage, consistent with the fact that it acts to reduce the local film thickness and increases the local electric field. For a fixed asperity width of

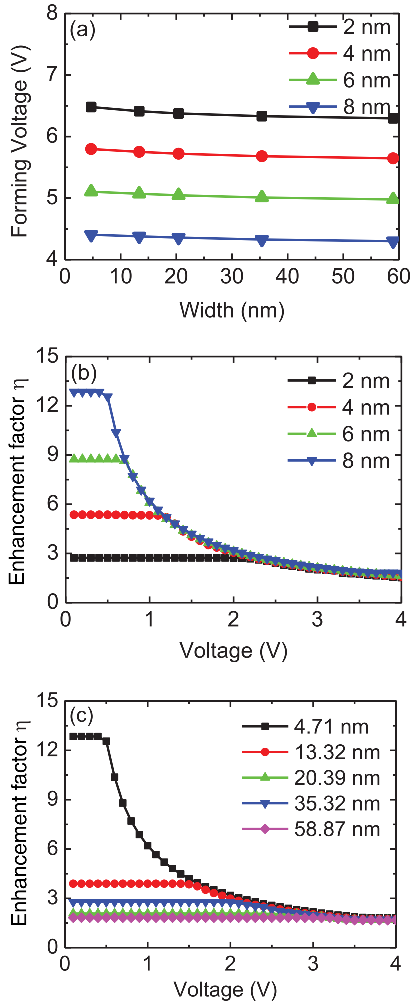


FIG. 7. (a) Forming voltage as a function of height and width of the asperity. (b) Electric-field-enhancement factor as a function of applied bias for different heights of the asperity but fixed width (4.7 nm). (c) The field enhancement as a function bias for different widths but fixed height (8 nm).

approximately 4.6 nm, the forming voltage is shown to decrease from approximately 6.5 V to approximately 4.4 V as the asperity height increases from 2 to 8 nm, a reduction that is of similar order to that observed experimentally. In contrast, changes in the width of the asperity have only a weak effect on the forming voltage, reducing it by just a few percent for width variations of more than an order of magnitude. This is a much smaller effect than expected from the calculated field enhancement.

As an approximation, the local electric field at the asperity $E_{\mathrm{loc}}$ can be expressed as $E_{\mathrm{loc}} = \eta E_{\mathrm{app}}$, where $\eta$ is a field-enhancement factor, and $E_{\mathrm{app}} = V_{\mathrm{app}}/d$ is the average electric field across the dielectric in the absence of any asperities, with $V_{\mathrm{app}}$ the applied voltage and $d$ the thickness of the layer. In general, $\eta$ has a complicated dependence on the height, width, and aspect ratio of an asperity (a simple model is described in Appendix D) [22]. This is reflected in Fig. 7(b), which shows the peak-enhancement factor as a function of applied bias for asperities of different height and constant width (approximately 4.7 nm). The tallest asperity (8 nm high) produces an initial field that is $>12$ times that of a flat electrode, partly due to the effective reduction in film thickness (a factor of 1.7) but mostly due to the height and shape of the asperity. The electric field enhancement factor initially remains constant and is determined by the electrode geometry only. However, the field enhancement begins to decrease for voltages above approximately 0.5 V, which corresponds to the critical electric field (approximately 2.9 MV/cm) for enhanced-defect generation. A similar effect is observed for asperities of different height, with the reduction in field enhancement occurring at different voltages but at the same peak field of approximately 2.9 MV/cm. Clearly, the generation and redistribution of defects acts to reduce the local field around the asperity (Appendix E), and at high voltages the field enhancement approaches a value similar to that calculated from the local film thickness (i.e., between 1.1 and 1.7 for peaks of heights of 2 nm and 8 nm, respectively). Significantly, this reduction in field enhancement occurs at voltages much less than the breakdown voltage. Figure 7 shows a similar effect as a function of asperity width for a constant asperity height of 8 nm. In this case, however, the saturation value is similar for all widths, consistent with the weak dependence of the breakdown voltage on the asperity width reported in Fig. 7(a). From these results, we conclude that the electric-field enhancement around electrode asperities can be reduced by the generation and migration (drift and diffusion) of defects and that as a consequence, the effect on dielectric breakdown is less than anticipated from geometric affects alone.

It is also worth noting that the diameter of conductive filaments increases with the increasing width of the asperity (Appendix F) and that this can have a direct impact on the dynamic evolution of the temperature and current in the filament. Indeed, for very sharp (small-width) asperities, the filament resistance can limit the current and thereby reduce the local temperature increase. As defect generation and migration are thermally activated, this effect reduces the rate of filament evolution and can act to increase the forming voltage. Such effects are responsible for the small increase in forming voltage with decreasing asperity width (increasing sharpness) evident in Fig. 7(a).

## IV. CONCLUSION

The surface roughness of electron-beam-evaporated Pt electrodes is shown to have a strong dependence on film thickness, consistent with the differential growth rate of Pt grains. The effect of electrode roughness on the forming characteristics of Pt/Ti/$\mathrm{HfO_2}$/Pt metal-insulator-metal RRAM structures is also investigated, and the electroforming voltage of $\mathrm{HfO_2}$ films is found to decrease with increasing electrode roughness. However, any advantage afforded by this effect is compromised by a concomitant increase in the forming voltage dispersion and an increase in the failure rate of the devices. Finite-element simulations of dielectric breakdown around an asperity provide considerable insight into the role of local electric-field enhancement. The electrode geometry was found to have a significant effect on the electric-field distribution, with field enhancements of more than an order of magnitude observed for high-aspect-ratio asperities. However, the peak field is shown to be moderated by the generation and redistribution of defects prior to dielectric breakdown. As a consequence, the effect of field enhancement on dielectric breakdown is less than anticipated from geometric affects alone. Finally, it is suggested that the increase in device failure rate with increasing electrode roughness is consistent with an increase in film defect density and effective device area and that these effects contribute to the reduction in breakdown voltage.

## ACKNOWLEDGMENTS

This work is made possible by funding from the Australian Research Council. S. K. N. wishes to acknowledge the Research School of Astronomy and Astrophysics for Ph.D. scholarships. The authors also acknowledge the NCRIS Australian Nanotechnology Fabrication Facility, ACT Node, for access to their research facilities and expertise.

## APPENDIX A: ASYMMETRIC DIELECTRIC BREAKDOWN CHARACTERISTICS OF Pt/Ti/$\mathrm{HfO_2}$/Pt

The asymmetric *I-V* characteristics of Pt/Ti/$\mathrm{HfO_2}$/Pt are shown in Fig. 8.

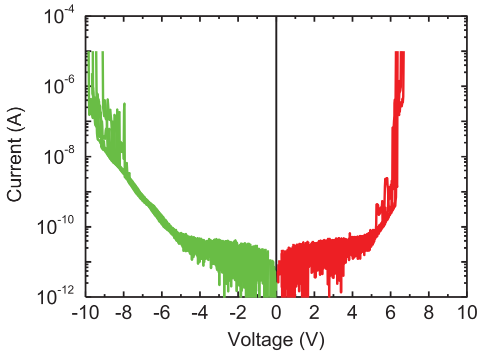


FIG. 8. Forming characteristics of the Pt/Ti/$HfO_2$/Pt device with both polarities.

## APPENDIX B: dc SWITCHING CHARACTERISTICS

The dc resistive switching characteristics of devices with the smoothest and roughest electrodes are compared in Fig. 9. It is evident that a wider distribution in switching parameters is observed for the rougher electrode. Indeed, in general, the device reliability and endurance are found to decrease with increasing electrode roughness. This is attributed to an increased probability for random filament generation during the set process. However, such behavior is also expected to depend on the device area, as both forming and the probability for random filament generation will depend on the number of dominant asperities within the active device area. One possible implication of roughness is that a smaller cell with a single dominant asperity (patterned electrode) may result in better uniformity and reliability.

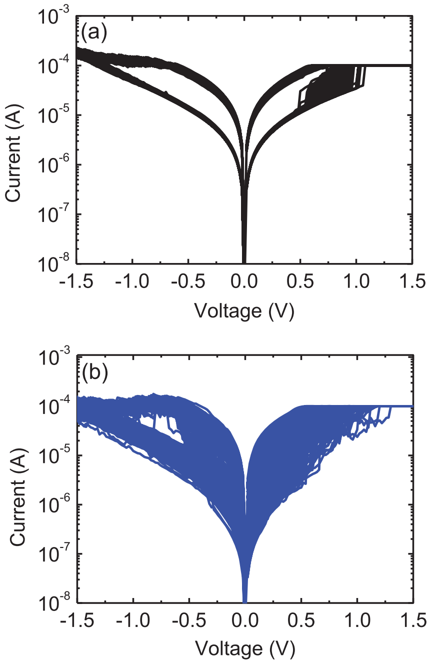


FIG. 9. dc switching characteristics of the Pt/Ti/$HfO_2$/Pt device with (a) 25-nm Pt and (b) 150-nm Pt bottom electrode.

## APPENDIX C: MODELING ASSUMPTIONS

According to the thermochemical model [17], the local electric field distorts molecular bonds and reduces the activation energy required for defect generation. These defects are assumed to be charged (e.g., positively charged vacancies ${V_O}^{2+}$) and to migrate by drift and diffusion driven by the local electric field and Joule heating. Their presence also increases the local conductivity of the dielectric ($HfO_2$) in the present case.

Defect generation, drift, and diffusion are described by the transport equation:

$$\frac{\partial n_D}{\partial t} = \nabla \cdot (\nabla n_D - \nu n_D) + G, \tag{C1}$$

where $D$ is defect diffusivity, $\mu$ is the defect mobility, $v$ the defect drift velocity, and $G$ the defect generation rate. The diffusion coefficient is assumed to be independent of defect concentration and its temperature dependence to be described by an Arrhenius relation:

$$D = D_0 \exp(-E_A/k_B T), \tag{C2}$$

where $D_0$ is a preexponential factor, $E_A$ is the defect migration energy, $k_B$ is the Boltzmann constant, and $T$ is the temperature. The ion mobility is assumed to be related to diffusivity through the Einstein relation:

$$\mu = \frac{qD}{k_B T}. \tag{C3}$$

The defect generation rate $G$ is defined as

$$G = G_0 \exp\left(\frac{E_b - \gamma\varepsilon}{k_B T}\right), \tag{C4}$$

where $G_0$ is a preexponential constant, $E_b$ is the energy to form an oxygen vacancy at zero electric field, $\gamma$ is the bond polarization factor [17], and $\varepsilon$ is the local electric field.

The electric field and current distributions are determined from the continuity equation

$$\nabla . \sigma \nabla \psi = 0, \tag{C5}$$

where $\sigma$ is the local electrical conductivity of the dielectric layer, and $\psi$ is the potential.

Following the work of Larentis *et al.* [54], the electrical conductivity $\sigma$ is assumed to vary between semiconducting and metallic behavior depending on the defect density. The electrical conductivity is represented by an Arrhenius equation $\sigma = \sigma_o \exp(-E_{AC}/k_B T)$ in which the activation energy $E_{AC}$ and preexponential factor $\sigma_o$ depend on the defect concentration, as shown in Figs. 11(a) and 11(b), respectively. The thermally activated conductivity is consistent with Poole-Frenkel conduction, which has relatively low-energy barriers for conduction [55]. Initially, activation energy $E_{AC}$ is considered to be 0.05 eV at $n_D = 0$, which is in agreement with the semiconducting state of $HfO_x$-based RRAM [56]. At a low concentration of defects, $E_{AC}$ decreases, for increasing $n_D$ becomes zero at $n_D = 6 \times 10^{21}$ cm$^{-3}$ [Fig. 11(a)]. This assumption is based on fact that the conducting filament is equivalent to a highly doped semiconductor, where the Fermi level is very close or even above the conduction or valance band [55,57]. The preexponential factor for insulating $HfO_2$ is taken as $\sigma_o = 10^{-6}\ \Omega^{-1}\,\mathrm{cm}^{-1}$ and that of the metallic filament as $\sigma_o = 10^{4}\ \Omega^{-1}\,\mathrm{cm}^{-1}$ at $n_D = 6 \times 10^{22}$ cm$^{-3}$ as shown in Fig. 10(b). The threshold value for conducting $HfO_2$ is $n_D = 6 \times 10^{21}$ cm$^{-3}$, which corresponds to 10% of the oxygen vacancy for *p*-type conductivity. These assumptions are based on the experimental results of Hildebrandt *et al.* [58].

The temperature distribution is determined from the Fourier heat equation:

$$\rho C_p \frac{\partial T}{\partial t} - \nabla . k_{\mathrm{th}} \nabla T = Q, \qquad \text{(C6)}$$

where $\rho$ is the mass density, $C_p$ is the specific heat, $k_{\mathrm{th}}$ is the thermal conductivity, and $Q = \sigma \nabla |\psi|^2$ is the heat-flux density supplied by the electric current. Following Larentis *et al.* [54], a linear relationship between $k_{\mathrm{th}}$ and $n_D$ is assumed as shown in Fig. 10(c). The minimum value of thermal conductivity $k_{\mathrm{HfO_2}}$ refers to the thermal conductivity of the insulating $HfO_2$. The maximum $k_{\mathrm{th}}$ value at high $n_D$ corresponds to that of the metallic conducting filament, i.e., thermal conductivity of hafnium, $k_{\mathrm{Hf}}$. All parameters and values used in the simulation are listed in Table I.

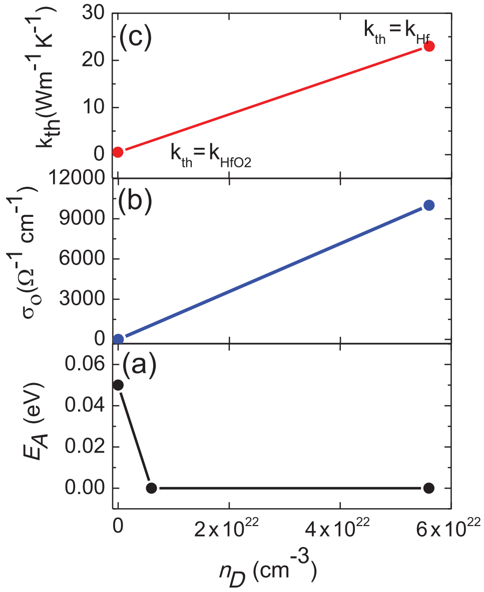


FIG. 10. (a) Activation energy $E_{AC}$ as a function of defect density $n_D$ (b) electrical conductivity preexponential factor $\sigma_0$, and (c) thermal conductivity $k_{\mathrm{th}}$ as a function of local defect density $n_D$.

TABLE I. Parameters for the simulation of dielectric break down in $HfO_2$.

| Parameter | Value | References |
|---|---|---|
| $D_o$ | $2 \times 10^{-3}$ cm$^2$ s | [20] |
| $E_A$ | 1 eV | [59] |
| $E_b$ | 2.8 eV | [58] |
| $\gamma$ | 10.13 eÅ | [60] |
| $n_A$ | $5.54 \times 10^{22}$ cm$^{-3}$ | [61] |
| $G_0$ | $7 \times 10^{13}$ Hz | [16] |
| $C_p$ | 60 Jkg$^{-1}$ K$^{-1}$ | [62] |
| $k_{\mathrm{HfO_2}}$ | 0.5 Wm$^{-1}$ K$^{-1}$ | [63] |
| $k_{\mathrm{Hf}}$ | 23 Wm$^{-1}$ K$^{-1}$ | [18] |

## APPENDIX D: GEOMETRIC DEPENDENCE OF FIELD ENHANCEMENT

The sensitivity of field-enhancement factor $\eta$ to the shape of the asperity can be estimated from comparison of $\eta$ versus aspect ratio $(h/w)$ [64]. The geometric field-enhancement factor $(\eta)$ is plotted as a function of the aspect ratio as shown in Fig. 11. The peak $\eta$ enhancement versus aspect ratio $(h/w)$ relation can be fitted by considering a power law, such that

$$\eta = p + m\left(\frac{h}{w}\right)^n, \qquad \text{(D1)}$$

where $p$, $m$, and $n$ are fitting parameters depending on the geometry of the asperity.

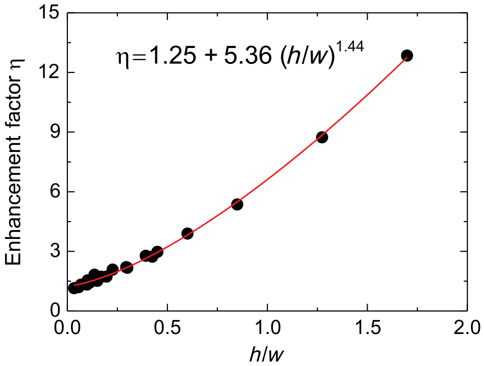


FIG. 11. The electric-field-enhancement $(\eta)$ factor as a function of the width of the asperity. The solid line is a fit by a power law.

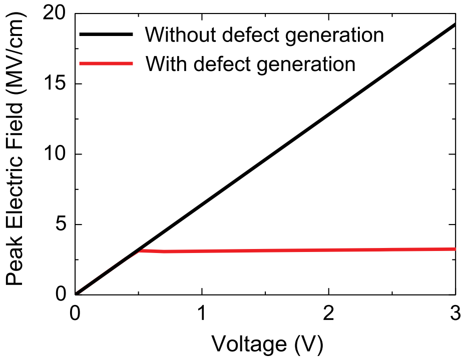


FIG. 12. Peak field as a function of the applied field with and without defect generation for an asperity of 8 nm height and width 4.71 nm.

## APPENDIX E: ELECTRIC-FIELD MODERATION

To demonstrate that the moderation of the electric field is caused by defects, simulations are performed with and without defect generation. Figure 12 shows the peak electric field as a function of applied voltage for an asperity of height 8 nm and width 4.71 nm. The peak field is observed to increase linearly with the voltage when the defect generation is neglected but it saturates at 2.9 MV/cm when the defect generation is included. It is clearly evident that the generation of defects acts to moderate geometric field enhancement.

## APPENDIX F: EFFECT OF ASPERITY WIDTH ON FILAMENT SIZE

The shape or size of the filament depends on the geometry of the asperity. Figure 13 shows the initial electric-field distribution and filament shape after a soft breakdown as a function of the asperity width. When a bias is applied to the top electrode, defects are generated at the tip of the asperity and gradually increase in concentration to form a defect cluster. From Figs. 13(a)–13(c), it is evident that the effective volume of the field enhancement is a function of the width of the asperity. Therefore, for a wider asperity, the defect generation is initiated over a broader volume and attracted toward the top electrode along the electric-field direction. This results in a wider filament formation as described in Figs. 13(d)–13(f).

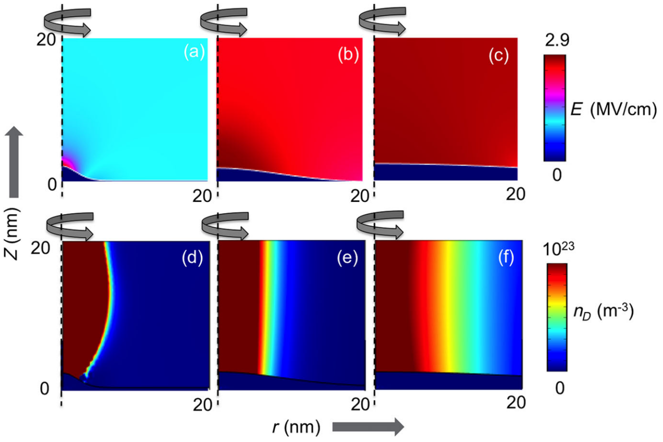


FIG. 13. 2D plot of (a)–(c) initial electric field and (d)–(f) defect distribution ($n_D$) at the forming voltage for the same height of asperity but a different width, which demonstrates that the filament size depends on the width.

[1] D. L. Crook, Method of determining reliability screens for time dependent dielectric breakdown, in *Proceedings of the IEEE 17th Annual Reliability Physics Symposium, 1979* (IEEE, New York, 1979), p. 1.

[2] T. DiStefano, Dielectric breakdown induced by sodium in MOS structures, J. Appl. Phys. **44**, 527 (1973).

[3] T. Tsuruoka, K. Terabe, T. Hasegawa, and M. Aono, Forming and switching mechanisms of a cation-migration-based oxide resistive memory, Nanotechnology **21**, 425205 (2010).

[4] J. Son and Y. -H. Shin, Direct observation of conducting filaments on resistive switching of NiO thin films, Appl. Phys. Lett. **92**, 222106 (2008).

[5] S. Krishnan, E. Stefanakos, and S. Bhansali, Effects of dielectric thickness and contact area on current–voltage characteristics of thin film metal–insulator–metal diodes, Thin Solid Films **516**, 2244 (2008).

[6] K.-H. Allers, P. Brenner, and M. Schrenk, Dielectric reliability and material properties of $Al_2O_3$ in metal insulator metal capacitors (MIMCAP) for RF bipolar technologies in comparison to $SiO_2$, SiN and $Ta_2O_5$, in *Proceedings of the IEEE Bipolar/BiCMOS Circuits and Technology Meeting, 2003* (IEEE, New York, 2003), p. 35.

[7] B. H. Lee, C. Y. Kang, P. Kirsch, D. Heh, C. D. Young, H. Park, J. Yang, G. Bersuker, S. Krishnan, and R. Choi, Electric-field-driven dielectric breakdown of metal-insulator-metal hafnium silicate, Appl. Phys. Lett. **91**, 243514 (2007).

[8] H. Frohlich, On the theory of dielectric breakdown in solids, Proc. R. Soc. A **188**, 521 (1947).

[9] S. K. Nandi, X. Liu, D. K. Venkatachalam, K. Belay, and R. G. Elliman, Self-assembly of an $NbO_2$ interlayer and configurable resistive switching in $Pt/Nb/HfO_2/Pt$ structures, Appl. Phys. Lett. **107**, 132901 (2015).

[10] S. K. Nandi, X. Liu, D. K. Venkatachalam, and R. G. Elliman, Threshold current reduction for the metal–insulator transition in $NbO_{2-x}$-selector devices: The effect of ReRAM integration, J. Phys. D **48**, 195105 (2015).

[11] J. J. Yang and R. S. Williams, Memristive devices in computing system: Promises and challenges, ACM J. Emerging Technol. Comput. Syst. **9**, 1 (2013).

[12] J. Y. Seok, S. J. Song, J. H. Yoon, K. J. Yoon, T. H. Park, D. E. Kwon, H. Lim, G. H. Kim, D. S. Jeong, and C. S. Hwang, A review of three-dimensional resistive switching cross-bar array memories from the integration and materials property points of view, Adv. Funct. Mater. **24**, 5316 (2014).

[13] G. H. Kim, J. H. Lee, Y. Ahn, W. Jeon, S. J. Song, J. Y. Seok, J. H. Yoon, K. J. Yoon, T. J. Park, and C. S. Hwang,

$32 \times 32$ crossbar array resistive memory composed of a stacked Schottky diode and unipolar resistive memory, Adv. Funct. Mater. **23**, 1440 (2013).

[14] 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, and S. Seo, 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).

[15] H. Jiang and Q. Xia, Effect of voltage polarity and amplitude on electroforming of $TiO_2$ based memristive devices, Nanoscale **5**, 3257 (2013).

[16] J. McPherson, J. Kim, A. Shanware, and H. Mogul, Thermochemical description of dielectric breakdown in high dielectric constant materials, Appl. Phys. Lett. **82**, 2121 (2003).

[17] J. McPherson and H. Mogul, Underlying physics of the thermochemical E model in describing low-field time-dependent dielectric breakdown in $SiO_2$ thin films, J. Appl. Phys. **84**, 1513 (1998).

[18] L. Vandelli, A. Padovani, L. Larcher, and G. Bersuker, Microscopic modeling of electrical stress-induced breakdown in poly-crystalline hafnium oxide dielectrics, IEEE Trans. Electron Devices **60**, 1754 (2013).

[19] G. Bersuker, D. Gilmer, D. Veksler, P. Kirsch, L. Vandelli, A. Padovani, L. Larcher, K. McKenna, A. Shluger, and V. Iglesias, Metal oxide resistive memory switching mechanism based on conductive filament properties, J. Appl. Phys. **110**, 124518 (2011).

[20] A. Padovani, L. Larcher, G. Bersuker, and P. Pavan, Charge transport and degradation in $HfO_x$ and $HfO_2$ dielectrics, IEEE Electron Device Lett. **34**, 680 (2013).

[21] S. K. Nandi, X. Liu, S. Li, D. K. Venkatachalam, K. Belay, and R. G. Elliman, Resistive switching behavior in $HfO_2$ with Nb as an oxygen exchange layer, in *Proceedings of the 2014 IEEE Conference on Optoelectronic and Microelectronic Materials and Devices (COMMAD), 2014* (IEEE, New York, 2014), p. 290.

[22] Y.-P. Zhao, G.-C. Wang, T.-M. Lu, G. Palasantzas, and J. T. M. De Hosson, Surface-roughness effect on capacitance and leakage current of an insulating film, Phys. Rev. B **60**, 9157 (1999).

[23] M. Lopes, S. dos Santos Fo, C. Hasenack, and V. Baranauskas, Si-$SiO_2$ electronic interface roughness as a consequence of Si-$SiO_2$ topographic interface roughness, J. Electrochem. Soc. **143**, 1021 (1996).

[24] K. Chen, M. Nielsen, S. Soss, E. J. Rymaszewski, T.-M. Lu, and C. Wan, Study of tantalum oxide thin film capacitors on metallized polymer sheets for advanced packaging applications, IEEE Trans Compon, Packag Manuf Technol, Part B **20**, 117 (1997).

[25] N. Gaillard, L. Pinzelli, M. Gros-Jean, and A. Bsiesy, *In situ* electric field simulation in metal/insulator/metal capacitors, Appl. Phys. Lett. **89**, 133506 (2006).

[26] Q. Xia, J. J. Yang, W. Wu, X. Li, and R. S. Williams, Self-aligned memristor cross-point arrays fabricated with one nanoimprint lithography step, Nano Lett. **10**, 2909 (2010).

[27] Y.-P. Zhao, G.-C. Wang, T.-M. Lu, G. Palasantzas, and J. T. M. De Hosson, Surface-roughness effect on capacitance and leakage current of an insulating film, Phys. Rev. B **60**, 9157 (1999).

[28] P. A. Kumar, B. Panda, S. Ray, B. Mathur, D. Bhattacharya, and K. Chopra, Effect of electrode microstructure on leakage current in lead–lanthanum–zirconate–titanate multilayer capacitors, Appl. Phys. Lett. **68**, 1344 (1996).

[29] S.-D. Cho and K.-W. Paik, Study on the amorphous $Ta_2O_5$ thin film capacitors deposited by dc magnetron reactive sputtering for multichip module applications, Mater. Sci. Eng. B **67**, 108 (1999).

[30] Y. S. Kim, Y. H. Lee, K. M. Lim, and M. Y. Sung, The effect of $Al-Ta_2O_5$ topographic interface roughness on the leakage current of $Ta_2O_5$ thin films, Appl. Phys. Lett. **74**, 2800 (1999).

[31] Y. Kim, M. Sung, Y. Lee, B. Ju, and M. Oh, The influence of surface roughness on the electric conduction process in amorphous $Ta_2O_5$ thin films, J. Electrochem. Soc. **146**, 3398 (1999).

[32] P. L. Grande, M. Vos, D. K. Venkatachalam, S. K. Nandi, and R. G. Elliman, Determination of thickness and composition of high-$k$ dielectrics using high-energy electrons, Appl. Phys. Lett. **103**, 071911 (2013).

[33] M. Vos, P. L. Grande, D. K. Venkatachalam, S. K. Nandi, and R. G. Elliman, Oxygen Self-Diffusion in $HfO_2$ Studied by Electron Spectroscopy, Phys. Rev. Lett. **112**, 175901 (2014).

[34] J. Zhang and J. B. Adams, Modeling and visualization of polycrystalline thin film growth, Comput. Mater. Sci. **31**, 317 (2004).

[35] A. W. Czanderna, T. E. Madey, and C. J. Powell, *Beam Effects, Surface Topography, and Depth Profiling in Surface Analysis* (Springer, New York, 1998).

[36] G. Palasantzas, Roughness spectrum and surface width of self-affine fractal surfaces via the $K$-correlation model, Phys. Rev. B **48**, 14472 (1993).

[37] G. Palasantzas, Erratum: Roughness spectrum and surface width of self-affine fractal surfaces via the $K$-correlation model, Phys. Rev. B **49**, 5785 (1994).

[38] Z. Chai, Y. Liu, X. Lu, and D. He, Reducing adhesion force by means of atomic layer deposition of ZnO films with nanoscale surface roughness, ACS Appl. Mater. Interfaces **6**, 3325 (2014).

[39] V. Jovanov, X. Xu, S. Shrestha, M. Schulte, J. r. Hüpkes, and D. Knipp, Predicting the interface morphologies of silicon films on arbitrary substrates: application in solar cells, ACS Appl. Mater. Interfaces **5**, 7109 (2013).

[40] D. Fabiani and L. Simoni, Discussion on application of the Weibull distribution to electrical breakdown of insulating materials, IEEE Trans. Dielectr. Electr. Insul. **12**, 11 (2005).

[41] S. Yao, Theoretical model of thin-film deposition profile with shadow effect, J. Appl. Phys. **50**, 3390 (1979).

[42] R. Karunasiri, R. Bruinsma, and J. Rudnick, Thin-Film Growth and the Shadow Instability, Phys. Rev. Lett. **62**, 788 (1989).

[43] M. S. Mathur, The engineering and optimization of the hafnium based metal oxide semiconductor structure, Ph.D. thesis, University of California, 2008.

[44] A. Chen, Area and thickness scaling of forming voltage of resistive switching memories, IEEE Electron Device Lett. **35**, 57 (2014).

[45] N. Gaillard, L. Pinzelli, M. Gros-Jean, and A. Bsiesy, *In situ* electric field simulation in metal/insulator/metal capacitors, Appl. Phys. Lett. **89**, 133506 (2006).

[46] M. Elrawemi, L. Blunt, L. Fleming, D. Bird, D. Robbins, and F. Sweeney, Modelling water vapour permeability through atomic layer deposition coated photovoltaic barrier defects, Thin Solid Films **570**, 101 (2014).

[47] M. Groner, J. Elam, F. Fabreguette, and S. M. George, Electrical characterization of thin $Al_2O_3$ films grown by atomic layer deposition on silicon and various metal substrates, Thin Solid Films **413**, 186 (2002).

[48] Y. Zhang, J. A. Bertrand, R. Yang, S. M. George, and Y. Lee, Electroplating to visualize defects in $Al_2O_3$ thin films grown using atomic layer deposition, Thin Solid Films **517**, 3269 (2009).

[49] J. Dendooven, K. Devloo-Casier, M. Ide, K. Grandfield, K. F. Ludwig, S. Bals, P. Van Der Voort, and C. Detavernier, *In situ* study of ALD processes using synchrotron-based x-ray fluorescence and scattering techniques, ECS Trans. **50**, 35 (2013).

[50] A. Lankhorst, B. Paarhuis, H. Terhorst, P. Simons, and C. Kleijn, Transient ALD simulations for a multi-wafer reactor with trenched wafers, Surf. Coat. Technol. **201**, 8842 (2007).

[51] H. Knoops, E. Langereis, M. van de Sanden, and W. Kessels, Conformality of plasma-assisted ALD: Physical processes and modeling, J. Electrochem. Soc. **157**, G241 (2010).

[52] J. Molina, R. Valderrama, C. Zuniga, P. Rosales, W. Calleja, A. Torres, J. D. Hidalga, and E. Gutierrez, Influence of the surface roughness of the bottom electrode on the resistive-switching characteristics of $Al/Al_2O_3/Al$ and $Al/Al_2O_3/W$ structures fabricated on glass at 300 °C, Microelectron. Reliab. **54**, 2747 (2014).

[53] H.-D. Kim, M. J. Yun, S. M. Hong, J. H. Park, D. S. Jeon, and T. G. Kim, Impact of roughness of bottom electrodes on the resistive switching properties of platinum/nickel nitride/nickel 1 × 1 crossbar array resistive random access memory cells, Microelectron. Eng. **126**, 169 (2014).

[54] S. Larentis, F. Nardi, S. Balatti, D. C. Gilmer, and D. Ielmini, Resistive switching by voltage-driven ion migration in bipolar RRAM—Part II: Modeling, IEEE Trans. Electron Devices **59**, 2468 (2012).

[55] D. Ielmini, F. Nardi, and C. Cagli, Physical models of size-dependent nanofilament formation and rupture in NiO resistive switching memories, Nanotechnology **22**, 254022 (2011).

[56] C. Walczyk, D. Walczyk, T. Schroeder, T. Bertaud, M. Sowinska, M. Lukosius, M. Fraschke, D. Wolansky, B. Tillack, and E. Miranda, Impact of temperature on the resistive switching behavior of embedded-based RRAM devices, IEEE Trans. Electron Devices **58**, 3124 (2011).

[57] T. Lee, R. Pashley, T. McGill, and J. Mayer, Investigation of tellurium-implanted silicon, J. Appl. Phys. **46**, 381 (1975).

[58] E. Hildebrandt, J. Kurian, M. M. Muller, T. Schroeder, H.-J. Kleebe, and L. Alff, Controlled oxygen vacancy induced $p$-type conductivity in $HfO_2$-$x$ thin films, Appl. Phys. Lett. **99**, 112902 (2011).

[59] M. Vos, P. Grande, D. Venkatachalam, S. Nandi, and R. Elliman, Oxygen Self-Diffusion in $HfO_2$ Studied by Electron Spectroscopy, Phys. Rev. Lett. **112**, 175901 (2014).

[60] T. J. Park, J. H. Kim, J. H. Jang, C.-K. Lee, K. D. Na, S. Y. Lee, H.-S. Jung, M. Kim, S. Han, and C. S. Hwang, Reduction of electrical defects in atomic layer deposited $HfO_2$ films by Al doping, Chem. Mater. **22**, 4175 (2010).

[61] M. A. Panzer, M. Shandalov, J. A. Rowlette, Y. Oshima, Y. W. Chen, P. C. McIntyre, and K. E. Goodson, Thermal properties of ultrathin hafnium oxide gate dielectric films, IEEE Electron Device Lett. **30**, 1269 (2009).

[62] N. P. Long, Y. Tanaka, Y. Uesugi, and Y. Yamaguchi, Numerical investigation of the effect of cathode holder shape on hafnium cathode evaporation for an oxygen plasma cutting arc, J. Phys. D **46**, 224012 (2013).

[63] J. Suné, M. Nafria, E. Miranda, X. Oriols, R. Rodriguez, and X. Aymerich, Failure physics of ultra-thin $SiO_2$ gate oxides near their scaling limit, Semicond. Sci. Technol. **15**, 445 (2000).

[64] G. S. Bocharov and A. V. Eletskii, Theory of carbon nanotube (CNT)-based electron field emitters, Nanomater. Nanotechnol. **3**, 393 (2013).