scientific reports

OPEN

Multilayer redox-based HfO$_x$/Al$_2$O$_3$/TiO$_2$ memristive structures for neuromorphic computing

Seongae Park, Benjamin Spetzler✉, Tzvetan Ivanov & Martin Ziegler

Redox-based memristive devices have shown great potential for application in neuromorphic computing systems. However, the demands on the device characteristics depend on the implemented computational scheme and unifying the desired properties in one stable device is still challenging. Understanding how and to what extend the device characteristics can be tuned and stabilized is crucial for developing application specific designs. Here, we present memristive devices with a functional trilayer of HfO$_x$/Al$_2$O$_3$/TiO$_2$ tailored by the stoichiometry of HfO$_x$ (x = 1.8, 2) and the operating conditions. The device properties are experimentally analyzed, and a physics-based device model is developed to provide a microscopic interpretation and explain the role of the Al$_2$O$_3$ layer for a stable performance. Our results demonstrate that the resistive switching mechanism can be tuned from area type to filament type in the same device, which is well explained by the model: the Al$_2$O$_3$ layer stabilizes the area-type switching mechanism by controlling the formation of oxygen vacancies at the Al$_2$O$_3$/HfO$_x$ interface with an estimated formation energy of ≈ 1.65 ± 0.05 eV. Such stabilized area-type devices combine multi-level analog switching, linear resistance change, and long retention times (≈ 10^7–10^8 s) without external current compliance and initial electroforming cycles. This combination is a significant improvement compared to previous bilayer devices and makes the devices potentially interesting for future integration into memristive circuits for neuromorphic applications.

Memristive devices are typically two-terminal resistive elements with a hysteretic current–voltage characteristic, which enables switching between different resistance states[1–3]. As analog resistive switching memory, memristive devices have emerged as one of the critical technologies for implementing neuromorphic computing systems[4,5], i.e., systems that mimic the structure or working mechanisms of the brain[6]. For such computing systems, memristive devices offer a variety of potential advantages compared to traditional memory devices such as flash or SRAM[4]. The simple two-terminal device structure of memristive elements permits a high integration density in crossbar arrays[7,8], and thereby, highly energy-efficient and parallelized in-memory computing. Additionally, the possibility of in-memory computing results in a low latency because it is not restricted by the data exchange between the memory and a central processing unit (von Neumann bottleneck)[4,5].

A major challenge for memristive devices in neuromorphic systems is a controllable and reproducible setting of resistance states for storing synaptic weights in artificial neuronal networks (ANNs). This requires high endurance and reproducibility of the device properties over many switching cycles[4,9]. In general, the requirements of the device parameters depend on the kind of weight update of the respective neuromorphic system. For example, the spike-timing dependent plasticity (STDP) rule is mainly relevant for spiking neural networks (SNN) and requires a time dependent and nonlinear resistance change[10]. In contrast, classical ANNs (such as deep neural networks or convolutional neural networks) do not require time dependency, but instead a linear resistance change and as many stable resistance values as possible[11]. Although the requirements imposed on the memory duration are lower in many areas of neuromorphic computing than in nonvolatile memories[12,13], too short state retention times preclude the devices from broad application in neuromorphic computing.

The properties of memristive devices depend on the underlying mechanisms of the resistive switching process, which can be broadly classified into filament-type switching and area-type switching[14,15]. Filamentary switching is characterized by the presence of (at least) one conductive filament, which provides an electrical pathway through the memristive material in the low resistive state. It typically must be induced with an initial forming procedure before operating the device[3,5]. The growth and disruption of the filament are stochastic processes and cause an abrupt change in the resistance between a high resistive state (HRS) and a low resistive state (LRS), usually without accessible intermediate resistance states (digital switching). Due to the high nonlinearity of the

Micro- and Nanoelectronic Systems, Institute of Micro and Nanotechnologies MacroNano, Technische Universität Ilmenau, Ilmenau, Germany. ✉email: benjamin.spetzler@tu-ilmenau.de

switching kinetics, additional electronics is typically required to stabilize the current–voltage characteristic by limiting the current through the device (current compliance). Limiting the current prevents uncontrolled filament formation and growth, leading to irreversible changes in the device characteristics and rendering the device useless[16]. From an application point of view, the highly nonlinear switching kinetics have enabled filamentary devices capable of long retention times > 10 years[17,18], high switching speeds < 1 ns[19,20], high endurance > 10^{12} cycles[7,21], and large on–off ratios (> 1000)[5,22]. These properties have made them attractive as a potential technology for nonvolatile memory approaches[23]. Major challenges are the low reproducibility of device parameters and endurance deterioration because of the underlying stochastic process[22], but also the required initial forming step and additional electronics for the current compliance, which can be disadvantageous for the integration of the devices into large arrays[4]. On the other hand, area-type switching is a less mature technology[14]. Depending on the specific device, various potentially involved microscopic processes are under debate[24,25]. Generally, the current through the device is assumed to be dominated by interface-limited conduction processes, such as the tunneling or emission through/over interface potential barriers. Switching has been explained as a modulation of the barrier via the drift and accumulation of mobile charge carriers under the application of an external voltage. In this context, electromigration of mobile vacancies has been discussed[26–30], as also the trapping dynamics of electrons[31,32] and the uniform modification of bulk conduction mechanisms[14]. Because of the spatially uniform contribution of conduction mechanisms to the total current, the conductance scales with the device area and the stochastic components of the microscopic processes average and yield more reproducible macroscopic device properties than filamentary devices. Area-type devices can be fabricated with a low variation in the device characteristics and a correspondingly high yield[33]. Additional advantages are many intermediate and continuously accessible resistance states (analog switching) and often the possibility of operating without initial forming cycles or external current compliance[34–36]. However, the main challenges of area-type devices are a typically short retention time (seconds to days) and large switching times (> 1 ms) compared to filamentary devices[12,37]. A combination of the beneficial properties of both switching types is generally desirable.

Among various types of memristive devices investigated for neuromorphic computing[1,5,9,38–40], memristive bilayer and multilayer oxide structures based on hafnium oxide and titanium oxide gained attention for the application as artificial synapses[35,41–43]. Such oxide structures can show filamentary- or area-type switching[23,42], depending on the details of the respective layer system. Compared to hafnium oxide single-layer devices, HfO_x/TiO_x bilayer structures demonstrated a variety of improvements[43], such as a smaller variability[44] and improved resistance modulation linearity[45], and a larger number of resistance states[46]. These bilayer structures are mainly of the filamentary type and often still require current compliance and an initial electroforming step. Previous studies[47,48] suggest that the filament formation in such systems is connected with the concentration of oxygen vacancies in the HfO_x layer and oxygen ions in the TiO_x layer. According to this work, the ions and vacancies are injected preferably at the HfO_x/TiO_x interface by a voltage-induced creation of Frenkel pairs, owing to much higher activation energy in the bulk material[47,48]. Consequently, altering the layer structure and interfaces could represent a way of modifying the switching type and further improve the device characteristics for the applications pursued.

This work presents memristive trilayer devices with a functional layer stack of HfO_x/Al_2O_3/TiO_2 sandwiched between TiN and Au electrodes. We discuss the influence of the additional aluminum oxide layer, operating conditions, and the composition of the hafnium oxide layer on the resistive switching characteristics. A model is devised to provide a microscopic interpretation of the experimental results and derive further implications for the design of memristive multilayer elements. The paper is structured as follows: after presenting the details of the device materials and geometries in "Memristive devices and materials", we analyze the current–voltage characteristics under different voltage operating conditions and for two different hafnium oxide compositions in the section "Current–voltage characteristics". In "Discussion of the switching mechanisms", the indications found for the underlying switching mechanisms are interpreted before they are quantified with a physics-based device model presented in "Physics-based compact model". Afterwards, we discuss application relevant device characteristics in "Characterization for neuromorphic applications", namely the state retention time and the linearity in the conductance update. Finally, the results are summarized, and a conclusion is drawn.

Results

Memristive devices and materials. All memristive devices presented here are fabricated with a 4-in. wafer technology, providing around 40,000 individual devices per wafer (Fig. 1a), arranged in groups of six (Fig. 1b). Each device can be electrically contacted via individual contact pads for the top electrode and a shared contact pad for the rear-side electrode. The top and rear-side electrodes are electrically insulated by a 180-nm-thick SiO_2 layer, encapsulating the functional layers underneath. The top electrode contacts a 30-nm-thick Au layer, located under the SiO_2 and on the top of the functional layer stack. Because the Au layer translates the electric potential applied to the top electrode to the functional layers, the lateral dimensions of the Au layer are used to define a device area A, as indicated by the dashed square in Fig. 1c. Within each group of six devices, the device area varies from a minimum of 10^2 μm^2 to a maximum of 50^2 μm^2. A schematic cross-section of the entire layer stack is illustrated in Fig. 1d. The functional layer stack comprises a sequence of HfO_x(3 nm)/Al_2O_3(2 nm)/TiO_2(15 nm), sandwiched between the Au contact layer and the TiN rear-side electrode. Here, the HfO_x is intended to introduce memristive behavior, the Al_2O_3 to modify the interface properties and the TiO_2 layer is beneficial because it forms well-defined interfaces with the TiN electrode and the Al_2O_3 interlayer. While the thickness of the layers is kept constant throughout the various devices analyzed, we focus on two stoichiometries of the hafnium oxide (HfO_x) layer with $x \approx 1.8$ and $x \approx 2$. The stoichiometries were confirmed by x-ray photoelectron spectroscopy (XPS) measurements on a test wafer with estimated uncertainties of ± 1%.
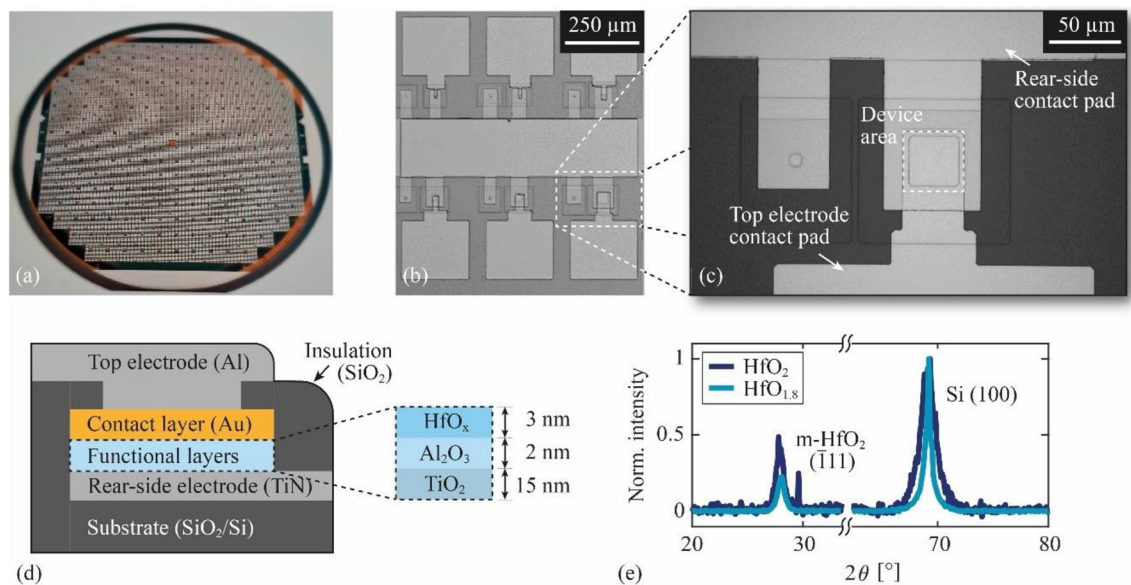
Figure 1. Illustration of the fabricated devices (**a**) 4-in. wafer with around 40,000 memristive devices, (**b**) arranged in groups of six with different device areas, (**c**) close-up top view of an individual device showing the top and rear-side electrode contact pads, and the device area is indicated. (**d**) Cross-sectional sketch along the dashed line in (**c**) of an example device with the configuration of the functional layers. (**e**) Normalized and background subtracted XRD intensity spectra of hafnium oxide, deposited with a dynamic (HfO_2) and a static ($HfO_{1.8}$) sputtering method on a Si substrate, indicating the presence of nanocrystalline m-HfO_2.
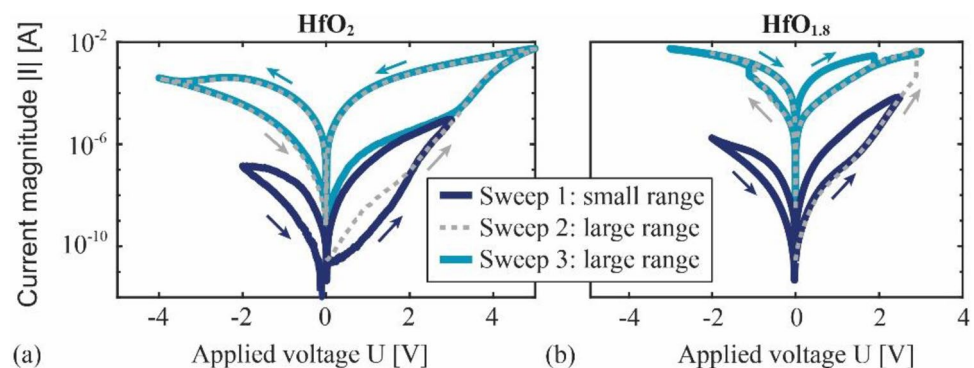

Figure 2. Results of three subsequently performed current–voltage measurements on two representative memristive devices with different stoichiometries of the HfO_x layer. (**a**) Measurements on a device with an HfO_2 layer and (**b**) device with an $HfO_{1.8}$ layer. The voltage was swept piecewise linearly, starting and ending at $U = 0$ V for 45 s.

The structural properties are determined with x-ray diffraction (XRD) measurements on hafnium oxide layers deposited on 4-in. Si wafers. The normalized and background subtracted intensity spectra in Fig. 1e show two distinct peaks at $2\theta = 69°$ and $2\theta = 28°$. While the former corresponds to the (100) planes of the Si substrate, the latter can be assigned to the ($\bar{1}11$) planes of monoclinic HfO_2 (m-HfO_2)[49]. Therefore, we assume the presence of nanocrystalline m-HfO_2 in our films, as demonstrated in previous investigations on hafnium oxide thin films[50–52]. From the difference in the intensity of the m-HfO_2 peak, larger average grain size and better oriented ($\bar{1}11$) planes are expected in the HfO_2 sample compared to the $HfO_{1.8}$ sample[50].

Current–voltage characteristics. We conducted quasi-static measurements of the current–voltage (I-U) hysteresis on two representative example devices with a device area of $A = 35^2$ μm², to identify the influence of the stoichiometry of the HfO_x layer on the electric device characteristics. All measurements were performed at room temperature without current compliance, within a time of $t = 45$ s per loop. The obtained I-U curves are shown in Fig. 2. Additional measurements on various devices demonstrate stable hysteresis curves over at least 100 cycles with only minor deviations in the current magnitudes ("Methods").

After an initial voltage sweep (Sweep 1 in Fig. 2), two curves (Sweep 2, Sweep 3 in Fig. 2) with an increased range of the applied voltage U were measured. The results of the three voltage sweeps are presented in Fig. 2a
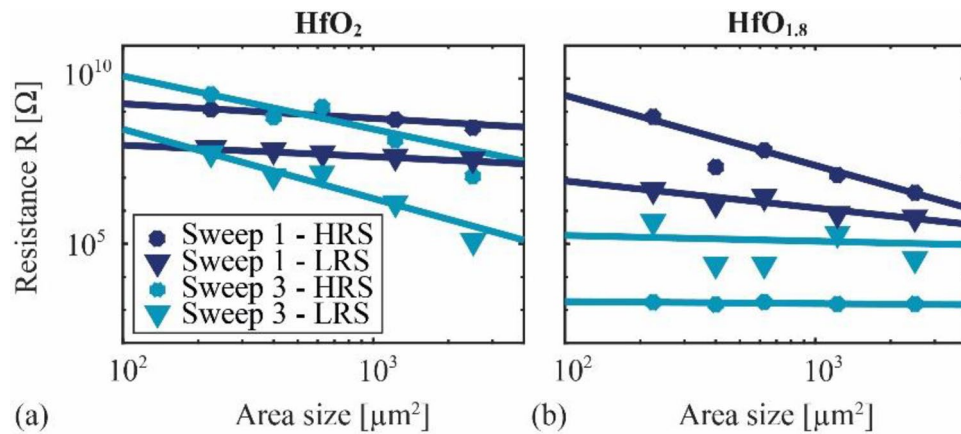
Figure 3. Dependency of the device resistances R_{HRS} and R_{LRS} on the active device area A in the high resistive state (HRS) and low resistive state (LRS), respectively, and for both voltage ranges. The plotted mean values were obtained from measurements on many devices, with (**a**) HfO$_2$ layers and (**b**) HfO$_{1.8}$ layers.

(HfO$_2$) and Fig. 2b (HfO$_{1.8}$). In the initial voltage sweep (Sweep 1), the I-U curves of both samples show a similarly distinct hysteresis with a smooth transition between the high resistive state (HRS) and the low resistive state (LRS). The I-U characteristics change notably upon increasing the voltage range in the second measurement (Sweep 2). The larger voltage range in both samples reveals a transition to a second hysteresis curve with a significantly larger maximum current and lower resistance. While this transition occurs continuously in the HfO$_2$ device, a jump is visible in the I-U curve of the HfO$_{1.8}$ device. An additional measurement (Sweep 3) confirms the reproducibility and irreversibility of the low-resistance hysteresis curves and reveals further differences and features in the I-U characteristics of the two samples. In the HfO$_2$ device, the hysteresis area increases compared to the initial sweep (Sweep 1), and the curve remains smooth. In contrast, in the HfO$_{1.8}$ device, the hysteresis area reduces, and the direction of the hysteresis loop is reversed. The transition from the HRS to the LRS occurs via discontinuities at $U \approx -1$ V and back at $U \approx 2$ V. While such discontinuities are an indication of filament-type switching[53], the continuous resistance change observed in all small-range sweeps and the large-range sweep of the HfO$_2$ device, is a typical characteristic of area-type switching[24].

These indications for the type of switching mechanisms can be confirmed by analyzing the dependency of the resistance $R := U/I$ on the device area. For that, we measured the I-U curves of many samples with different device areas A in both voltage ranges to extract the resistances in the high and the low resistive state, labeled as R_{HRS} and R_{LRS}, respectively. The mean values of R_{HRS} and R_{LRS} are plotted in Fig. 3a (HfO$_2$), and Fig. 3b (HfO$_{1.8}$), together with fits of the data sets to guide the eye. The relative standard deviations of the mean values are provided in Table 2 ("Methods") because they are so small that they would not be clearly visible in Fig. 3, which indicates a good device-to-device reproducibility. A reduction of R_{HRS} and R_{LRS} with increasing device area is visible in all measurements of the small voltage range (Sweep 1). This trend is stronger in the measurements of the large voltage range (Sweep 3) of the HfO$_2$ devices, while no clear dependency on the area is visible in the large voltage range measurements on the HfO$_{1.8}$ devices (Sweep 3, HRS). Consistent with the I-U measurements, the area independent resistance of the HfO$_{1.8}$ devices supports the presence of a dominant filamentary conduction mechanism, and the area dependency of all other configurations confirms a dominant area-type conduction mechanism[14,23]. Hence, we identified two different switching mechanisms in the same device, as well as their dependency on the voltage treatment and the composition of the hafnium oxide layer. An explanation for the observed behavior, i.e., the reversed hysteresis direction, the voltage tunability of the I-U characteristics, and the current self-compliance, requires a discussion of the underlying switching mechanisms in the following section.

Discussion of the switching mechanisms. The current–voltage characteristics and the observed change from an area to a filament-type switching behavior are explained in this section with a qualitative model considering the presence and dynamics of charged point defects. Based on the assumptions and measurements, we then estimate the formation energy of these defects.

Following ab initio calculations, double positively charged oxygen vacancies V_O^{2+} and negatively charged interstitials O_i^{2-} are mobile in HfO$_2$ with activation energies of diffusion of 0.69 eV (V_O^{2+})[54,55] and 0.6 eV (O_i^{-2})[56]. However, their formation via the Frenkel mechanism $O_O^0 \rightarrow V_O^{2+} + O_i^{2-}$ in the bulk of HfO$_2$ can be omitted at room temperature owing to large activation energies of 5–9 eV[57,58], also in the presence of high electric fields[58]. Instead, oxygen vacancies could be created at the HfO$_x$/Al$_2$O$_3$ interface by incorporating oxygen into the aluminum oxide layer, functioning as an oxygen reservoir/scavenger, which increases the formation energy of oxygen interstitials and reduces the formation energy of oxygen vacancies[59]. This mechanism was shown to be energetically favored over the formation of Frenkel pairs in the bulk of hafnium/hafnium oxide layers[60], and could similarly apply here. Oxygen vacancies are also expected to be introduced during the sputtering process with densities depending on the stoichiometries of the deposited films. Therefore, the switching processes in our devices are likely dominated by the dynamics and formation of V_O^{2+} as illustrated in Fig. 4, while oxygen interstitials can be omitted.
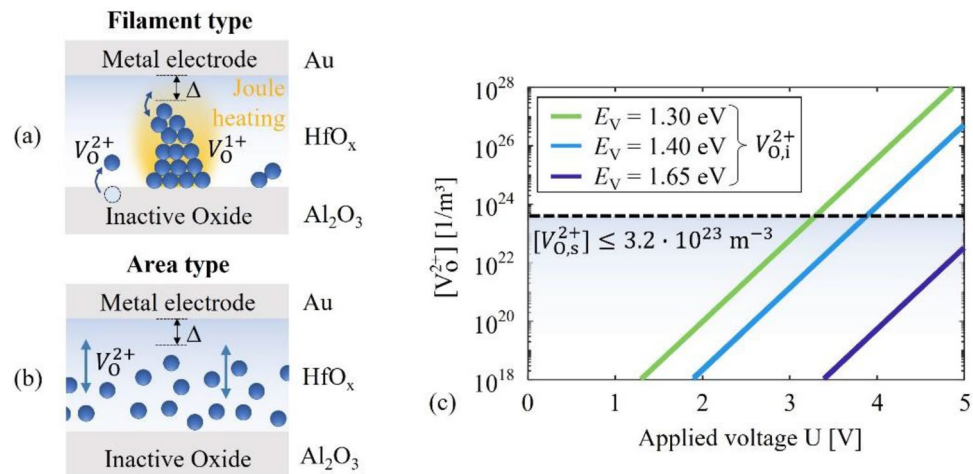
Figure 4. Illustration of the role of defect dynamics for the two switching types and the formation of oxygen vacancies V_O^{2+} at the interface. (**a**) Filamentary switching is based on the formation, growth, and disruption of an oxygen vacancy filament in defect rich HfO$_x$. (**b**) Area-type switching in highly stoichiometric hafnium oxide, where the drift and diffusion of mobile vacancies change the potential landscape. (**c**) Density $[V_O^{2+}]$ of double positively charged oxygen vacancies V_O^{2+} in the HfO$_2$ layer, calculated as a function of the applied voltage U and plotted together with the stoichiometric vacancy density $\left[V_{O,s}^{2+}\right]$ (maximum indicated by black, dashed line) in the HfO$_2$ layer, which is still consistent with the XPS measurements. The electric field-induced vacancy density $[V_{O,i}^{2+}]$ is plotted for three different formation energies E_v.

The stoichiometry of the HfO$_{1.8}$ layer is close to the computationally estimated ideal stoichiometry (HfO$_{1.5}$–HfO$_{1.75}$)[61] for the nucleation of a filament of oxygen vacancies (Fig. 4a). This process might be further fostered by the voltage-induced formation of vacancies during the large-range voltage sweep (Fig. 2b). Following the results of ab initio calculations[55,62], the double-positive charge state of oxygen vacancies is only energetically favored over single-charged vacancies if the vacancies are sufficiently apart from each other. As V_O^{2+} vacancies reach a filamentary cluster, they preferably trap charges to occupy a neutral or single-charged state. The single-charged vacancies V_O^+ are immobile in the bulk with activation energies of diffusion of ≈ 2 eV[55] but comparatively mobile within and alongside the filament with much smaller activation energies of 1.05 eV and 0.8 eV, respectively[55]. Vacancies in these charge states can return to the mobile double-charged state (cohesion-isolation transition) by releasing electrons. The transition energy barrier (≈ 1 eV) corresponds to electric field magnitudes, which can be achieved at small applied voltages of a few volts[55,62]. Further, the large current densities expected in the filament cause significant Joule heating and a corresponding increase in the defect mobilities[63]. Via these mechanisms, the filament can grow, and an applied voltage can modulate the electrode-filament gap and the interface potential. From our measurements, the filamentary switching process in the HfO$_{1.8}$ device is also connected with a reversed direction of the hysteresis loop. This reversal could indicate a local change from n-type to p-type conduction[23] because a reduced interface barrier for electrons corresponds to an increased barrier for holes. Consistent with experimental results, a change from n-type to p-type conduction can be induced in HfO$_x$ by increasing the density of oxygen vacancies[64].

Area-type resistive switching is assumed to occur at small vacancy densities, where filament formation is not favored, by the homogeneous drift and formation of V_O^{2+} as illustrated in Fig. 4b. Within this model, the charged point defects alter the interface potential as they drift in the applied electric field, which changes the current through the device. The transition of the measured I-U characteristics (Fig. 2) to a larger hysteresis with a smaller overall resistance can be explained by the voltage-induced formation of additional vacancies at the HfO$_x$/AlO$_x$ interface. The range of the required formation energy of this process can be estimated from the expected density of oxygen vacancies and the voltage range in which it is expected to increase significantly, i.e., the voltage range of the large-range voltage sweep in Fig. 2a. We use a rate equation to describe the formation of electric field-induced vacancies, derived from the Butler-Volmer equation[47]. Further, considering the uncertainty of $\pm 1\%$ in the stoichiometry ("Memristive devices and materials" section) and the density of oxygen sites in ideal monoclinic HfO$_2$, we estimate that the density $[V_{O,s}^{2+}]$ of stoichiometric vacancies $V_{O,s}^{2+}$ in our HfO$_2$ layers must be $[V_{O,s}^{2+}] \leq 3.2 \cdot 10^{23}$ m^{-3} to be consistent with the XPS measurements. Example results of the electric field-induced vacancy density $[V_{O,i}^{2+}]$ as a function of the applied voltage U are plotted in Fig. 4c. Results are shown for three different formation energies E_V together with the maximum of $[V_{O,s}^{2+}]$ indicated by a black dashed line. From the current–voltage measurements in Fig. 2a, a significant increase in the vacancy density is expected for applied voltages between 3 and 5 V, which corresponds to a formation energy of approximately $E_V > 1.4$ eV in Fig. 4c if we assume that the maximum of $\left[V_{O,s}^{2+}\right]$ is present in the HfO$_2$ layer. For smaller formation energies $[V_O^{2+}]$ reaches unrealistically high values at $U = 5$V, e.g., $\left[V_O^{2+}\right] > 10^{28}$ 1/m^2 for $E_V \approx 1.3$ eV. While such high vacancy densities have been used for filamentary regions[65], we consider it too high for area-type switching because it by far exceeds the density of oxygen lattice sites of $\approx 3.2 \cdot 10^{25}$ m^{-3} in ideal monoclinic HfO$_2$. A more precise
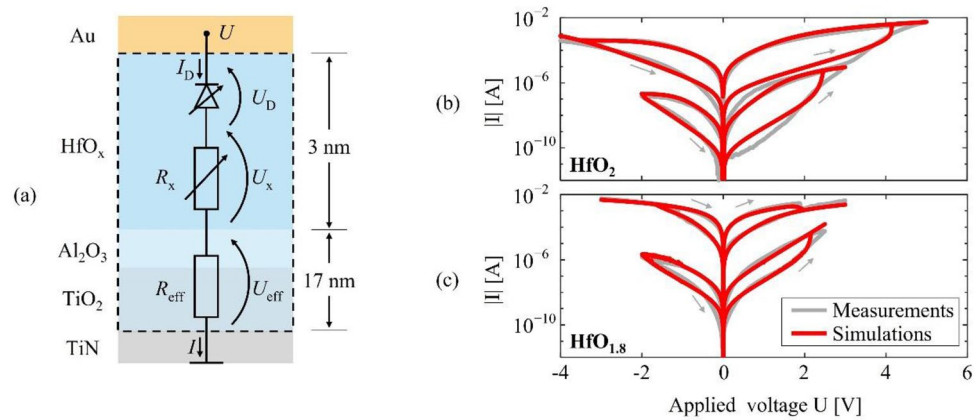
Figure 5. Comparison of simulations with measurements of the current–voltage characteristics. (**a**) Illustration of the equivalent circuit of the physics-based compact model used. (**b**) Comparison of simulations with measurements from Fig. 2 over the small-range and the large-range voltage sweep of the HfO_2 device and (**c**) the $HfO_{1.8}$ device.

estimation of E_V, including an upper limit, requires knowing the range of the vacancy density in the HfO_2 layer. For such an estimation, a model is presented in the next section.

Physics-based compact model.   The following model is based on a simplified version of the mechanisms described in Section C, and it will be used to discuss the physics of the switching processes quantitatively. The equivalent circuit of the model is shown in Fig. 5a.

Within the model, we assume that the electromigration of mobile double-charged oxygen vacancies V_O^{2+} in the hafnium oxide layer dominates the hysteresis in the current–voltage characteristics. Based on this assumption, we represent the contribution of the other layers by an effective resistance R_{eff} and omit mobile ions in the TiO_2 layer. This is justified by its large thickness and correspondingly small electric field compared to the HfO_x layer. For area-type switching, the resistance R_x of the HfO_x layer is approximated to be constant depending only on the average dopant concentration, while for filamentary switching, it is a function of the filament length. The current through the Au/HfO_x interface is represented by a diode current I_D and follows a modified Shockley equation with a barrier potential and ideality factors that depend on the concentration of mobile dopants at the HfO_x/Au interface. The dynamics of V_O^{2+} is modeled using the expression of Mott and Gurney[66] for the ionic current as a function of the electric field. Capacitive effects are neglected in good approximation because of the low frequencies used in all measurements here. On this time scale, also Joule heating can be calculated using the stationary equilibrium solution and an effective thermal resistance, which unifies heat loss.

Despite the approximations made in the model, the simulations fit the measured current–voltage characteristics very well (Fig. 5b,c). Distinct features such as the discontinuities and the reversed direction of the hysteresis in the filament-type measurement are reproduced (Fig. 5c), as well as the nonlinearities and asymmetry of the maximum currents. With ≤ 0.1 eV, the values for the Schottky barrier lowering found in the area-type devices are of a typical order of magnitude reported previously[67]. Also, the obtained activation energies of diffusion (≈ 0.4–0.7 eV) roughly match the theoretical value of 0.69 eV[54,55] for ideal single crystalline HfO_2 but are overall smaller. Such a deviation is expected because of the reduced activation energy of diffusion along the grain boundaries present in our HfO_2 layer. Minor deviations occur mainly around the maximum positive voltage in the simulations of the area-type curves and might be a consequence of the simplified description of the dopant dynamics and of omitting inhomogeneities. Further, additional contributions to the hysteresis from the dynamics of electrons and holes being trapped and released from interface or volume trap states cannot be entirely excluded[24]. Such trap states are introduced by the V_O^{1+} and V_O^{2+}[68–76], and therefore, their contribution to the hysteresis could change with the applied voltage, composition, and microstructure[76,77]. Electron trap states have been analyzed in other Al_2O_3/HfO_2 structures, which were designed as charge-trapping memories with retention times of up to 10^8 s[78–80]. Such a mechanism could contribute to the hysteresis curves in our devices, too.

From the model, we obtain a vacancy density range of approximately $10^{18} - 10^{21}$ m^{-3} for the low-defect area-type devices (HfO_2) and consistently much larger values of $5 \cdot 10^{23} - 6 \cdot 10^{24}$ m^{-3} for the high-defect area-type device ($HfO_{1.8}$). These values match the expected stoichiometric vacancy densities of $[V_{O,s}^{2+}] < 3.2 \cdot 10^{23}$ m^{-3} (HfO_2) and $\left[V_{O,s}^{2+}\right] = 3.2 \cdot 10^{24}$ m^{-3} ($HfO_{1.8}$), and are smaller than the vacancy densities obtained for the filament-type device ($\approx 10^{25} - 10^{27}$ m^{-3}) in good agreement with previous investigations[44,65]. Hence, the results reflect the described microscopic mechanisms very well. Using the range of the vacancy density in the HfO_2 device, we estimate an energy of approximately $E_V \approx 1.65$ eV ± 0.05 eV as indicated in Fig. 4 for the formation of V_O^{2+} at the Al_2O_3/HfO_2 interface. This value is much smaller than the formation energy 5–9 eV[57,58] of Frenkel pairs in the bulk and larger than 0.9 eV[47] for forming Frenkel pairs at TiO_x/HfO_x interfaces of bilayer devices without stable area-type switching. The result is consistent with our assumptions and with the fact that Al_2O_3 is known as a poor oxygen vacancy former[62] with large activation energy for the diffusion of oxygen ions[45,81]. Based on the
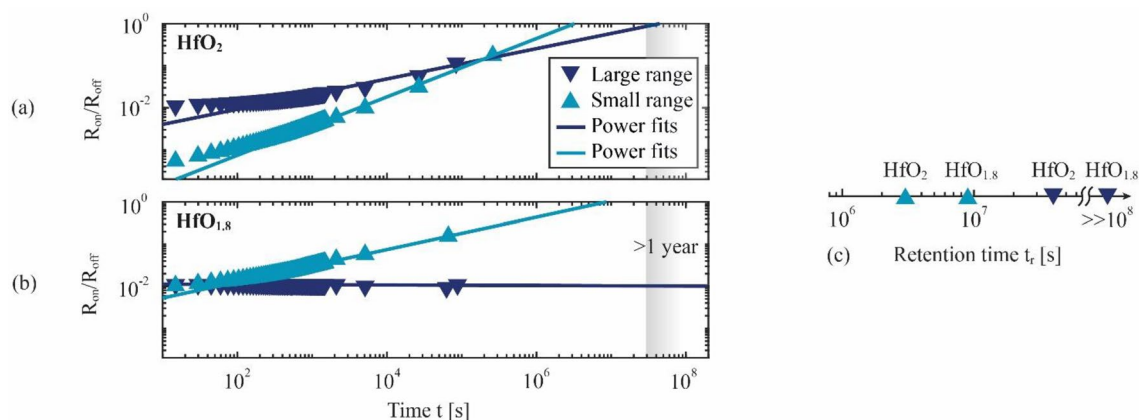
Figure 6. Measured resistance ratios R_{on}/R_{off} as a function of time and power-law fits of the function $R_{on}/R_{off} = \beta t^{\alpha}$ for (**a**) the HfO$_2$ device and (**b**) the HfO$_{1.8}$ device. (**c**) Retention time t_r for the two hafnium oxide compositions and the two voltage ranges, estimated from the values of the power functions at $R_{on}/R_{off} = 1$.

above estimations, the aluminum oxide layer is expected to stabilize the area-type switching process by limiting the formation and dynamics of oxygen vacancies, which is also quantitatively consistent with our model.

Characterization for neuromorphic applications. Two important device properties for neuromorphic applications are the time a device remains in the low resistive state after a set pulse was applied (state retention time) and the linearity of the conductance as a function of the number of applied SET and RESET pulses (linearity in conductance update). In the following, we characterize these two properties for the two different compositions and the two voltage ranges.

The retention characteristics of the LRS can be measured as the resistance R_{on} of the LRS as a function of time t. To obtain $R_{on}(t)$, the respective memristive device is initially set to the LRS of either the large-range or the small-range voltage characteristic (Fig. 2) by applying a rectangular voltage pulse with a correspondingly large or small voltage amplitude. The evolution of $R_{on}(t)$ is probed with a series of reading voltage pulses and normalized to the corresponding resistance R_{off} in the HRS. The resulting on–off ratio $R_{on}(t)/R_{off}$ is evaluated by fitting the Curie-von Schweidler power law $R_{on}/R_{off} = \beta t^{\alpha}$[82] to the measurements with fitting parameters β and α. Besides describing the dynamics of capacitive discharge and charge trapping[82–85], this power law has been used previously to characterize the retention characteristics of double-barrier memristive devices[34] and is used here to enable a quantitative comparison. We define the state retention time t_r via $R_{on}(t = t_r)/R_{off} = 1$ and estimate it by extrapolating the measured data with the power law.

The results of the measurements and fits are presented in Fig. 6a (HfO$_2$) and b (HfO$_{1.8}$), and the calculated retention times t_r are summarized in Fig. 6c. Overall, the fits match very well, and the on–off ratios follow the expected power law. A clear time dependency is visible in all data sets except for the measurement on HfO$_{1.8}$ with a large set voltage amplitude. Within the measurement accuracy R_{on}/R_{off} is constant for this data set over the entire period tracked, which indicates a retention time $t_r \gg 10^8$ s. All other values of t_r are between some 10^6–10^7 s. For both samples, a longer retention time t_r is observed in the measurement series with a large set-voltage amplitude compared to those with a small set-voltage amplitude. More specifically, the retention time can be improved by choosing the corresponding set voltage amplitude from ≈ 30 days to ≈ 460 days (HfO$_2$) and from ≈ 115 days to an estimate of several years (HfO$_{1.8}$).

Such high retention times of several years are often observed in filament-type devices[17,18], while much shorter retention times are reported for area-type devices[24,34]. Consequently, the results from the retention measurements are consistent with the analysis of I-U characteristics in Section B.

Another critical parameter for the application of memristive devices in neuromorphic computing is the linearity in the conductance as a function of the number of applied SET pulses. For linearity analysis, pulse measurements are conducted on the two memristive devices from Fig. 2, starting in the HRS of either the small-range or large-range voltage characteristics. A series of 21 SET voltage pulses and 20 RESET voltage pulses are applied, all rectangular and with a pulse length of 50 ms. A reading voltage pulse follows every SET and RESET pulse to probe the conductance.

The measured conductance G is normalized to its maximum value G_{max} and plotted as a function of the pulse count N_P in Fig. 7a (HfO$_2$ device) and Fig. 7b (HfO$_{1.8}$ device). In the measurements on both devices, a significant difference can be observed depending on the voltage range used for the SET pulses. For the measurements on the HfO$_2$ device, both set voltage ranges (small range and large range) yield qualitatively very similar conductance curves. Here, G/G_{max} increases continuously with the number of SET pulses ($N_P = 1-21$) until the LRS is reached and reduces continuously during the RESET process ($N_P = 22-41$) after an initially large drop in the conductance during the first RESET pulse. The conductance curves of the two voltage modes of the HfO$_2$ device differ mainly quantitatively, with a slightly larger nonlinearity and altered slope in the large-range measurement compared to the small-range measurement.
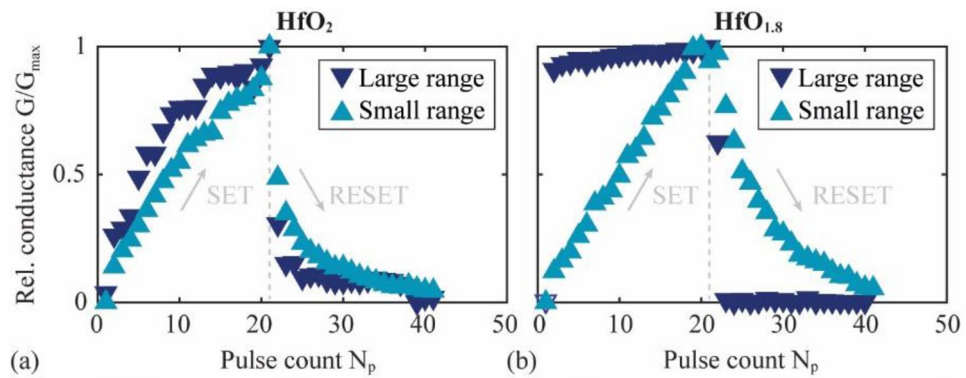
Figure 7. Conductance G/G_{max} normalized to the maximum value G_{max} of the conductance G as a function of the number of applied voltage pulses N_p for (**a**) the HfO$_2$ device and (**b**) the HfO$_{1.8}$ device, and two different amplitude ranges of the applied SET/RESET voltage pulses.

A similar conductance curve is observed in the small-range measurements on the HfO$_{1.8}$ device (Fig. 7b) but with improved linearity and continuity, i.e., without a sudden drop in the conductance after the first RESET pulses. In the large voltage range, entirely different behavior of the conductance update is found. Here, an abrupt transition between the LRS ($G/G_{max} \approx 1$) and HRS ($G/G_{max} \approx 0$) occurs after the first SET and RESET pulse, respectively, and in between, G/G_{max} changes slightly but almost linearly. This behavior of the conductance update curve differs qualitatively from the other three measurement sets, consistent with the current–voltage measurements in Fig. 2b.

Discussion and conclusions

We presented trilayer HfO$_x$/Al$_2$O$_3$/TiO$_2$ memristive devices with two hafnium oxide stoichiometries ($x = 1.8$, $x = 2$) and analyzed their electrical characteristics for neuromorphic applications. We demonstrated the possibility of tuning the switching mechanism by altering the hafnium oxide stoichiometry and showed how the Al$_2$O$_3$ interlayer could extend the range in which area-type switching can be tuned. The Al$_2$O$_3$ interlayer leads to stable area-type switching at small voltages for both hafnium oxide stoichiometries. A change from an area-type switching mechanism to a filament-type is induced at a larger voltage amplitude in the HfO$_{1.8}$ device, with strong indications for a transition from an n-type semiconductor to a p-type semiconductor. The origin of this transition and the switching mechanisms is discussed in detail and used for setting up a physics-based model capable of reproducing distinct features of the measured current–voltage characteristics. The model is based on an equivalent-circuit formulation and can be easily integrated into large-scale simulations for the design of memristive circuits. Quantitatively consistent with our model, we estimated an activation energy of ≈ 1.65 eV ± 0.05 eV for Frenkel defects at the HfO$_x$/Al$_2$O$_3$ interface. This is larger than the values of HfO$_x$/TiO$_x$ interfaces[47] and can explain the stable area-type switching behavior we observe in the trilayer devices. The stabilization leads to various additional advantages in performance compared to previously presented HfO$_x$/TiO$_2$ bilayer devices[67], such as the possibility of operating without current compliance and initial electroforming cycles in devices of both hafnium oxide stoichiometries.

Further, the state retention time and linearity in the conductance update were characterized. Depending on the stoichiometry and voltage range applied, we observed retention times from several days to months in the area-type devices and several years for the filamentary device with on–off ratios of 10^2–10^3 Consistent with the switching type, a binary conductance update is observed in the filamentary device, and an analog-type conductance update with many continuously accessible resistance states in the area-type devices. Compared to the other trilayer device configurations, the area-type HfO$_{1.8}$ devices show much-improved linearity in the conductance update while keeping a large on–off ratio and a long retention time. These properties could make the devices promising for classical artificial neural networks such as deep neural networks and convolutional neural networks. Area-type switching in devices of this stoichiometry, and thereby, such a combination of properties, was essentially made possible by the stabilizing aluminum oxide layer. The nonlinear, time dependent resistance change in the HfO$_2$ devices could be on the other hand suitable for implementing the STDP rule in spiking neural networks.

In conclusion, we demonstrated a technological methodology for tuning multilayer redox-based memristive devices. The insights gained on stabilizing the area-type switching mechanism are an essential step towards tailored devices unifying the beneficial characteristics of filamentary and area-type devices. We achieved promising properties with trilayer memristive elements and provided a model, which can be integrated into large-scale circuit simulations to support the design of memristive circuits for neuromorphic learning schemes in the future.

Methods

Device fabrication and setup. The TiN bottom electrode was deposited by reactive DC magnetron sputtering on top of a SiO$_2$/Si substrate and afterward patterned with photolithography and a subsequent lift-off process. On top of the rear-side electrode, the titanium oxide (TiO$_2$) layer was deposited using a reactive gas
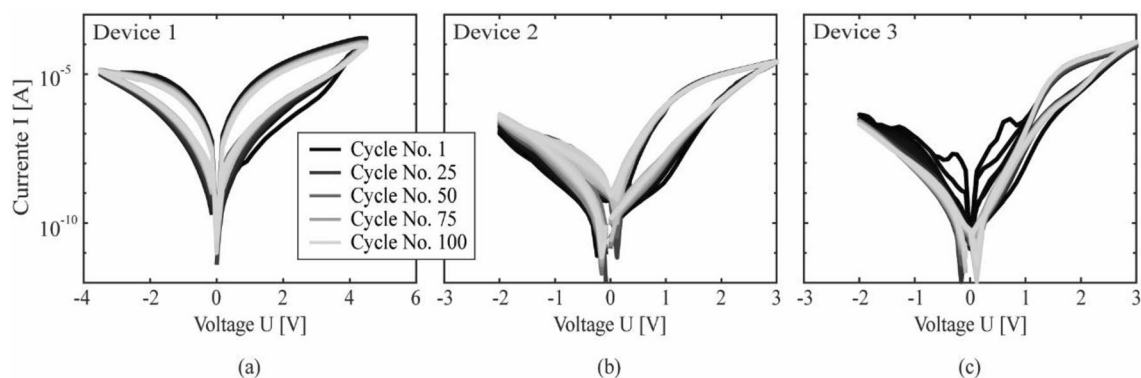
Figure 8. Current–voltage loops measured over a series of 100 periods on three representative example devices; (**a**) device 1: HfO_2, device area 20^2 μm^2, large range, (**b**) device 2: HfO_2, device area 25^2 μm^2, small range, and (**c**) device 3: $HfO_{1.8}$, device area 10^2 μm^2, small range.

mixture of O_2 and Ar with a volume ratio of 10/40. For the aluminum oxide layer (Al_2O_3), a 2-nm-thick Al layer was sputter deposited and subsequently thermally oxidized under an O_2 atmosphere for 30 min. During the oxidization, an of Ar/O_2 gas mixture was used with a volumetric ratio of 30/10 and a working pressure of 0.05 mbar. The stoichiometry of the polycrystalline monoclinic hafnium oxide (HfO_x) layer was tuned to obtain either an atomic ratio of O/Hf of $x \approx 2$ deposited homogeneously on a rotating substrate or an atomic ratio of $x \approx 1.8$ with a static wedge deposition. An O_2/Ar gas mixture with a volume ratio of 10/29 is used during the deposition process. The Au contact layer was electron-beam evaporated on the top of the functional trilayer and patterned via photolithography and a lift-off process. The devices were encapsulated by a 120-nm-thick SiO_2 layer before the Al layer was evaporated and patterned to create the conduction lines. Details on the XPS measurements and the fabrication process are provided in Ref.[67]. The variation of electrical properties of reactively sputtered metal oxides has been investigated elsewhere[86].

Current–voltage characteristics. All electrical measurements were conducted with a source-measurement unit (Keysight B2901A). During the measurement, a voltage U is applied to the top electrode and swept piecewise linearly between the start and end values ($U = 0V$) and the respective maximum and minimum values. The maximum and minimum voltages are set to values just below the expected electrical breakdown voltage, which was determined in advance by measurements on similar devices. During the measurements, the rear-side electrode was grounded. The device-to-device variability was tested on several devices by probing the current over a series of 100 voltage cycles. Representative example measurements on three different devices are shown in Fig. 8.

XRD measurements. X-ray diffraction (XRD) spectra were measured of two samples comprising different hafnium oxide layers on a Si substrate. The hafnium oxide layer was sputter deposited for 900 s on a rotating substrate to achieve a homogeneous HfO_2 layer and in a static wedge-deposition mode to obtain a gradient HfO_x layer. From the two 4-in. test wafers, samples were selected from the same position in the center of the wafer, where a similar thickness of ≈ 30 nm of hafnium oxide is expected and a composition of $HfO_{1.8}$ of the wedged layer. The XRD system (Bruker D8 Discover, Bruker Corporation, USA) consists of a copper Kα radiation source and a Lynxeye XE-T detector with an energy resolution < 380 eV at 8 keV. The measured data are filtered to remove Kα-contributions, the background is subtracted, and they are normalized to the Si (100) peak at $2\theta = 69°$.

Area dependency. For the analysis of the area dependency in Fig. 3, the resistance values were extracted from the I-U curves at a voltage of 1 V and only for the large-range measurements on the $HfO_{1.8}$ device at 0.1 V. The measurements were performed on many devices and then averaged. An overview of the number of devices used for calculating the mean values and standard deviations is provided in Table 1 and the standard deviations are summarized in Table 2. Owing to the fabrication process, a significantly smaller number of devices with an $HfO_{1.8}$ layer was available compared to the number of devices with an HfO_2 layer.

Pulse measurements. For all reading voltage pulses, amplitudes of 1.2 V (small range) and 0.1 V (large range) are used for the $HfO_{1.8}$ device, and 1.2 V in both voltage ranges for the HfO_2 device. For the linearity of conductance measurements (Fig. 7), SET/RESET voltage pulses with amplitudes of 2.5 V/− 1.5 V (small range) and 3.5 V/− 3 V (large range) are applied to the HfO_2 device and 2.5 V/− 1.5 V (small range) and − 3 V/3 V (large range) to the $HfO_{1.8}$ device. We used pulse lengths of 50 ms with a frequency of 150 ms. For the retention test, we used amplitudes of the initial SET pulses of 3.5 V (small range) and 5 V (large range) for the HfO_2 device and 3 V (small range) and − 2 V (large range) for the $HfO_{1.8}$ device. The pulse lengths were 10 ms (small range) and 50 ms (large range) for the HfO_2 device and 10 ms for the $HfO_{1.8}$ device in both voltage ranges.

Device area A (μm^2)	HfO$_2$		HfO$_{1.8}$	
	Small range	Large range	Small range	Large range
15^2	79	127	3	3
20^2	108	145	4	5
25^2	208	181	4	1
35^2	316	179	4	6
50^2	213	153	6	8

Table 1. Overview of the number of devices used to calculate the mean values in Fig. 3 and the standard deviations in Table 2.

Device area A (μm^2)	HfO$_2$				HfO$_{1.8}$			
	Small range		Large range		Small range		Large range	
	LRS	HRS	LRS	HRS	LRS	HRS	LRS	HRS
15^2	32	29	27	28	16	25	66	15
20^2	35	36	16	14	20	14	27	6
25^2	36	34	6	6	12	31	–	
35^2	34	39	1	1	9	4	96	17
50^2	36	36	5	5	10	6	60	19

Table 2. Relative standard deviation of the resistance measurements in Fig. 3 in percent.

Fit parameter	HfO$_2$		HfO$_{1.8}$	
	Small range	Large range	Small range	Large range
α	$3.023 \cdot 10^{-5}$	$1.754 \cdot 10^{-3}$	$2.176 \cdot 10^{-3}$	$1.142 \cdot 10^{-2}$
β	$6.944 \cdot 10^{-1}$	$3.603 \cdot 10^{-1}$	$3.839 \cdot 10^{-1}$	$-6.427 \cdot 10^{-3}$

Table 3. Values of the fit parameters for the function $R_{on}/R_{off} = \beta t^\alpha$, plotted in Fig. 6.

Retention fits. The values of the parameters α and β, which result from the fit of the power function $R_{on}/R_{off} = \beta t^\alpha$ to the measured values of R_{on}/R_{off} as a function of time t in Fig. 6 are provided in the following Table 3. For calculating the resistance ratios R_{on}/R_{off}, the on-state resistances R_{on} are normalized to the measured off-state resistances $R_{off} = 290$ MΩ (HfO$_2$, small range), $R_{off} = 0.7$ MΩ (HfO$_2$, large range) and $R_{off} = 80$ MΩ (HfO$_{1.8}$, small range), $R_{off} = 0.16$ MΩ (HfO$_{1.8}$, large range).

Vacancy densities in area-type devices. The density of oxygen sites $n_{max} = N_O/V_c$ of ideal monoclinic HfO$_2$ is estimated from the number $N_O = 4$ oxygen ions within one unit cell with volume $V_c \approx (5 \text{ nm})^2$ to $n_{max} = 3.2 \cdot 10^{25}$ m^{-3}. Considering the fractions of $< 1\%$ and 10% of occupied oxygen sites from the stoichiometry measured via XPS, we estimate for HfO$_2$ a density of stoichiometric vacancies of $[V_{O,s}] < 3.2 \cdot 10^{23}$ m^{-3} and for HfO$_{1.8}$ a density of $[V_{O,s}] = 3.2 \cdot 10^{24}$ m^{-3}. These vacancies can occupy various stable charged states, V_O^0, V_O^{1+} and V_O^{2+}, following the reactions $V_O^0 \rightarrow V_O^{1+} + e^-$ and $V_O^{1+} \rightarrow V_O^{2+} + e^-$. The corresponding densities $[V_O^0], [V_O^{1+}], [V_O^{2+}]$ can be calculated considering their ionization energies relative to the conduction band edge[65]. With activation energies of $E_O^{0\rightarrow1} \approx 0.63$ eV for the neutral-single charge transition[57], and $E_O^{1\rightarrow2} = 0.8$ eV for the single charge-double charge transition[57], we obtain at $U = 0$V concentrations of $[V_O^{+2}] \approx [V_O]$ for HfO$_2$ and HfO$_{1.8}$. For estimating the change of the electric-field induced vacancy density as a function of time, we consider the following rate equation[47]

$$\frac{d[V_O^{2+}]}{dt} = v_0 N_{0,s} \exp\left(-\frac{qE_V - 0.5zaqE}{k_B T}\right),$$
(1)

with the phonon frequency v_0, the charge number $z = +2$ and hopping distance $a \approx 0.5$ nm of oxygen vacancies, the positive elementary charge q, the electric field E across the HfO$_x$ layer, the Boltzmann constant k_B, the temperature $T = 300$ K, and the effective formation energy E_V of V_O^{2+}. The electric field E is obtained from the voltages solved for in the compact model. Because the reaction occurs at the HfO$_x$/AlO$_x$ interface and not in the entire volume, we use an effective volume density $N_{0,s} = n_{max} \cdot t_s/t_{HfO} \approx 5 \cdot 10^{24}$ m^{-3} of oxygen sites, scaling the ideal density of oxygen sites n_{max} by the ratio of surface thickness t_s to film thickness t_{HfO}. Further we assume that $t_s \approx a$, the hopping distance of the vacancies. Details on the parameters are provided in "Simulation parameters".

Physics-based model. In the following, all equations used for the compact model are presented and explained. The contributions of the Al_2O_3 and the TiO_2 layer to the current–voltage characteristics are considered by one effective resistance R_{eff} with a corresponding current I_{eff}. This current depends on the voltage amplitude U_{eff} over these layers and is scaled by the fitting parameter $I_{eff,0}$[67]

$$I_{eff} := I_{eff,0}\sinh(U_{eff}). \tag{2}$$

Within the hafnium oxide layer, we consider the drift of mobile point defects and define a normalized concentration x_r ($0 \leq x_r \leq 1$) via the maximum concentration x_{max} and the minimum concentration x_{min} of mobile dopants at the Au/HfO$_x$ interface

$$x_r := \frac{x - x_{min}}{x_{max} - x_{min}} \text{ with } x_{min} \leq x \leq x_{max}. \tag{3}$$

The average dopant concentration is then given by

$$\overline{x} := \frac{1}{2}(x_{max} + x_{min}). \tag{4}$$

The change of dopant concentration x with time t at the Au/HfO$_x$ interface is approximated using the (net) ionic current I_x, the charge number z of the drifting dopants, the cross-section area A_c of the conducting channel, and the channel length, which equals the thickness t_{HfO} of the HfO$_x$ layer[87]

$$\frac{\partial x}{\partial t} := \frac{I_x}{zqA_c t_{HfO}} \text{ with } I_x = zqA_c \overline{x}\overline{v}. \tag{5}$$

For the area-type case, the cross-section area A_c is set to the active device area $A_c = A$, while for filamentary conduction, it is set to the cross-section area A_f of the filament $A_c = A_f = \pi r_f^2$ with the filament radius r_f. The drift velocity $\overline{v}$ has been formulated by Mott and Gurney[66] as a function of the applied electric field E across the HfO$_x$ layer

$$\overline{v} = v_0 a \exp\left(\frac{\Delta W}{k_B T}\right) 2\sinh\left(\frac{zqaE}{2k_B T}\right), \text{ with } E = \frac{U_x}{t_{HfO}}, \tag{6}$$

given by the voltage U_x across the HfO$_x$ layer with a thickness t_{HfO}. In this equation, v_0 is the attempt frequency (phonon frequency) for overcoming the potential barrier ΔW, which separates the local minima. The distance between the local minima is given by the hopping distance a. The Boltzmann constant is denoted by k_B and the temperature by T. The electric resistance of the HfO$_x$ layer is approximated to be independent of x for area-type conduction

$$R_x = \frac{t_{HfO}}{|z|q\mu_n A\overline{x}}, \tag{7}$$

and linearly dependent on x for filamentary-type conduction

$$R_x(x) = R^{LRS}x_r + R^{HRS}(1 - x_r), \tag{8}$$

with the electric resistance R^{LRS} of the low resistance state, and R^{HRS} of the high resistance state

$$R^{LRS} = \frac{t_{HfO}}{|z|q\mu_p A_f x_{max}}, R^{HRS} = \frac{t_{HfO}}{|z|q\mu_p A_f x_{min}}. \tag{9}$$

Here, the electron and hole mobilities are denoted as μ_n and μ_p, respectively. For a phenomenological description of the current through the Schottky contact at the Au/HfO$_x$ interface, a Schockley-like expression is used[88]

$$I_D = \begin{cases} I_{D,0}\left(\exp\left(\frac{qU_D}{n(x)k_B T}\right) - 1\right) & \text{for } U - U_D \geq 0 \\ -I_{D,0}\left(\exp\left(-\frac{\alpha_D qU_D}{k_B T}\right) - 1\right) & \text{for } U - U_D < 0 \end{cases}, \tag{10}$$

with the voltage U_D across the interface, the applied voltage U, the ideality factor $n(x)$, and a fitting factor α_D included to describe the reversed current. The saturation current $I_{D,0}$ is described by[88]

$$I_{D,0} = A^* A_c T^2 \exp\left(-\frac{q\varphi_D(x)}{k_B T}\right), \tag{11}$$

with the effective Richardson constant $A^* = 1.202 \cdot 10^6 A/(m^2 K^2)$. The effective Schottky barrier $\varphi_D(x)$ and the ideality factor $n(x)$ are described as functions of the concentration x of mobile dopants at the Au/HfO$_x$ interface to cover various effects, which can cause a change in the barrier height. In linear approximation, it is

$$\varphi_D(x) = \varphi_D^{HRS}x_r + \varphi_D^{LRS}(1 - x_r),$$
$$n(x) = n^{HRS}x_r + n^{LRS}(1 - x_r). \tag{12}$$

	HfO$_{1.8}$		HfO$_2$	
Parameter	Small range	Large range	Small range	Large range
ϕ_B^{HRS} [eV]	0.65	0.31	0.73	0.51
ϕ_B^{LRS} [eV]	0.58	0.34	0.63	0.41
n^{HRS}	12	7	12	24
n^{LRS}	11	3	9	15
α_D	0.080	0.100	0.100	0.055
$I_{el,O}$ [A/m^2]	$1 \cdot 10^{12}$	$2 \cdot 10^{14}$	$1 \cdot 10^{12}$	$1 \cdot 10^{12}$
ΔW [eV]	0.385	0.84	0.67	0.72
x_{min} [1/m^3]	$5 \cdot 10^{23}$	$1 \cdot 10^{25}$	$1 \cdot 10^{18}$	$5 \cdot 10^{19}$
x_{max} [1/m^3]	$6 \cdot 10^{24}$	$1.5 \cdot 10^{27}$	$1 \cdot 10^{19}$	$3 \cdot 10^{21}$

Table 4. Device-dependent fitting parameters used for the simulations in Fig. 5. All parameters are defined in "Physics-based model", where the model is described.

Parameter	Value	Parameter	Value
ν_0 [Hz]	10^{-12}[47]	μ_n [m^2/(Vs)]	10^{-5}[67]
t_{HfO} [nm]	3	μ_p [m^2/(Vs)]	10^{-4}[64]
a [nm]	0.5	r_f [nm]	50
$R_{T,eff}$ [W/K]	$2.5 \cdot 10^4$	A [μm^2]	35^2
T_0 [K]	300	z	2

Table 5. Model parameters used for the simulations in Fig. 5.

Here, φ_D^{HRS} and φ_D^{LRS} are the effective Schottky barriers in the high resistive state and low resistive state, respectively, and n^{HRS} and n^{LRS} the corresponding ideality factors. Considering Joule heating, we calculate the quasi-static equilibrium temperature T from

$$T = T_0 + \frac{UI}{A_s h_{eff}} := T_0 + \frac{UI}{R_{T,eff}}, \tag{13}$$

with the environmental temperature $T_0 = 300$ K, the current $I = I_D$ through the device, the surface area A_s of the device through which heat is lost, and the effective heat-transfer coefficient h_{eff}, which are unified in an effective thermal resistance $R_{T,eff}$. Applying Kirchhoff's circuit laws to the equivalent circuit in Fig. 5a, together with the above expressions, a system of equations is set up and solved numerically for the unknown voltages U_D, U_x, and U_{eff}, as functions of the applied voltage.

Simulation parameters. The parameters used for the simulations in Fig. 5 are summarized in Tables 4 and 5. The effective hopping distance $a \approx 0.5$ nm is estimated from the lattice constant[89] and is similar to the length of the lowest energy hopping pathway of V_O^{2+}[54]. For the filament radius r_f, a value is used consistent with other studies on filamentary devices[47,65]. The value of the effective thermal resistance $R_{T,eff}$ is adjusted to reach a maximum temperature of 550 K[63] for the case of filamentary switching.

Data availability
The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

References
1. Sung, C., Hwang, H. & Yoo, I. K. Perspective: A review on memristive hardware for neuromorphic computation. *J. Appl. Phys.* **124**, 151903 (2018).
2. Serrano-Gotarredona, T., Masquelier, T., Prodromakis, T., Indiveri, G. & Linares-Barranco, B. STDP and STDP variations with memristors for spiking neuromorphic learning systems. *Front. Neurosci.* **7**, 2 (2013).
3. Ielmini, D. & Waser, R. (eds) *Resistive switching. From fundamentals of nanoionic redox processes to memristive device applications* (Wiley-VCH, 2016).
4. Zhao, M., Gao, B., Tang, J., Qian, H. & Wu, H. Reliability of analog resistive switching memory for neuromorphic computing. *Appl. Phys. Rev.* **7**, 11301 (2020).
5. Zhu, J., Zhang, T., Yang, Y. & Huang, R. A comprehensive review on emerging artificial neuromorphic devices. *Appl. Phys. Rev.* **7**, 11312 (2020).

6. Potok, T. E. *et al.* A study of complex deep learning networks on high-performance, neuromorphic, and quantum computers. *J. Emerg. Technol. Comput. Syst.* **14**, 1–21 (2018).

7. Yang, J. J., Strukov, D. B. & Stewart, D. R. Memristive devices for computing. *Nat. Nanotechnol.* **8**, 13–24 (2013).

8. Zhu, G. *et al.* Study on high-density integration resistive random access memory array from multiphysics perspective by parallel computing. *IEEE Trans. Electron Devices* **66**, 1747–1753 (2019).

9. Choi, S., Yang, J. & Wang, G. Emerging memristive artificial synapses and neurons for energy-efficient neuromorphic Computing. *Adv. Mater.* **32**, e2004659 (2020).

10. Ziegler, M. & Kohlstedt, H. *Memristor Computing Systems* 247–272 (Springer, 2022).

11. Chandrasekaran, S., Simanjuntak, F. M., Saminathan, R., Panda, D. & Tseng, T.-Y. Improving linearity by introducing Al in HfO_2 as a memristor synapse device. *Nanotechnology* **30**, 445205 (2019).

12. Solan, E. *et al.* An enhanced lumped element electrical model of a double barrier memristive device. *J. Phys. D Appl. Phys.* **50**, 195102 (2017).

13. Dirkmann, S., Hansen, M., Ziegler, M., Kohlstedt, H. & Mussenbrock, T. The role of ion transport phenomena in memristive double barrier devices. *Sci. Rep.* **6**, 35686 (2016).

14. Bagdzevicius, S., Maas, K., Boudard, M. & Burriel, M. Interface-type resistive switching in perovskite materials. *J. Electroceramics* **39**, 157–184 (2017).

15. Waser, R. *et al.* In *Resistive Switching* (eds Ielmini, D. & Waser, R.) 1–30 (Wiley-VCH, 2016).

16. Kim, S., Kim, T.-H., Kim, H. & Park, B.-G. Current suppressed self-compliance characteristics of oxygen rich TiO_y inserted Al_2O_3/TiO_x based RRAM. *APL Mater.* **117**, 202106 (2020).

17. Ninomiya, T. *et al.* Improvement of data retention during long-term use by suppressing conductive filament expansion in TaO_x bipolar-ReRAM. *IEEE Electron Device Lett.* **34**, 762–764 (2013).

18. Ninomiya, T. *et al.* Conductive filament scaling of TaO_x bipolar ReRAM for improving data retention under low operation current. *IEEE Trans. Electron. Devices* **60**, 1384–1389 (2013).

19. Choi, B. J. *et al.* High-speed and low-energy nitride memristors. *Adv. Funct. Mater.* **26**, 5290–5296 (2016).

20. Menzel, S., von Witzleben, M., Havel, V. & Böttger, U. The ultimate switching speed limit of redox-based resistive switching devices. *Faraday Discuss.* **213**, 197–213 (2019).

21. Lee, M.-J. *et al.* 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**, 625–630 (2011).

22. Zhu, Y. B., Zheng, K., Wu, X. & Ang, L. K. Enhanced stability of filament-type resistive switching by interface engineering. *Sci. Rep.* **7**, 43664 (2017).

23. Sawa, A. Resistive switching in transition metal oxides. *Mater. Today* **11**, 28–36 (2008).

24. Sawa, A. & Meyer, R. In *Resistive Switching* (Ielmini, D. & Waser, R.) 457–482 (Wiley-VCH, 2016).

25. Park, C., Seo, Y., Jung, J. & Kim, D.-W. Electrode-dependent electrical properties of metal/Nb-doped $SrTiO_3$ junctions. *J. Appl. Phys.* **103**, 54106 (2008).

26. Muenstermann, R., Menke, T., Dittmann, R. & Waser, R. Coexistence of filamentary and homogeneous resistive switching in Fe-doped $SrTiO_3$ thin-film memristive devices. *Adv. Mater.* **22**, 4819–4822 (2010).

27. Nian, Y. B., Strozier, J., Wu, N. J., Chen, X. & Ignatiev, A. Evidence for an oxygen diffusion model for the electric pulse induced resistance change effect in transition-metal oxides. *Phys. Rev. Lett.* **98**, 146403 (2007).

28. Seong, D., Jo, M., Lee, D. & Hwang, H. HPHA effect on reversible resistive switching of Pt/Nb-doped $SrTiO_3$ Schottky junction for nonvolatile memory application. *Electrochem. Solid-State Lett.* **10**, H168–H170 (2007).

29. Jeon, S. H., Park, B. H., Lee, J., Lee, B. & Han, S. First-principles modeling of resistance switching in perovskite oxide material. *Appl. Phys. Lett.* **89**, 42904 (2006).

30. Baikalov, A. *et al.* Field-driven hysteretic and reversible resistive switch at the Ag–$Pr_{0.7}Ca_{0.3}MnO_3$ interface. *Appl. Phys. Lett.* **83**, 957–959 (2003).

31. Fujii, T. *et al.* Hysteretic current–voltage characteristics and resistance switching at an epitaxial oxide Schottky junction $SrRuO_3$/$SrTi_{0.99}Nb_{0.01}O_3$. *Appl. Phys. Lett.* **86**, 12107 (2005).

32. Sawa, A., Fujii, T., Kawasaki, M. & Tokura, Y. Hysteretic current–voltage characteristics and resistance switching at a rectifying Ti/$Pr_{0.7}Ca_{0.3}MnO_3$ interface. *Appl. Phys. Lett.* **85**, 4073–4075 (2004).

33. Hansen, M., Zahari, F., Ziegler, M. & Kohlstedt, H. Double-barrier memristive devices for unsupervised learning and pattern recognition. *Front. Neurosci.* **11**, 91 (2017).

34. Hansen, M. *et al.* A double barrier memristive device. *Sci. Rep.* **5**, 13753 (2015).

35. Bousoulas, P. *et al.* Engineering amorphous-crystalline interfaces in TiO_{2-x}/TiO_{2-y}-based bilayer structures for enhanced resistive switching and synaptic properties. *J. Appl. Phys.* **120**, 154501 (2016).

36. Lee, J. H., Park, J. H., Dongale, T. D. & Kim, T. G. Vacancy-modulated self-rectifying characteristics of NiO_x/Al_2O_3-based nanoscale ReRAM devices. *J. Alloys Compd.* **821**, 153247 (2020).

37. Jeon, K. *et al.* Self-rectifying resistive memory in passive crossbar arrays. *Nat. Commun.* **12**, 2968 (2021).

38. Cao, G. *et al.* 2D material based synaptic devices for neuromorphic computing. *Adv. Funct. Mater.* **31**, 2005443 (2021).

39. van de Burgt, Y., Melianas, A., Keene, S. T., Malliaras, G. & Salleo, A. Organic electronics for neuromorphic computing. *Nat. Electron.* **1**, 386–397 (2018).

40. Boybat, I. *et al.* Neuromorphic computing with multi-memristive synapses. *Nat. Commun.* **9**, 2514 (2018).

41. Li, J. *et al.* Tuning analog resistive switching and plasticity in bilayer transition metal oxide based memristive synapses. *RSC Adv.* **7**, 43132–43140 (2017).

42. Guo, T. *et al.* From memristive materials to neural networks. *ACS Appl. Mater. Interfaces* **12**, 54243–54265 (2020).

43. Lee, H. Y. *et al.* *Electron Devices Meeting, 2008. IEDM 2008. IEEE International* 1–4 (IEEE, 2008).

44. Hardtdegen, A. *et al.* Improved switching stability and the effect of an internal series resistor in HfO_2/TiO_x bilayer ReRAM Cells. *IEEE Trans. Electron Devices* **65**, 3229–3236 (2018).

45. Li, J. *et al.* *2018 China Semiconductor Technology International Conference (CSTIC)* 1–3 (IEEE, 2018).

46. Stathopoulos, S. *et al.* Multibit memory operation of metal-oxide bi-layer memristors. *Sci. Rep.* **7**, 17532 (2017).

47. Dirkmann, S., Kaiser, J., Wenger, C. & Mussenbrock, T. Filament growth and resistive switching in hafnium oxide memristive devices. *ACS Appl. Mater. Interfaces* **10**, 14857–14868 (2018).

48. Cüppers, F. *et al.* Exploiting the switching dynamics of HfO_2-based ReRAM devices for reliable analog memristive behavior. *APL Mater.* **7**, 91105 (2019).

49. Balogh-Michels, Z. *et al.* Crystallization behavior of ion beam sputtered HfO_2 thin films and its effect on the laser-induced damage threshold. *J. Eur. Opt. Soc.-Rapid Publ.* **17**, 1–8 (2021).

50. Modreanu, M. *et al.* Investigation of thermal annealing effects on microstructural and optical properties of HfO_2 thin films. *Appl. Surf. Sci.* **253**, 328–334 (2006).

51. Li, F. M. *et al.* High-density remote plasma sputtering of high-dielectric-constant amorphous hafnium oxide films. *Phys. Status Solidi B* **250**, 957–967 (2013).

52. Kim, H., McIntyre, P. C. & Saraswat, K. C. Effects of crystallization on the electrical properties of ultrathin HfO_2 dielectrics grown by atomic layer deposition. *Appl. Phys. Lett.* **82**, 106–108 (2003).

53. Waser, R., Dittmann, R., Staikov, G. & Szot, K. Redox-based resistive switching memories—nanoionic mechanisms, prospects, and challenges. *Adv. Mater.* **21**, 2632–2663 (2009).
54. Capron, N., Broqvist, P. & Pasquarello, A. Migration of oxygen vacancy in HfO_2 and across the HfO_2/SiO_2 interface: A first-principles investigation. *Appl. Phys. Lett.* **91**, 192905 (2007).
55. Duncan, D., Magyari-Kope, B. & Nishi, Y. Filament-induced anisotropic oxygen vacancy diffusion and charge trapping effects in hafnium oxide RRAM. *IEEE Electron Device Lett.* **37**, 400–403 (2016).
56. Foster, A. S., Shluger, A. L. & Nieminen, R. M. Mechanism of interstitial oxygen diffusion in hafnia. *Phys. Rev. Lett.* **89**, 225901 (2002).
57. Foster, A. S., Lopez Gejo, F., Shluger, A. L. & Nieminen, R. M. Vacancy and interstitial defects in hafnia. *Phys. Rev. B* **65**, 174117 (2002).
58. Schie, M., Menzel, S., Robertson, J., Waser, R. & de Souza, R. A. Field-enhanced route to generating anti-Frenkel pairs in HfO_2. *Phys. Rev. Mater.* **2**, 35002 (2018).
59. Guo, Y. & Robertson, J. Materials selection for oxide-based resistive random access memories. *Appl. Phys. Lett.* **105**, 223516 (2014).
60. O'Hara, A., Bersuker, G. & Demkov, A. A. Assessing hafnium on hafnia as an oxygen getter. *J. Appl. Phys.* **115**, 183703 (2014).
61. McKenna, K. P. Optimal stoichiometry for nucleation and growth of conductive filaments in HfO_x. *Model. Simul. Mater. Sci. Eng.* **22**, 25001 (2014).
62. Kamiya, K. *et al.* Generalized mechanism of the resistance switching in binary-oxide-based resistive random-access memories. *Phys. Rev. B* **87**, 155201 (2013).
63. Yalon, E. *et al.* Thermometry of filamentary RRAM devices. *IEEE Trans. Electron Devices* **62**, 2972–2977 (2015).
64. Hildebrandt, E. *et al.* Controlled oxygen vacancy induced *p*-type conductivity in HfO_{2-x} thin films. *Appl. Phys. Lett.* **99**, 112902 (2011).
65. Marchewka, A. *et al.* Nanoionic resistive switching memories: On the physical nature of the dynamic reset process. *Adv. Electron. Mater.* **2**, 1500233 (2016).
66. Mott, N. F. & Gurney, R. W. *Electronic Processes in Ionic Crystals* 2nd edn. (Oxford at the Clarendon Press, 1950).
67. Park, S. *et al.* Engineering method for tailoring electrical characteristics in $TiN/TiO_x/HfO_x/Au$ bi-layer oxide memristive devices. *Front. Nanotechnol.* **3**, 29 (2021).
68. Gavartin, J. L., Muñoz Ramo, D., Shluger, A. L., Bersuker, G. & Lee, B. H. Negative oxygen vacancies in HfO_2 as charge traps in high-*k* stacks. *Appl. Phys. Lett.* **89**, 82908 (2006).
69. Perevalov, T. V., Ivanov, M. V. & Gritsenko, V. A. Electronic and optical properties of hafnia polymorphs. *Microelectron. Eng.* **88**, 1475–1477 (2011).
70. Ribes, G. *et al.* Origin of Vt instabilities in high-*k* dielectrics Jahn-Teller effect or oxygen vacancies. *IEEE Trans. Device Mater. Relib.* **6**, 132–135 (2006).
71. Bersuker, G. *et al.* In *IEEE International Reliability Physics Symposium Proceedings, 2006. 44th Annual* 179–183 (IEEE, 2006).
72. Perevalov, T. V., Aliev, V., Gritsenko, V. A., Saraev, A. A. & Kaichev, V. V. Electronic structure of oxygen vacancies in hafnium oxide. *Microelectron. Eng.* **109**, 21–23 (2013).
73. Torii, K. *et al.* In *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004* 129–132 (2004).
74. Lyons, J. L., Janotti, A. & van de Walle, C. G. The role of oxygen-related defects and hydrogen impurities in HfO_2 and ZrO_2. *Microelectron. Eng.* **88**, 1452–1456 (2011).
75. Bersuker, G. *et al.* Metal oxide resistive memory switching mechanism based on conductive filament properties. *J. Appl. Phys.* **110**, 124518 (2011).
76. Gritsenko, V. A., Perevalov, T. V. & Islamov, D. R. Electronic properties of hafnium oxide: A contribution from defects and traps. *Phys. Rep.* **613**, 1–20 (2016).
77. Bousoulas, P., Michelakaki, I., Skotadis, E., Tsigkourakos, M. & Tsoukalas, D. Low-power forming free $TiO_{2-x}/HfO_{2-y}/TiO_{2-x}$-trilayer RRAM devices exhibiting synaptic property characteristics. *IEEE Trans. Electron Devices* **64**, 3151–3158 (2017).
78. Salomone, L. S. *et al.* Deep electron traps in HfO_2-based metal-oxide-semiconductor capacitors. *Thin Solid Films* **600**, 36–42 (2016).
79. Maikap, S. *et al.* Charge trapping characteristics of atomic-layer-deposited HfO_2 films with Al_2O_3 as a blocking oxide for high-density non-volatile memory device applications. *Semicond. Sci. Technol.* **22**, 884–889 (2007).
80. You, H.-W. & Cho, W.-J. Charge trapping properties of the HfO_2 layer with various thicknesses for charge trap flash memory applications. *Appl. Phys. Lett.* **96**, 93506 (2010).
81. Heuer, A. H. Oxygen and aluminum diffusion in α-Al_2O_3: How much do we really understand?. *J. Eur. Ceram. Soc.* **28**, 1495–1507 (2008).
82. Westerlund, S. & Ekstam, L. Capacitor theory. *IEEE Trans. Dielectr. Electr. Insul.* **1**, 826–839 (1994).
83. Zafar, S., Callegari, A., Gusev, E. & Fischetti, M. V. Charge trapping related threshold voltage instabilities in high permittivity gate dielectric stacks. *J. Appl. Phys.* **93**, 9298–9303 (2003).
84. Zhang, H. J., Zhang, X. P., Shi, J. P., Tian, H. F. & Zhao, Y. G. Effect of oxygen content and superconductivity on the nonvolatile resistive switching in $YBa_2Cu_3O_{6+x}$/Nb-doped $SrTiO_3$ heterojunctions. *Appl. Phys. Lett.* **94**, 92111 (2009).
85. Mikheev, E., Hoskins, B. D., Strukov, D. B. & Stemmer, S. Resistive switching and its suppression in Pt/Nb:$SrTiO_3$ junctions. *Nat. Commun.* **5**, 1–8 (2014).
86. Zahari, F. *et al.* Correlation between sputter deposition parameters and I-V characteristics in double-barrier memristive devices. *J. Vacuum Sci. Technol. B: Nanotechnol. Microelectron. Mater. Process. Meas. Phenom.* **37**, 61203 (2019).
87. La Torre, C., Zurhelle, A. F., Breuer, T., Waser, R. & Menzel, S. Compact modeling of complementary switching in oxide-based ReRAM devices. *IEEE Trans. Electron Devices* **66**, 1268–1275 (2019).
88. Sze, S. M. & Ng, K. K. *Physics of Semiconductor Devices* (Wiley, 2006).
89. Laudadio, E., Stipa, P., Pierantoni, L. & Mencarelli, D. Phase properties of different HfO_2 polymorphs: A DFT-based study. *Curr. Comput.-Aided Drug Des.* **12**, 90 (2022).

Acknowledgements

Author contributions

S.P., M.Z. and T.I. performed and designed the experiments. B.S. and M.Z. developed the model and performed the simulations. S.P., M.Z. and B.S. analyzed the results. B.S. and S.P. wrote the manuscript with contributions from all authors. All authors reviewed the manuscript.

Funding

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to B.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.