TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY-NC-ND

Citation (APA)
Hua, E., Spyrou, T., Ahmadi, M., Xun, H., Gebregiorgis, A., Gaydadjiev, G., Hamdioui, S., Ishihara, R., Abunahla, H., & More Authors (2026). PdNeuRAM: forming-free, multi-bit Pd/HfO2 ReRAM for energy-efficient neuromorphic computing. *Communications Engineering*, *5*(1), Article 97. https://doi.org/10.1038/s44172-026-00650-3

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

communications engineering

A Nature Portfolio journal

Article

https://doi.org/10.1038/s44172-026-00650-3

PdNeuRAM: forming-free, multi-bit Pd/ HfO_2 ReRAM for energy-efficient neuromorphic computing

Check for updates

Erbing Hua[1], Theofilos Spyrou[1], Majid Ahmadi[2,3], Abdul Momin Syed[4], Hanzhi Xun[1], Laurentiu Braic[5], Nimer Wehbe[6], Ewout van der Veer[2,3,7], Nazek Elatab[4], Anteneh Gebregiorgis[1], Georgi Gaydadjiev[1], Beatriz Noheda[2,3], Said Hamdioui[1], Ryoichi Ishihara[1,8] & Heba Abunahla[1]

Memristor technology offers a promising route toward energy-efficient computing but faces challenges including resistance drift, variability, and the need for electroforming. Filamentary resistive random-access memory, one of the most studied memristive platforms, typically requires a high-voltage electroforming step to initiate conductive filaments, leading to increased power overhead and reduced endurance. Here we report HfO_2-based forming-free memristive devices (PdNeuRAM) that operate at low voltages, support multi-bit functionality, and exhibit reduced variability. Through combined electrical and materials characterization, we identify a Pd-O-Hf interfacial configuration that lowers oxygen-vacancy formation and migration barriers, creating a dense network of shallow defect states. Together with a Ti top electrode acting as an oxygen reservoir and an ultrathin (5 nm) HfO_2 layer, this interfacial engineering enables charge redistribution at room temperature and eliminates the need for electroforming. The fabricated devices provide tunable resistance states and reduce programming and read energy by 43% and 38%, respectively, in spiking neural network inference tasks. These results provide mechanistic insight into forming-free resistive switching and demonstrate the potential of Pd/HfO_2 devices for energy-efficient neuromorphic computing.

In this era defined by Artificial General Intelligence and Internet of Things technologies[1], Computing-in-Memory, where computation and storage seamlessly converge in a single physical locale, has ascended as a next-generation computing paradigm[2]. Among the many candidate technologies for computing cells, ReRAM has captivated both academia and industry with its widely studied attributes, including non-volatility, zero leakage power, a simplified structure, high integration density, rapid switching speeds, and inherent CMOS compatibility[1,3,4]. However, beneath these advantages lie formidable challenges, most notably, the prerequisite for electroforming pre-processing, where the electroforming phase necessitates the application of substantially elevated voltages to establish Conductive Filaments (CF) within the resistive medium[4]. It not only intensifies energy consumption and complicates fabrication but also compromises device reliability and endurance[5–7]. Moreover, the pronounced current density during ReRAM write operations further exacerbates concerns about power consumption and long-term durability, often forcing designers to enlarge ReRAM cell dimensions, a trade-off that diminishes circuit packing density and inflates manufacturing costs[5,8].

To address these challenges, HfO_2 has been widely explored as a premier dielectric candidate, celebrated for its superior breakdown voltage, robust thermal stability, and exceptional CMOS compatibility. These attributes promise not only a reduction in operational power consumption but also enhanced thermal performance over alternative dielectrics[9–12]. Furthermore, the incorporation of multi-bit technology in HfO_2 boosts memory density and slashes manufacturing costs[13–15], offering advantages for neuromorphic computing applications. Yet, current specialized fabrication methods, such as thermal annealing[16], X-ray irradiation[4], exotic element doping[16,17], and plasma treatments[17,18], are designed to introduce

[1]Department of Quantum and Computer Engineering, Delft University of Technology, Delft, The Netherlands. [2]Zernike Institute for Advanced Materials, University of Groningen, Groningen, The Netherlands. [3]CogniGron center, University of Groningen, Groningen, The Netherlands. [4]Computer Electrical Mathematical Science and Engineering Division, Electrical and Computer Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. [5]Core Labs, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. [6]Imaging and Characterization Core Lab, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. [7]Faculty of Physics, University of Duisburg-Essen, Duisburg, Germany. [8]QuTech, Delft University of Technology, Delft, The Netherlands. e-mail: e.hua@tudelft.nl; h.n.abunahla@tudelft.nl

defect states and reduce forming voltages. Aforementioned techniques will increase manufacturing costs and complexity, typically demand substantial energy, and may inadvertently impair transistor performance[19]. Therefore, there is a need for simpler, energy-efficient fabrication processes for forming-free ReRAM devices.

In this study, inspired by the high efficiency of biological neural cells, as illustrated in Fig. 1a, b, which underpin the acceleration of neuromorphic devices for fully connected artificial neural networks (Fig. 1c)[1],

Results

Device characterization

We designed crossbar arrays of devices with 2 μm × 2 μm node size (see "Method"), as displayed in Fig. 1d, e. The successful realization of the designed layout and stack is validated by optical microscopy and high-resolution scanning transmission electron microscopy (HRSTEM), as depicted in Fig. 1f, g. The distinct layers of the device stacks are clearly visible in the cross-sectional HRSTEM image, as shown in Fig. 1g, where the HfO_{2-x} layer exhibits a thickness of approximately 5 nm, sandwiched by 5 nm layers of Ti and Pd electrodes. Furthermore, HRSTEM-EELS imaging of PdHT (Fig. 1h) confirms the precise elemental distribution.

To examine the electrical properties of the fabricated devices, I-V characterizations were performed, as depicted in Fig. 2. Measurements on the electrical properties of the PdHT devices reveal three pivotal attributes for energy-efficient computing, including electroforming-free operation, low operating voltages, and tunable conductance. As illustrated in Fig. 2a, the PdHT devices manifest forming-free bipolar switching behavior. A statistical analysis of 42 randomly selected devices from the same die (Fig. 2b) reveals that the V_{SET} predominantly centers around 0.56 V, while the V_{RESET} is near −0.58 V. For comparison, similar characterizations were conducted on PtHT devices fabricated using an identical process except for choosing Pt as the top and bottom electrodes (TE/BE). As shown in Fig. 2d, these PtHT devices show bipolar switching, which is similar to PdHT devices, but they require an additional electroforming step (as illustrated in Fig. 2e), a result consistent with previous reports[20]. The Device-to-Device variability (D2D) test (Fig. 2e) confirms that PtHT devices require an electroforming voltage of approximately 2.3 V. Moreover, the statistical analysis presented in Table 1 confirms an alleviation in variability and a reduction in both V_{SET} and V_{RESET} for the PdHT structure compared to both the PtHT devices and the state-of-the-art reports[4,16,17], as detailed in Table 2. Table 2 presents a comprehensive comparison of state-of-the-art forming-free ReRAM devices, encompassing both non-HfO_2 and HfO_2-based systems. This analysis reveals that these devices typically exhibit one or more inherent limitations: they require elevated SET/RESET voltages, offer a constrained range of resistance states, or depend on specialized treatments to eliminate the conventional electroforming process. In contrast, our work addresses several of these limitations while maintaining low-voltage forming-free operation. We note that the present reliability metrics reflect a deliberate optimization toward forming-free, low-voltage analog switching; long-term non-volatile storage and online training are therefore beyond the current scope.

Furthermore, the reliability of the PdHT devices was also studied and certified by further electrical measurements. The multi-bit capability of the PdHT devices was evaluated using a linear voltage sweeping mode. As shown in Fig. 2c, the resistance is continuously modulated by varying the V_{RESET}, thereby precisely controlling the rupture degree of the CF. To ascertain the stability of the distinct resistance states, a READ-number dependent stability test was conducted (Fig. 2f), which demonstrates analog state functionality that permits the storage of eight tunable multi-bit weights within a single memory cell. Furthermore, retention (Fig. 2g) and endurance (Fig. 2h) assessments reveal the stable resistance states between LRS and High Resistance States (HRS), with non-volatility, where the resistance states remain stable for over 4.5×10^4 s, and the device can endure 2.5×10^3 cycles without noticeable performance degradation, though the endurance is not its optimum pattern. (tests under higher temperatures and a higher number of cycles are needed for further investigation) Additionally, to gain deeper insight into the tunable resistance states, we conducted a comprehensive analysis of the conductance variation versus $V_{RESET,stop}$, as depicted in Fig. 2i. The results reveal an exponential gradual decline in conductance with increasing $V_{RESET,stop}$, indicating that the extent of CF rupture can be modulated by adjusting the amplitude of $V_{RESET,stop}$. This tunability may be further enhanced by controlling additional pulse parameters, such as pulse numbers, pulse amplitudes, and pulse width of $V_{RESET,stop}$. Although the baseline memory window of 19× is modest relative to forming-based devices, it is fully adequate for stable multi-level neuromorphic operation. Importantly, by increasing the SET compliance current, the MW can be expanded up to 65× as explained in the Supplementary Fig. 14, demonstrating that compliance-current engineering provides a straightforward route to further optimize the conductance range.

Understanding the forming-free behavior

We proposed a possible explanation for the forming-free behavior (see Fig. 3a) based on a suite of comprehensive characterizations, including the Pd/HfO_{2-x} interface imaging by HRSTEM-iDPC (iDPC stands for integrated Differential Phase Contrast), which allows us to image light atoms in the presence of heavy atoms (as shown in Figs. 3b and S5a), the surface profiles of Pd and Pt thin films obtained via Atomic Force Microscopy (AFM) (Fig. S6a, b), the Arrhenius activation energy measurements for both PdHT and PtHT memristors using a semiconductor analyzer (Fig. 3c), and the elemental distribution profiles acquired from Rutherford Backscattering Spectroscopy (RBS) (Fig. 3d).

The activation energies for the pristine PtHT and PdHT devices are compared in Fig. 3c via the Arrhenius equation, $\rho(T) = \rho_0 \exp(E_a/k_B T)$, where $\rho(T)$ denotes the resistivity, T is the absolute temperature in Kelvin, ρ_0 is the pre-exponential factor, E_a represents the activation energy, and k_B is the Boltzmann constant. As illustrated in Fig. 3c, the activation energy exhibits two distinct regimes: a Low-Temperature (LT) regime (300–400 K) marked by lower activation energies (denoted as E_{a1}) and a High-Temperature (HT) regime (400–470 K) characterized by larger activation energies (denoted as E_{a2}).

Notably, the pristine PdHT devices exhibit activation energies of E_{a1} = 0.055(8) eV and E_{a2} = 0.22(2) eV, which are lower than those of the PtHT devices (E_{a1} = 0.124(4) eV and E_{a2} = 0.502(6) eV). The activation energy E_{a2} of PtHT is possibly attributed to the diffusion barrier associated with the migration of doubly positively charged oxygen vacancies inside HfO_{2-x}, in accordance with previous reports[21–24]. It may arise from the diffusion barrier of doubly negatively charged interstitial oxygen ions, based on existing ab initio calculations[22,24,25]. Meanwhile, it is noteworthy that the Frenkel-Pair (FP) dissociation process in HfO_2 typically requires activation energies on the order of 5.2 eV[22,26]. Furthermore, if the FP process comprises a doubly positively charged pair ($V_O^{2+} + O^{2-}$), its formation energy is 5.8 eV[27], a value closely related to the electroforming process, as reported previously[22,28], implying that FP dissociation is highly unlikely to occur under both LT and HT conditions without an applied voltage bias.

In pristine PtHT devices, the presence of V_O^{2+} is expected in amorphous hafnia devices and is introduced during the sputtering process[29], though it remains at a low level (see more details in the ESI part 1 about doubly positively charged oxygen vacancies formation). In contrast, in PdHT devices, the activation energy barriers seem to correspond to those reported for singly positively charged (V_O^{+}) defects or interstitial oxygen O_i^{2-} defects. This difference is likely due to the incorporation of Pd in HfO_{2-x} at the Pd/HfO_{2-x} interface (as observed in Fig. 3b), adsorbing the oxygen ions around the Pd atoms (Pd-O_i^{2-}), generating more oxygen vacancies along the conductive path. Interstitial Pd atoms alter the defect landscape in hafnia by creating more oxygen vacancies and donating electrons, thereby stabilizing oxygen vacancies as singly positively charged (V_O^{+}) defects and forming V_O^{+}-O_i^{2-} pairs, which serve as traps for electrons hopping[27]. Additional activation energy measurements on the HRS state of both PdHT and PtHT memristors (Fig. S3) demonstrate that after electroforming (i.e., completion of FP dissociation) in PtHT devices, the dominant charges exhibit an activation energy of approximately 0.22(2) eV, similar to that observed in PdHT

devices, which is also consistent with the fact that after electroforming of the PtHT devices, the IV characteristics of both types of devices are basically indistinguishable.

Considering all the above, the forming-free phenomenon likely originates from Pd atoms initially forming a conductive bridge through bonding oxygen ions from hafnia, generating more oxygen vacancies (see Fig. 3a). The incorporation of Pd supplies electrons that effectively reconstruct oxygen vacancies/ions in the Pd-O-Hf configuration relative to the conventional Hf-O configuration in PtHT memristors. Furthermore, under minimum bias, it provides the maximum electrons among the Pd bridges,

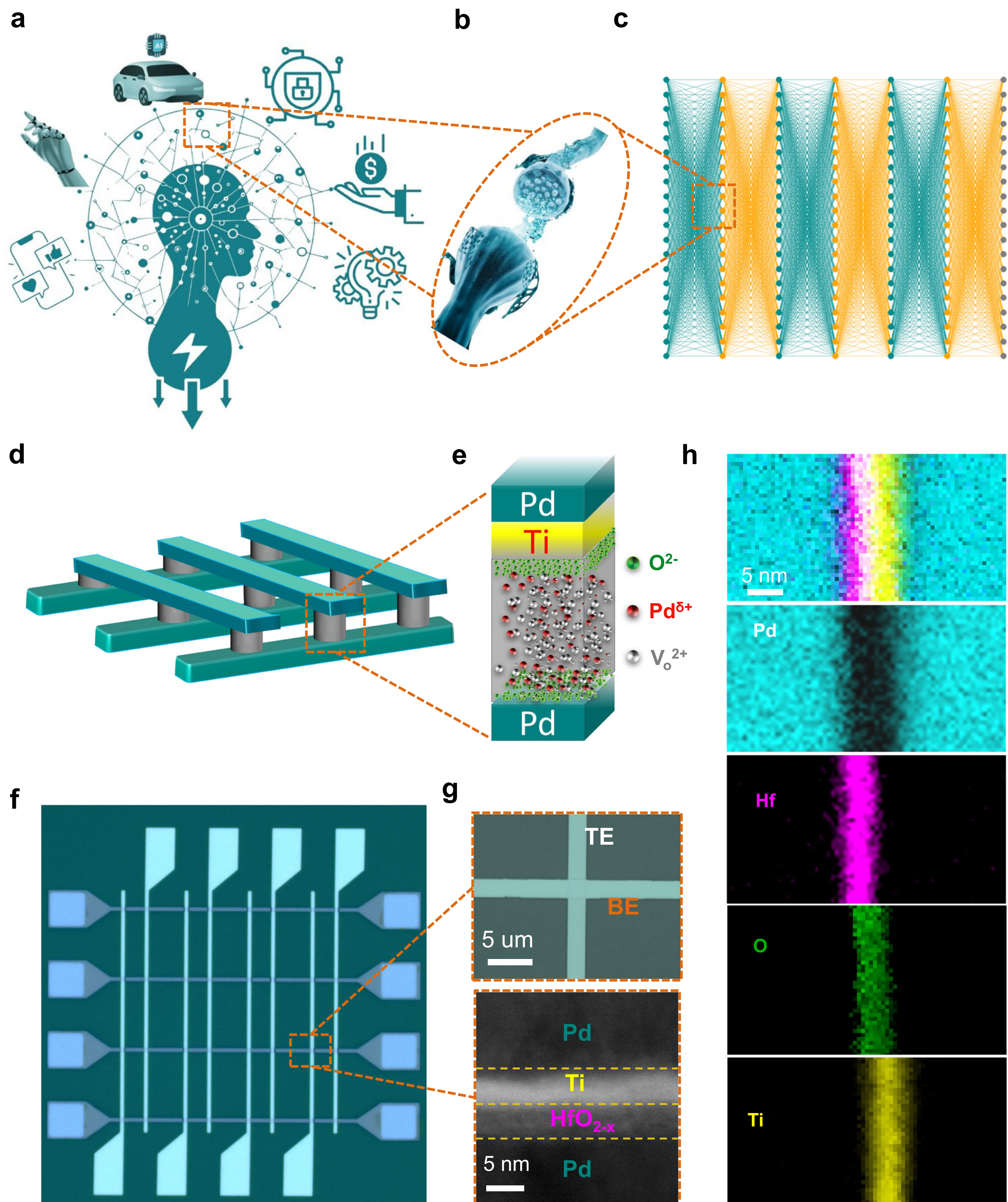

Fig. 1 | Overview of neuromorphic systems and PdNeuRAM device structure. **a** Schematic overview of energy-efficient neuromorphic system applications. **b** Schematic of a biological synapse. **c** Fully connected artificial neural network architecture. **d** Crossbar array schematic. **e** Schematic of the Pd/HfO_{2-x}/Ti/Pd ReRAM cell. **f** SEM image of fabricated 4 × 8 crossbar array. **g** Optical image of a single cell (top) and cross-sectional STEM image (bottom). **h** Cross-sectional EELS elemental mapping of a PdHT device in the low-resistance state (LRS).

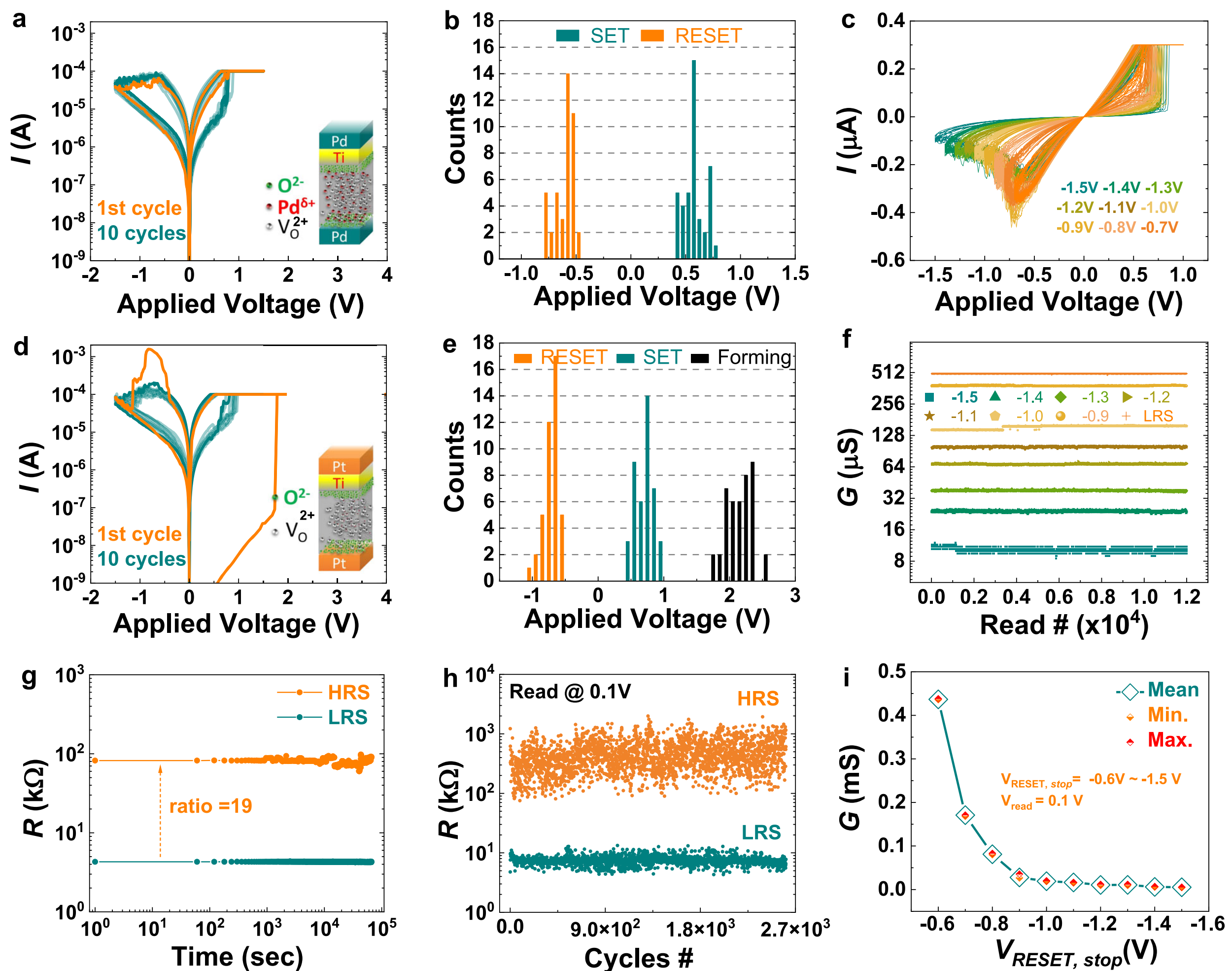

Fig. 2 | Electrical characterization of PdHT and PtHT devices. a Current-voltage (I-V) characteristics of forming-free PdHT devices. **b** Statistical distribution of V_{SET} and V_{RESET} measured from 42 devices on the same die. **c** Multi-level resistance tuning by varying $V_{RESET,stop}$. **d** I-V characteristics of PtHT devices. **e** Statistical distribution of V_{SET}, V_{RESET}, and forming voltage for PtHT devices ($n = 42$ devices). **f** Read stability of eight conductance states measured over 1.2×10^4 read cycles at 0.1 V. **g** Retention characteristics over 4.5×10^4 s with 60 s read interval. **h** Endurance performance over 2.5×10^3 switching cycles, where orange and dark cyan represent high resistance states (HRS) and low resistance states (LRS), respectively. **i** Conductance variation as a function of $V_{RESET,stop}$ (20 ms step duration). (Dark cyan, orange, and red colors represent mean values, minimum (Min.) values, and maximum (Max.) values, respectively.

Table 1 | Comparison of operating voltages and their variability for PtHT and PdHT devices, including cycle-to-cycle (C2C) and D2D

Structures	C2C (V_{SET})	D2D (V_{SET})	V_{SET}	V_{RESET}
PtHT	11.0%	19.2%	0.68 V	−0.72 V
PdHT	7.7%	16.8%	0.56 V	−0.58 V

and its inherent special configuration provides electron injection by effectively providing electrons to defective sites and eases the forming process[22] and conductive path for the initial states.

To substantiate the hypothesis of the Pd-induced conductive bridges in HfO_{2-x} of PdHT, RBS measurements were conducted. As depicted in Fig. 3d, the HfO_{2-x} layer in the PdHT memristors exhibits a higher Pd concentration compared to the Pt concentration in PtHT devices. Remarkably, even after annealing the PdHT and PtHT devices at 573 K for 5 min, the Pd atomic ratio within the HfO_{2-x} layer remains higher than that of Pt in the PtHT devices, as further illustrated in Fig. S7a. This observation agrees with the lower migration barrier of Pd into semiconductors compared with other metals such as Li, Cu, Ag, Pt, and Au[30]. As further demonstrated in Fig. S16, the SIMS depth profiles clearly reveal Pd-O-Hf interdiffusion, providing direct chemical evidence for the modified interfacial region that underpins forming-free conduction in PdHT devices.

We employed HRSTEM-HAADF to examine the Pd/HfO_{2-x} interface. As shown in Fig. 3b, the images reveal that the Pd-Hf atoms are clearly observed above Pd electrodes taken in the [011] zone, showing a region where crystallized HfO_{2-x} (evidenced by small, moderately bright atoms, marked in purple spheres as shown in the most right HRSTEM-iDPC image) intermixed with Pd (characterized by larger, highly bright atoms, marked in orange spheres), which is further emphasized in Fig. S5b with red circles. For further validation of the Pd-O-Hf configuration, please refer to supplementary part 2 for further discussion on it.

The influence of Pd and the thickness of HfO_{2-x} or Ti on forming voltages was examined via the I-V characteristics of various devices with various electrodes (see Figs. S4a, b). The critical role of Pd electrodes is underscored by their ability to markedly reduce or even eliminate the need for electroforming voltages, in stark contrast to the PtHT structure (Fig. S4c). Moreover, as the HfO_2 thickness increases, a distinct electroforming behavior emerges (Fig. S4d), likely due to the finite extent of the Pd-O-Hf configuration within HfO_2, which results in an exponential increase in resistance with a linear increase in thickness. This is also consistent with the previous report[7,22] that forming-free behavior is correlated with the thickness of HfO_{2-x} and oxygen distribution, where the Pd-O-Hf mechanism operates effectively in

Table 2 | State-of-the-art forming-free ReRAM devices

Types	TE/BE	Th_{OEL}[a]	P^{b}_{SET}	MW[c]	$V_{SET/RESET}$	MLR[d]	Treatment
non-HfO_2	W/Pd[53]	20 nm WO_x	1 mW	~20	1.0, −3.0	No	oxygen plasma
	Ta/Pt[54]	7 nm Ta_2O_5	n.a.	≥10	1.2, −0.9	No	600 °C annealing
	$SrTiO_3$/Pt[55]	400 nm MoO_3	n.a.	~10	3.5, −3.8	No	proton injection
	Al/Si[56]	150 nm SiN_x	0.16 mW	~10	3.2, −1.5	No	hydrogen plasma
	TiN/TiN[57]	5 nm $HfTiO_x$	0.7 mW	≥30	1.0, −1.0	2	400 °C annealing
HfO_2	Ge/Pd[16]	5 nm HfO_2	0.16 mW	≥500	3.2, −0.8	No	600 °C & doping
	Al/Si[58]	60 nm HfO_x	25.9 mW	~3	2.9, −2.5	No	550 °C & doping
	Pt/TiN[4]	20 nm HfO_2	6.0 mW	≥10	0.70, −0.70	No	X-Ray
	TiN/TiN[7]	3 nm HfO_x	0.225 mW	~10	0.75, −0.6	2	450 °C annealing
	Pd/Pd	5 nm HfO_{2-x}	0.056 mW	~65	0.56, −0.58	≥3	No [This work]

[a]Thickness of oxygen exchange layer.
[b]For consistency with prior literature, Table 2 reports the peak SET power estimated as $P_{SET} = V_{SET}I_{CC}$. Detailed pulse-integrated switching energies under various pulse widths are provided in (Table S12).
[c]Memory Window.
[d]Multi-level resistance (MLR) and number of bits.

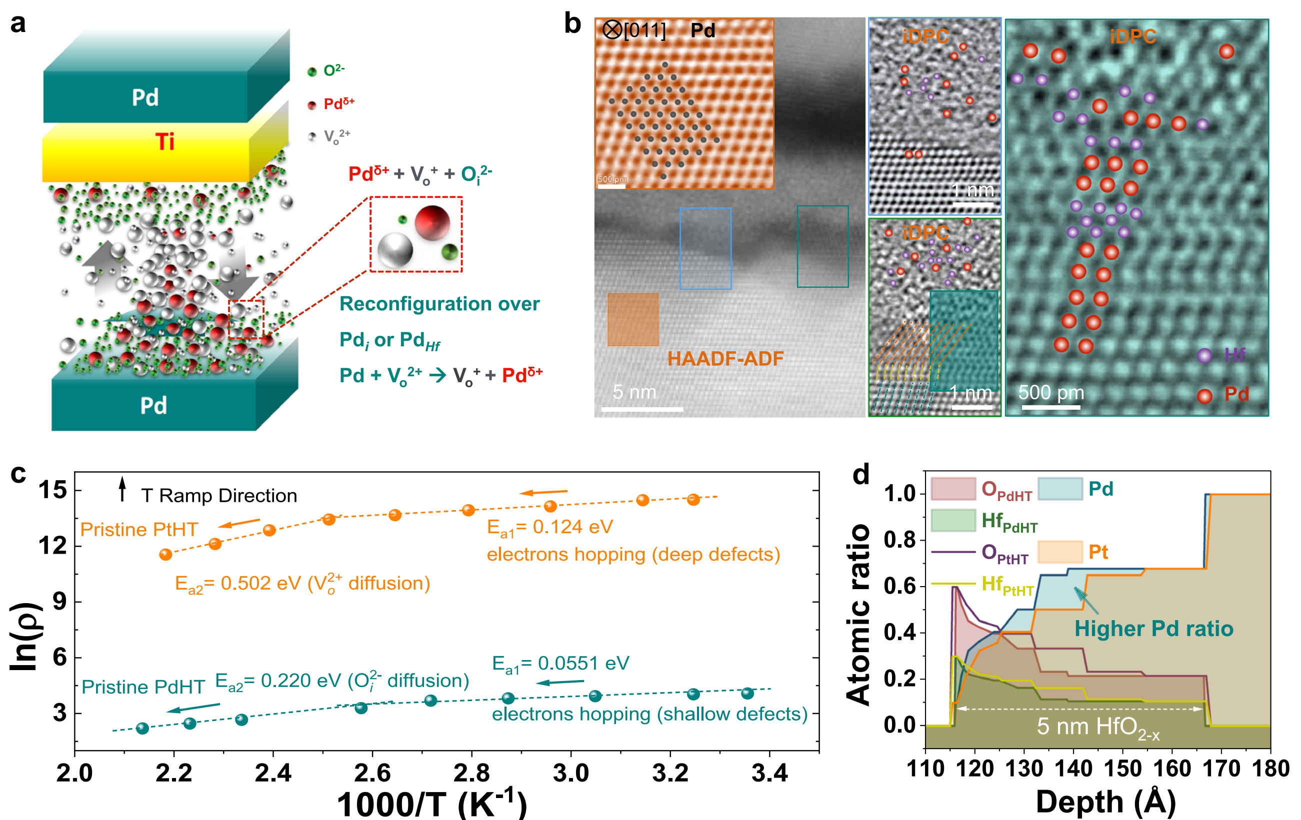

Fig. 3 | Mechanistic investigation of forming-free behavior. a Proposed forming-free mechanism schematic. **b** STEM-HAADF (scanning transmission electron microscopy-high-angle annular dark-field) and iDPC (integrated differential phase contrast) images of the PdHT interface. **c** Activation energy extraction from Arrhenius plots for pristine PtHT and PdHT devices. **d** Rutherford backscattering spectroscopy (RBS) elemental depth profile of the PdHT stack.

the 5 nm regime, validating its suitability for scaled, forming-free integration. Additionally, the thickness of the Ti examination is as illustrated in (Fig. S13).

It is important to note that the Ti top electrode plays a distinct and more limited role in this mechanism. Ti forms a thin TiO_x interfacial layer that acts as an oxygen reservoir during operation, assisting vacancy creation and recombination near the top interface and thereby smoothing the RESET transition. However, Ti alone does not generate the extended shallow-defect network required to bypass electroforming. Control devices with identical 10 nm HfO_2 thickness but varied Ti thickness at the top electrode (30 nm Pd/0 nmTi/10 nm HfO_2/30 nm Pd versus 30 nm Pd/3 nm Ti/10 nm HfO_2/30 nm Pd) still require a forming step and show limited cycle endurance, even though Ti improves RESET regularity (Fig. S13). These observations, together with literature reports[22,31,32] that Ti primarily acts as an oxygen exchange layer rather than an intrinsic forming-free trigger, confirm that the forming-free behavior in our 5 nm Pd/HfO_2/Ti/Pd devices is governed by the Pd-O-Hf-induced shallow-defect landscape in an ultrathin oxide, with Ti mainly assisting vacancy supply and RESET stability rather than replacing the forming step.

Investigation of conduction mechanism

In ReRAM devices, multiple conduction mechanisms[33,34], such as Ohmic conduction, trap-mediated conduction, trap-assisted tunneling (TAT),

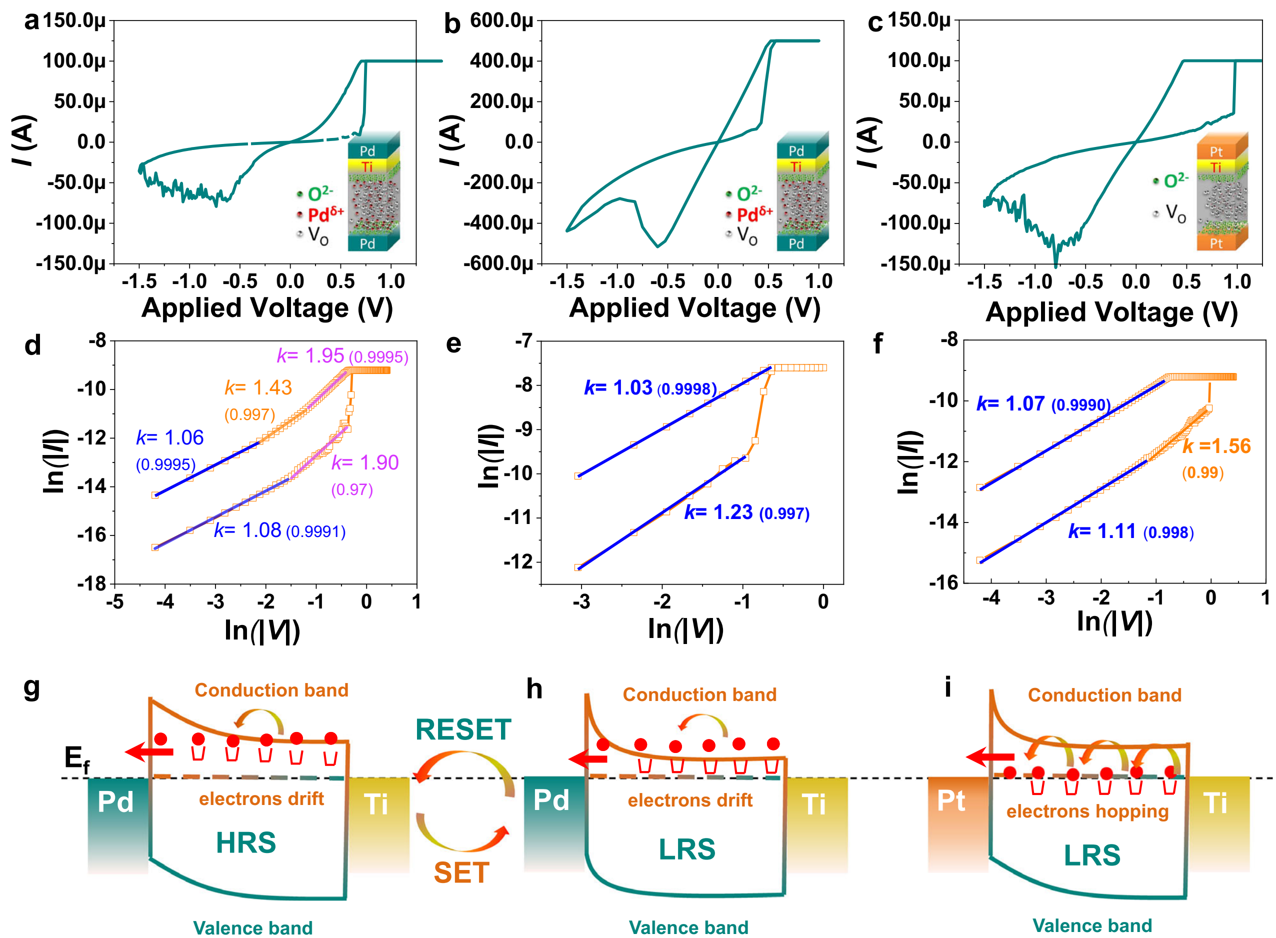

Fig. 4 | Investigation of conduction mechanism from I-V curves fitting for PtHT and PdHT. **a** I-V curves of PdHT device with 0.1 mA compliance current, **b** 0.5 mA compliance current. **c** Present PtHT device I-V curve with 0.1 mA compliance current. **d–f** The fitted I-V plots of positive bias parts **a–c**, respectively. **g, h** Illustration of HRS and LRS band diagrams of PdHT and electrons drifting between shallow vacancy states, respectively. **i** The band diagram of PtHT LRS states and electron hopping between deep defective sites.

space-charge-limited conduction (SCLC), Poole–Frenkel emission, Schottky emission, and valence change-based mechanisms, collectively dictate the overall device behavior. In particular, SCLC is a prominent mechanism observed in ReRAM devices during the high-voltage regime[34]. Here, the current is constrained by a space-charge-limited region in which charge carriers, injected from the electrodes, traverse localized trap states within the resistive switching material. The SCLC behavior is critically influenced by the trap density and trap energy levels, which govern the mobility and transport of the charge carriers.

A trap-controlled SCLC regime can be segmented into two distinct portions: the trap-unfilled (trap-limited) regime and the trap-filled (trap-free) regime. In the Mott-Gurney region, characterized by the power law $I \propto V^2$, a steep increase in current is observed at high electric fields. This behavior is typically identified by an initial Ohmic conduction at low fields, followed by a transition to a power-law dependence as electrode-injected electrons surpass the equilibrium concentration. Consequently, SCLC conduction is more likely when the electrode contact exhibits high carrier injection efficiency[34]. The SCLC conduction mechanism is quantitatively described by the Mott–Gurney law (Eq. (1)), which relates the current density J to the applied voltage V and material properties:

$$J = \frac{4}{9}\epsilon_0 \mu \frac{V^2}{d^3} \quad (1)$$

Here, ϵ_0 denotes the dielectric permittivity, μ is the carrier mobility, d represents the film thickness, and V is the applied voltage.

The slope (k) of the double logarithmic I-V curve evolves from approximately 1.0 (indicative of trap-limited Ohmic conduction) through 1.5 (reflecting the trap-unfilled transitional region) to nearly 2.0 (characteristic of trap-free SCLC) for both HRS and LRS, as observed in Fig. 4a and d for PdHT devices under a 0.1 mA compliance current. On the contrary, when the compliance current (CC) is increased to 0.5 mA, the slope k for HRS(see Fig. 4b, e) approaches around 1, indicating that both the HRS and LRS are dominated by Ohmic conduction in both HRS and LRS, a consequence of reduced interface state density at the Pd/HfO_{2-x} interfacial region owing to filament formation. Conversely, PtHT devices exhibit ohmic conduction under a 0.1 mA compliance current CC for both HRS and LRS (Fig. 4c, f). Beyond the representative cases shown in Fig. 4a, b, the compliance current provides a continuous tuning knob for filament strength and multi-level conductance spacing. As shown in (Fig. S14), increasing the SET compliance current progressively lowers the LRS and expands the memory window, while gradual RESET behavior is preserved across the entire compliance-current range. An optimal compliance-current window between 0.1 and 0.5 mA is identified, where stable multi-level states, sufficient conductance spacing, gradual RESET dynamics, and controlled leakage are simultaneously achieved. These results indicate that at the Pd/HfO_{2-x} interface in PdHT devices, shallow defect states are preferentially formed, enabling electrons to enter the conduction band (see Fig. 4g, h). In contrast, PtHT devices tend to form deep defect states, wherein electrons tunnel into vacancy defect sites[34] (see Fig. 4i).

For completeness, we also analysed the same I–V data using the standard diagnostic forms for Schottky emission, Poole–Frenkel emission,

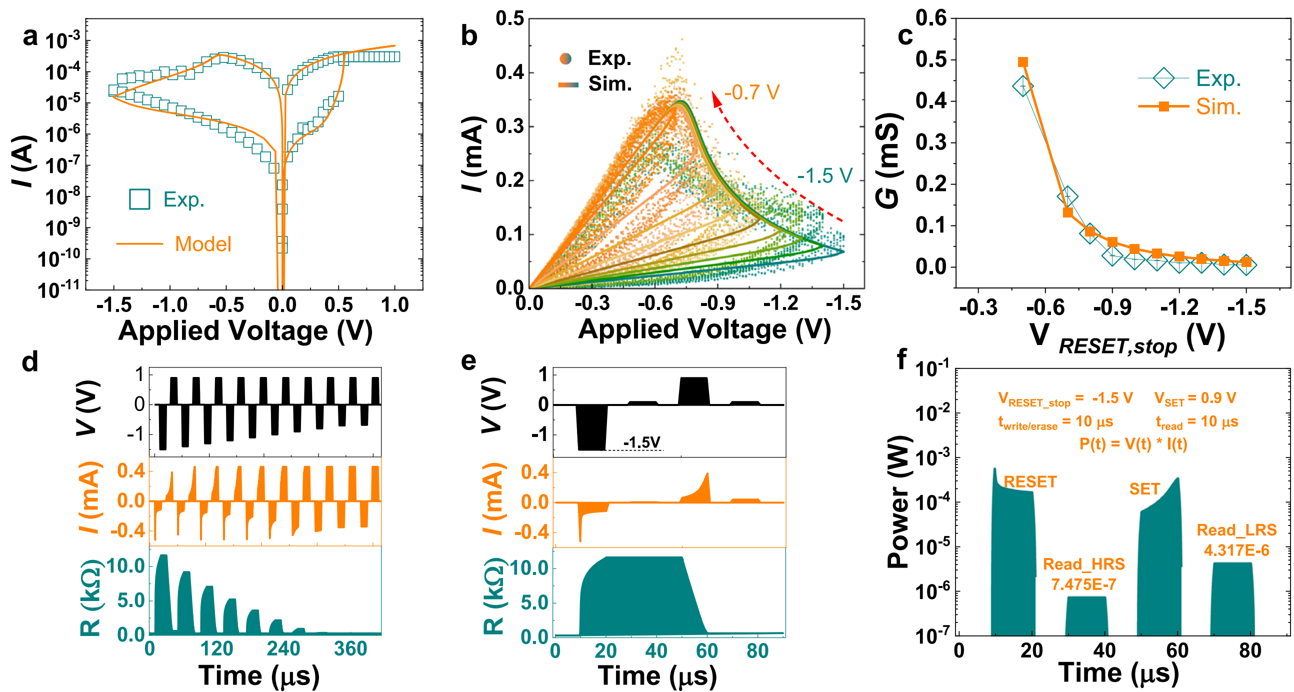

Fig. 5 | Power consumption evaluation. **a** Experimental (Exp.) DC I-V curve calibration with JART memristor model. **b** MLR verification using DC sweeping by changing V_{RESET} compared to simulated results (Sim.) using JART model. **c** Comparison of conductance gradual change for experimental (Exp.) and simulation results (Sim.). **d** Eight resistance states programming schemes, including the results of currents and resistance over time. **e** Programming scheme of the resistance state with $V_{RESET,\ stop} = -1.5$ V. **f** Energy consumption evaluation for programming the conductance states in (**e**).

and tunneling conduction. As detailed in Supplementary Fig. S15, these models either fail to produce linear diagnostic plots or require unphysical parameter values and are therefore ruled out as the dominant transport mechanisms. In contrast, the PdHT devices exhibit trap-controlled SCLC at low compliance current, transitioning toward Ohmic-like conduction at higher compliance current, consistent with a progressive increase in shallow defect density near the Pd-O-Hf interface.

Taking into consideration all the aforementioned analysis in Figs. 3 and 4, including the activation energy study, HRSTEM observations of Pd atoms, EELS, and RBS elemental distribution profiles of PdHT, and I-V curve fitting, in PdHT devices, the Pd-O-Hf configuration, including the Pd-induced conductive bridges and electron injection from Pd atoms in HfO_{2-x}, facilitates oxygen vacancies reconfiguration and electron drift toward the Pd electrode via tunneling. It hereby contributes to the elimination of the electroforming process. After the RESET operation for PdHT devices (from initial state to HRS), the Pd-induced conductive bridges will be partially damaged irreversibly by scavenging the oxygen vacancies through oxygen ions from TiO_x, leaving the shallow defective states thanks to the oxygen affinity nature of Pd. In contrast, in PtHT devices, electrons predominantly hop between deep defect sites generated by high oxygen vacancy concentrations after the electroforming process under higher electric fields. Although the Pd–O–Hf interfacial complexes introduce shallow defect states that lower the activation energy for carrier hopping and enable electroforming-free, trap-assisted transport, shallow traps alone do not suppress the current magnitude. Instead, the reduced read current in PdHT devices originates from the conduction morphology: Pd-induced defects form a spatially distributed, interface-localized percolation network rather than a narrow metallic filament as in PtHT devices. This electroforming-induced-filament-free, trap-limited conduction geometry results in lower effective carrier mobility, higher resistance, suppressed current overshoot, and therefore reduced read and write energy.

Device calibration and electrical simulation

To evaluate the energy consumption of the PdHT and PtHT devices, we calibrated the DC I-V curves of the real devices using the JART model[35]. In the case of PdHT, as illustrated in Fig. 5, the simulated I-V performance exhibited excellent concordance with the experimental results (see Fig. 5a). Regarding the multilevel resistance state performance, the rupture of the CF displayed a well-controlled, gradual transition upon tuning the $V_{RESET,\ stop}$, as demonstrated in Fig. 5b, where multiple distinguishable I-V curves were observed from Cadence simulation, which implies that the results of PdHT device calibration agree well with experimental results. Similarly, the incremental change in conductance by varying biasamplitude for $V_{RESET,\ stop}$ showed agreement between experimental observations and simulation results (Fig. 5c). Furthermore, distinct resistance states can be programmed using variable RESET stop voltages with a 10 μs pulse duration, leading to the conclusion that 8 discernible resistance states can be established with corresponding programming schemes, as depicted in Fig. 5d. The detailed programming protocol was further illustrated in Fig. 5e, where each programmed resistance state is successfully achieved using a writing/reading pulse of 10 μs at $-1.5 V/100$ mV, respectively. To assess the power consumption during programming, we integrated the instantaneous power $P(t) = I(t) \cdot V(t)$ over time, as shown in Fig. 5f, where the energy consumption evaluation is as the equation described:

$$E = \int_{t_0}^{t_1} P(t)\, dt = \int_{t_0}^{t_1} I(t)\, V(t)\, dt$$

where the limits t_0 and t_1 bracket the programming interval. Here, both RESET/SET energy consumptions contribute to the total programming energy, while the sum of HRS and LRS reading energies represents the total energy consumption during reading. The programming energy consumption for 8 distinct conductance states for both PdHT and PtHT devices is summarized in Tables S1, S2, respectively.

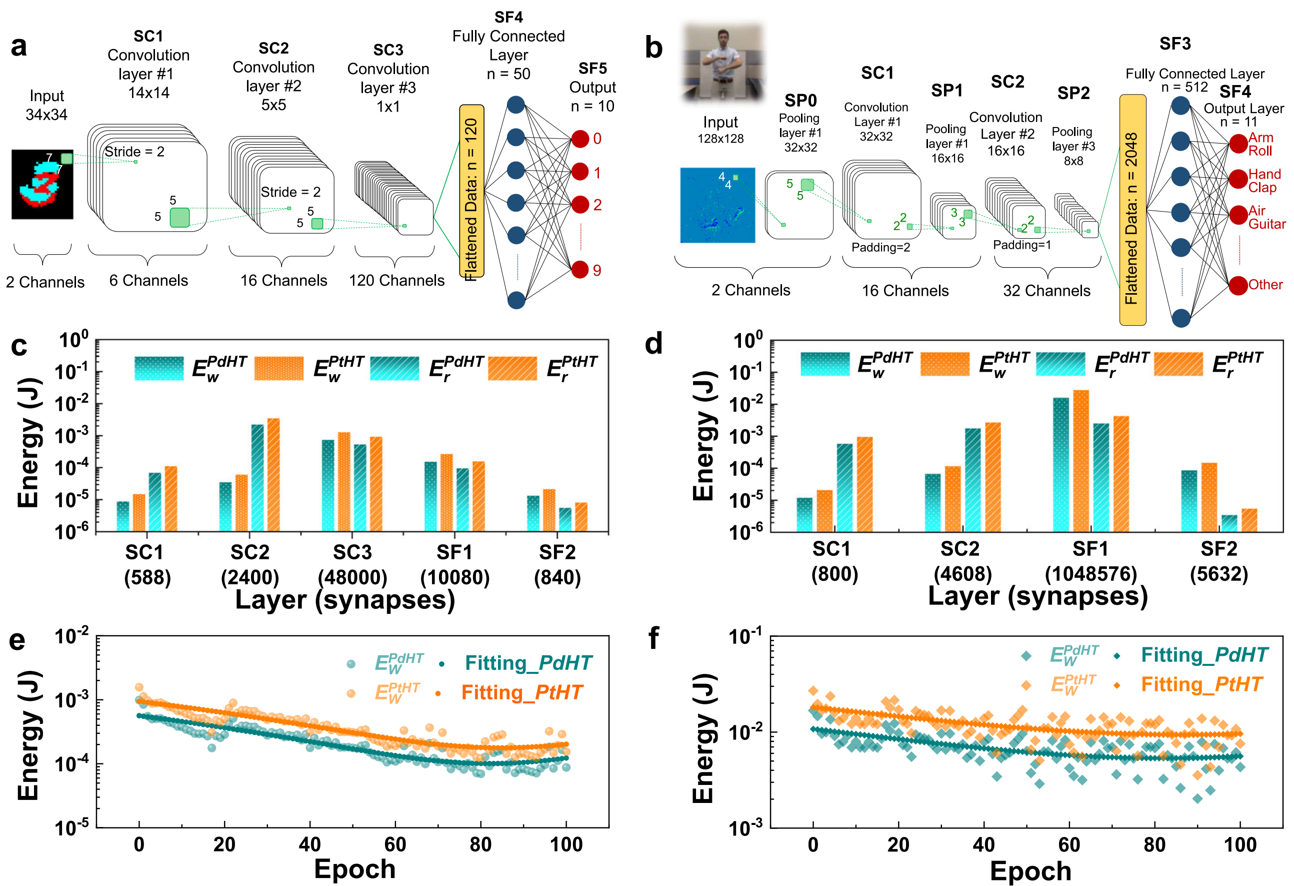

Fig. 6 | System-level energy evaluation in spiking neural networks. a, **b** Network architectures for N-MNIST and DVS128 Gesture datasets. **c**, **d** Per-layer write and read energy consumption after one-time programming. **e**, **f** Write energy per training epoch during online learning. Energy values are obtained by summing the programming or read energy across all devices storing synaptic weights.

System-level energy simulation for neuromorphic applications

For the demonstration of the proposed PdHT ReRAM device in realistic neuromorphic applications, we designed two deep convolutional SNNs targeted at the classification of the N-MNIST[36] (Fig. 6a) and IBM DVS128 Gesture[37] (Fig. 6b) datasets, respectively. These networks are implemented in Python using primitives from the open-source Spike Layer Error Reassignment (SLAYER)[38] and PyTorch[39] frameworks, and they are trained with a variant of the backpropagation algorithm. Spiking neurons are modeled on the Spike Response Model (SRM), an advanced generalization of the ubiquitous Integrate-and-Fire neuron model[40]. At the output, the class corresponding to the neuron that produced the highest number of spikes is selected as the winning class. For further methodological details, please refer to the Methods section.

To estimate the energy consumption at the system level, we perform a series of experiments and report the overall read and write energy measured collectively for all ReRAM devices needed to store the synaptic weights across each network. In the first set of experiments, we train the two SNNs in software, and then we write the final synaptic weights directly onto the ReRAM devices. For simplicity and generalization purposes, there is no limit to the maximum number of employed devices. The first two bars in Fig. 6c, d report the write energy, E_{write}, per layer on a logarithmic scale for the N-MNIST and Gesture SNNs, respectively. As anticipated, the write energy scales proportionally with the number of synapses, since a greater number of weights necessitates more devices. Notably, the PdHT devices achieve an overall reduction in write energy of 43% with respect to that of PtHT for each SNN.

In the case of online learning, the synaptic weights are updated and written back on the ReRAM devices at every epoch of the training of each SNN, and the overall write energy per epoch is illustrated in Fig. 6e, f for the two case studies, respectively. Initially, the synaptic weights are set to small random values, resulting in a peak write energy equal to the total write energy occurring by summing the corresponding bars for all layers in Fig. 6c, d. This is because in the first iteration, almost all ReRAM devices need to be programmed to a new value. As training progresses, only those devices that require updates contribute to the write energy, leading to a gradual reduction that mirrors the convergence of the synaptic weights, thereby lowering the overall energy consumption.

Regarding the read energy E_{read}, during inference, represented by the rightmost two bars in Fig. 6c, d, the energy consumption per layer exhibits a pattern distinct from that of the write energy. In SNNs, each layer functions as a filter that reduces the number of spikes transmitted to subsequent layers, resulting in progressively sparser spiking activity. Therefore, the read energy depends not only on the number of synapses but also on the density of incoming spikes. For example, layer SF1 of the Gesture SNN consumes only 4× more energy than its predecessor, despite having 228× more synapses. Overall, both networks exhibit an approximate 38% reduction in read energy during the complete inference of their respective test sets. The average energy per spike during inference is summarized in Supplementary Table 7 and demonstrates a comparable energy reduction when utilizing the proposed PdHT devices over PtHT ones. The reduced read energy in PdHT devices originates from their intrinsically higher resistance levels under identical read bias conditions, resulting in a lower read current and thus reduced $E_{read} = \int V_{read}(t) I_{read}(t)\, dt$. This behavior is a direct consequence of the forming-free Pd-O-Hf defect landscape, which suppresses dense, forming-induced conductive filaments and stabilizes shallow-defect-dominated transport compared to PtHT devices.

PdHT and PtHT devices inherently operate in different resistance regimes due to their distinct defect landscapes: as shown by the activation-

energy and conduction-mechanism analyses (Figs. 3 and 4), forming-free PdHT devices develop sparse, shallow defect networks with intrinsically higher resistance, whereas PtHT devices require electroforming and form denser, deep-level filaments. Accordingly, energy efficiency is evaluated in each device native, physically stable operating window, as artificially normalizing resistance would require forcing the devices into non-native and non-physical regimes.

Discussion

In this work, we reported the successful CMOS-compatible fabrication of PdHT ReRAM devices, PdNeuRAM. High-resolution imaging analyses, including SEM, HRSTEM, and AFM, confirm well-defined device stacks with minimal interfacial roughness. Electrical characterization reveals forming-free switching at relatively low operating voltages, with V_{SET} around 0.56 V and V_{RESET} averaged at −0.58 V. Notably, the PdHT devices exhibit reduced C2C and D2D variability comparable to their PtHT counterparts, indicating robust switching behavior and enhanced reliability.

From previous studies on oxygen chemisorption on Pd and Pt films and Schottky barriers[41] and MIGS[42,43], it seems unlikely that the Schottky barrier differences cause the electroforming differences. The absence of an electroforming step in PdHT devices is primarily attributed to the presence of Pd atoms within the HfO_{2-x} layer for a few nanometers, acting as Pd-induced conductive bridges, thus eliminating the need for a high-voltage electroforming process. Furthermore, regardless of the role of Pd in the conduction mechanism, the PdHT presents similar resistive switching properties to that of PtHT, as indicated by the identical activation energy, which is in turn in agreement with the behavior reported in the previous HfO_{2-x} devices[4,7].

The Pd-induced oxygen exchange reaction produces nanoscale clusters of shallow traps that are spatially localized. These nanoscale trap clusters define the effective conduction cross-section and therefore do not scale with device area. This behavior is consistent with reported Pd/Hf/HfO_2/Pd ReRAM scaling, where localized conduction leads to weak or non-monotonic area dependence[44–47]. Because the switching volume is governed by a nanoscale Pd-O-Hf interfacial region rather than the lithographic footprint, PdHT remains compatible with post-CMOS scaling and preserves analog programmability even at reduced device dimensions.

In addition to geometric scalability, the defect landscape in PdHT can be systematically tuned through the SET CC. As shown in Fig. S14, low CC values favor the formation of sparse, shallow defect states associated with the Pd-O-Hf interface, enabling fine-grained analog weight updates and smooth conductance modulation. Increasing the CC facilitates the formation of denser and deeper defect states, which stabilize the conductive pathway and expand the available memory window up to ~65×. This CC-dependent defect evolution offers a practical means of selecting the optimal operating regime: high-precision incremental updates at low CC, or enhanced state robustness and dynamic range at higher CC.

Such tunability naturally aligns with neuromorphic inference workloads, where synaptic weights are updated or refreshed during operation and do not require long-term non-volatile retention. The demonstrated eight-level programmability and multi-hour stability are therefore sufficient for practical inference tasks, while the ability to access deeper, more stable defect configurations at higher CC provides a route for applications requiring larger memory windows or increased state robustness. Furthermore, established materials-engineering strategies, such as inserting thin diffusion-barrier layers (e.g., Al_2O_3, TaO_x)[48–50], optimizing the Ti oxygen-reservoir thickness[22,31,32], and engineering interface dipoles[34,51,52], offer clear pathways to further enhance endurance and defect stability without compromising the forming-free nature of the PdHT stack.

Current-voltage (I-V) analysis indicates that both PdHT and PtHT devices operate under an SCLC mechanism. At low bias, conduction follows an ohmic behavior. As the voltage increases, trap-mediated processes dominate; injected carriers fill shallow defect states, transitioning the conduction to a quadratic dependence on voltage ($I \propto V^2$), consistent with the Mott-Gurney law. In PdHT devices, intrinsic insulating layer defects and interfacial imperfections, enhanced by Pd-O-Hf interactions, create shallower defect states that promote efficient electron transport and lower variability. In contrast, PtHT devices exhibit deeper defect states, which vary with electroforming voltage and material modifications, leading to increased variability.

The effect of CC was further examined, revealing that higher I_{CC} values produce denser filaments and lower LRS resistance, transitioning the conduction from SCLC-like to Ohmic (Fig. 4a, b). A systematic I_{CC} sweep (Fig. S14) confirms that tuning CC reliably expands the multi-level conductance window while preserving gradual RESET behavior. The observation that PdHT devices biased with a 0.5 mA CC and PtHT devices biased with a 0.1 mA CC both exhibit Ohmic conduction can be understood in terms of compliance-current-controlled filament formation and stack-dependent interface properties. In our devices, the CC sets the extent of oxygen-vacancy accumulation and the effective cross-sectional area of the conductive filament in HfO_2. At low compliance currents, transport is dominated by narrow or partially formed filaments and localized trap-assisted conduction, resulting in non-linear I-V characteristics. As the CC is increased, the filament thickens and becomes more continuous, such that the device resistance is increasingly governed by a quasi-metallic conductive path and the series resistance of the contacts, which produces an approximately linear (Ohmic) I-V response within the applied read window. Because the Pd/HfO_2 interface differs from Pt/HfO_2 in terms of oxygen affinity, interface dipole formation, and band alignment, PdHT devices require a higher CC to reach this filament-dominated regime than PtHT devices. Once both stacks are operated above their respective thresholds, their low-bias conduction is controlled by similar quasi-metallic filaments, which explains the convergence of the extracted Ohmic conduction behavior despite the different CC values.

We employ 10μs pulses to obtain high-fidelity multi-level calibration for system-level SNN evaluation, where precise conductance separability and reproducible energy accounting are required. While shorter programming pulses generally necessitate higher voltages due to the well-known voltage–time trade-off, PdHT devices experimentally retain forming-free bipolar switching down to the nanosecond regime, as demonstrated in Supplementary Fig. 12. This robustness is attributed to the shallow Pd–O–Hf defect landscape that underpins low-voltage, forming-free switching (Figs. 3–4), lowering the activation barrier for defect motion. Consistent with this mechanism, recent forming-free ReRAM demonstrations report sub-μs multi-level operation[6]. Although the present work focuses on μs-scale pulses for accurate multi-level calibration and SNN benchmarking, the combined experimental evidence and interfacial analyses support the compatibility of PdHT devices with high-speed neuromorphic operation, with verify-and-program schemes enabling further voltage scaling in practical arrays.

The integration of these forming-free multibit ReRAM devices into neuromorphic systems has been demonstrated by employing them as synaptic elements in SNNs. Using three-device (3 × 3-bit) configurations to represent 9-bit quantized synaptic weights, the networks are implemented for real-world pattern recognition tasks on the N-MNIST and IBM's DVS128 Gesture datasets. The SNN architectures, developed using frameworks such as SLAYER and PyTorch, achieve competitive classification accuracies (94.6% for N-MNIST and 85.6% for IBM's DVS128 Gesture). However, energy consumption analyses reveal that the PdHT devices yield substantial energy savings, with write and read operations reduced by 43% and 38%, respectively. This energy efficiency, combined with stable multi-level conductance programming and robust endurance, makes the proposed PdNeuRAM technology highly promising for next-generation, low-power neuromorphic computing applications.

Methods

ReRAM cell fabrication

First, the SiO_2 substrate was sequentially cleaned by fuming nitric acid, acetone, isopropyl alcohol, and deionized (DI) water in an ultrasonic oscillator. Then, the photoresist was coated and patterned, and after development, the metal films were deposited on the chemically cleaned SiO_2

substrate through electron beam evaporation in a vacuum chamber with a pressure of 10^{-8} torr. 5 nm Ti (the adhesion layer) and 50 nm Pd (BE) were deposited by AJA ATC 2400 Sputtering System under a vacuum level of 10^{-8} torr; after lift-off process, 5 nm or 10 nm HfO_2 (the oxide layer) was deposited by Sputter (Alliance Concepts system) under a vacuum level of 10^{-7} torr; and it is followed by spin coating and patterning with photolithography; after the development process, 5 nm or 10 nm Ti (interface layer) and 50 nm Pd (TE), which were deposited by ATC 2400 Sputtering System under a vacuum level of 10^{-8} torr. The nodes with dimensions are 5 μm × 5 μm were fabricated by negative photoresist (AZnLof@2020) and positive photoresist (S1805) for photolithography (Heidelberg microMLA), as the profiles demonstrated in Supplementary Fig. 2. The pad dimensions are 150 μm × 150 μm. The BE via was etched by CHF_3/Ar gases (Sentech Etchlab 200) for 270 s. The optical image of the device was obtained by an optical microscope (Olympus BX51) at 50×. Scanning Electron Microscope (Hitachi S-4800) was applied to observe the details of the nodes and cross-section.

Electrical performance characterization

To study the electrical performance of both structure cells, a probe station (CASCADE) equipped with a semiconductor analyzer (Keysight B1500A) was utilized to conduct the electrical measurement for the ReRAM cell at room temperature (300 K), during which a bias voltage was applied to the TE while the BE was grounded. The double linear sweeping bias is from 0 to 1.0/1.2 V with a step 100 mV for the SET sweeping process and 0 to −1.5 V step −100 mV for the RESET sweeping process.

HRSTEM characterization

In our study, we implemented a multifaceted STEM methodology combining HAADF, EELS, and iDPC techniques to achieve a comprehensive characterization of advanced materials. We began with STEM-HAADF imaging, which utilizes a high-angle annular dark-field detector to collect thermally diffuse scattered electrons. This method provides Z-contrast images where the intensity scales roughly as $Z^{1.7}$, allowing us to resolve atomic columns and differentiate heavy and light elements based on their scattering power. Precise alignment and calibration are performed with dwell times in the order of μs to achieve high resolution and reliable compositional mapping.

STEM-EELS was employed to complement the HAADF images by providing elemental and chemical state information through electron energy loss spectroscopy. EELS spectra, collected in parallel with HAADF imaging, enabled the identification of subtle variations in bonding and composition at the nanoscale.

To enhance the detection of light elements and obtain electrostatic potential maps, we applied the iDPC method. In this mode, the segmented DF4 detector is selected via the Velox DPC/iDPC control panel, and the resulting differential signals are integrated. This process improves image contrast and reduces sensitivity to defocus and thickness variations, yielding an image where contrast is roughly proportional to atomic number.

RBS characterization

RBS was employed to determine the elemental depth profiles of the layers in the forming-free PdHT devices. RBS measurements were conducted using a Kobe Steel HRBS-V500 system with a He^+ ion beam accelerated at 400 keV. The beam was directed onto the sample at an incidence angle of 45°, and backscattered ions were detected at a scattering angle of 107.5°. A 512-channel detector with an energy resolution of 2 keV was used. The RBS data was analyzed using the Kobe Steel AnalysisIB software. The measurements revealed the Pd ratio in the HfO_{2-x} layer, confirming the intermixing of Pd-O-Hf. This configuration is critical for reducing the oxygen diffusion barrier and facilitating the forming-free behavior.

Activation energy measurement

The activation energy of the forming-free HfO_2-based ReRAM devices was measured using a Keithley B1500A Semiconductor Parameter Analyzer in conjunction with a thermionic heater. In this setup, the devices were mounted on a temperature-controlled stage that provided stable and uniform heating over a defined temperature range. To prevent oxidation and moisture interference during high-temperature measurements, a continuous flow of high-purity N_2 gas was maintained throughout the experiment. The B1500A recorded the current-voltage (I-V) characteristics at incremental temperature steps. The temperature-dependent resistivity was then derived from the I-V data, and an Arrhenius plot of $\ln(\rho)$ versus $1/T$ was constructed. The slope of the linear fit, when divided by Boltzmann's constant (k_B), yielded the activation energy. This method ensured precise thermal control and reproducible measurements, critical for understanding the conduction mechanisms in the devices.

Device calibration and power consumption evaluation

The device model is developed based on the JART model, which accurately replicates the electrical behavior of the forming-free PdHT ReRAM devices. Calibration of the model is achieved by fitting the experimental DC I-V curves, obtained using a Keisight B1500A, with the simulation results. In our setup, both DC sweep and pulse-based measurement configurations are implemented using Cadence simulation tools. The DC sweep methodology enables the extraction of key parameters such as SET/RESET voltages and memory window, while pulse-based simulations reveal the dynamic switching characteristics and multi-resistance states under various programming schemes. This integrated simulation approach provides critical insights into the energy consumption during programming and reading operations, ensuring consistency between experimental and simulated results.

Device-based system-level SNN simulation

The synaptic weights of the two SNNs are quantized to integer values of 9 bits, which are represented by a combination of three multistate ReRAM devices with 8 states each, i.e., $3 \times 3-$ bit ReRAM devices. In the presented experiments, we performed the training and inference for the two case studies and reported the write and read energy footprint of the PdHT and PtHT ReRAM devices in different scenarios. The focus is kept on the energy impact of the crossbar arrays hosting the devices since it is not dependent on the actual realization of the periphery circuitry, allowing space for generalization, no matter the actual implementation of the neuromorphic hardware accelerator. In all cases, we present a comparison between state-of-the-art PtHT ReRAM devices and the proposed PdHT ones.

N-MNIST SNN: The N-MNIST dataset is a neuromorphic, that is, spiking, version of the MNIST dataset, which comprises images of handwritten arithmetic digits in grayscale format[36]. It consists of 7×10^4 sample images that are generated from the saccadic motion of a Dynamic Vision Sensor (DVS) in front of the original images in the MNIST dataset. The samples in the N-MNIST dataset are not static, i.e., they have a duration in time of 300 ms for each. The dataset is divided into a training set of 6.0×10^4 samples and a test set of 1.0×10^4 samples. The SNN architecture shown in Fig. 6a comprises 3 convolutional layers (SC1, SC2, and SC3) followed by 2 fully connected ones (SF4 and SF5). The classification accuracy on the test set is 94.56%, which is comparable to the performance of state-of-the-art level-based DNNs.

Gesture SNN: The IBM's DVS128 Gesture dataset consists of 29 individuals performing 11 hand and arm gestures in front of a DVS, such as hand waving and air guitar, under 3 different lighting conditions[37]. Samples from the first 23 subjects are used for training, and samples from the last 6 subjects are used for testing. In total, the dataset comprises 1342 samples, each of which lasts about 6 s, making the samples 20 times longer than those in N-MNIST. To speed up the neuromorphic simulations, we trimmed the length of the samples to about 1.5 s. The proposed SNN architecture consists of a pooling layer SP0 to reduce the input samples, 2 convolutional layers (SC1 and SC2), followed by a pooling layer each. The data coming from the last pooling layer is flattened and fed to 2 fully connected layers (SF3 and SF4). The network architecture is presented in Fig. 6b. The network performs with an accuracy of 85.61% in the test set, which is acceptable

considering the shortened samples in the dataset and the shallower architecture compared to the originally proposed architecture in ref. 37.

In both use cases, the SNNs are implemented with SRM spiking neurons, effectively capturing the temporal dynamics of spiking activity. Training uses a variant of the backpropagation algorithm, in which error is calculated on the probability of each neuron to change spike state, i.e., fire a spike if it was in a resting state or stop firing, in the next timing instance[38].

Data availability

The datasets generated and analysed during the current study, including raw I-V measurements, statistical device data, activation energy fitting data, and system-level SNN energy evaluation results, are available from the corresponding author upon reasonable request.

Code availability

The custom code used for device calibration, SNN simulation, and energy evaluation is available from the corresponding author upon reasonable request.

Received: 11 July 2025; Accepted: 15 March 2026;
Published online: 30 March 2026

References

1. Kudithipudi, D. et al. Neuromorphic computing at scale. *Nature* **637**, 801–812 (2025).
2. Hamdioui, S. et al. Memristor based computation-in-memory architecture for data-intensive applications. In *2015 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 1718–1725 (IEEE, 2015).
3. Chen, I.-W. & Lee, J. Ultrathin film resistive memory devices. US Patent Application 14/507,957 (2016).
4. Wang, Y.-B. et al. Forming-free HfO_2-based resistive random access memory by x-ray irradiation. *IEEE Trans. Electron Devices* **69**, 6705–6709 (2022).
5. Tsai, C.-Y., Ting, Y.-W. & Huang, K.-C. Structure and method for a forming-free resistive random access memory with multi-level cell. US Patent 9,053,781 (2015).
6. Park, J. et al. Multi-level, forming and filament free, bulk switching trilayer rram for neuromorphic computing at the edge. *Nat. Commun.* **15**, 3492 (2024).
7. Chen, Y.-S. et al. An ultrathin forming-free hfox resistance memory with excellent electrical performance. *IEEE Electron Device Lett.* **31**, 1473–1475 (2010).
8. Huang, Y.-J., Pan, S. C. & Lee, S.-C. RRAM devices. US Patent 9,525,008 (2016).
9. Stecconi, T. et al. Filamentary taox/hfo_2 reram devices for neural networks training with analog in-memory computing. *Adv. Electron. Mater.* **8**, 2200448 (2022).
10. Wedig, A. et al. Nanoscale cation motion in TaO_x, HfO_x and TiO_x memristive systems. *Nat. Nanotechnol.* **11**, 67–74 (2016).
11. Nukala, P. et al. Reversible oxygen migration and phase transitions in hafnia-based ferroelectric devices. *Science* **372**, 630–635 (2021).
12. Falcone, D. F. et al. All-in-one analog AI hardware: On-chip training and inference with conductive-metal-oxide/HfO_x ReRAM devices. *Adv. Funct. Mater.* **35**, 2504688 (2025).
13. Beckmann, K., Holt, J., Manem, H., Van Nostrand, J. & Cady, N. C. Nanoscale hafnium oxide rram devices exhibit pulse dependent behavior and multi-level resistance capability. *Mrs Adv.* **1**, 3355–3360 (2016).
14. Nakayama, K., Takata, M., Kasai, T., Kitagawa, A. & Akita, J. Pulse number control of electrical resistance for multi-level storage based on phase change. *J. Phys. D Appl. Phys.* **40**, 5061 (2007).
15. Xu, C., Niu, D., Muralimanohar, N., Jouppi, N. P. & Xie, Y. Understanding the trade-offs in multi-level cell reram memory design. In *Proc. 50th Annual Design Automation Conference*, 1–6 (2013).
16. Ding, X. et al. Forming-free HfO_x-based resistive memory with improved uniformity achieved by thermal annealing-induced self-doping of Ge. *IEEE Trans. Electron Devices* **70**, 1671–1675 (2023).
17. Kim, W. et al. Forming-free metal-oxide reram by oxygen ion implantation process. In *2016 IEEE International Electron Devices Meeting (IEDM)*, 4–4 (IEEE, 2016).
18. Wu, E., Ando, T. & Jamison, P. Reversing a decades-long scaling law of dielectric breakdown for reram forming voltage reduction-modeling competition among defect generation and annihilation processes. In *2023 International Electron Devices Meeting (IEDM)*, 1–4 (IEEE, 2023).
19. Ikraiam, F. A. Investigating the effects of ionizing radiation on CV characteristics of SOI-MOS capacitors. In *Proc. 14th Arabic Conference of the Arab Union for Astronomy and Space Sciences: AUASS-CONF23, 13-16 November 2023, Sharjah, United Arab Emirates*, vol. 420, 266 (Springer Nature, 2025).
20. Hardtdegen, A., Zhang, H. & Hoffmann-Eifert, S. Tuning the performance of Pt/HFO_2/Ti/pt reram devices obtained from plasma-enhanced atomic layer deposition for HFO_2 thin films. *ECS Trans.* **75**, 177 (2016).
21. 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).
22. Traoré, B. et al. HfO_2-based RRAM: electrode effects, Ti/HfO_2 interface, charge injection, and oxygen defects diffusion through experiment and *Ab Initio* calculations. *IEEE Trans. Electron Devices* **63**, 360–368 (2015).
23. Mueller, M. P. & De Souza, R. A. SIMS study of oxygen diffusion in monoclinic HfO_2. *Appl. Phys. Lett.* **112**, 052902 (2018).
24. Clima, S. et al. First-principles simulation of oxygen diffusion in HfO_x: Role in the resistive switching mechanism. *Appl. Phys. Lett.* **100**, 133102 (2012).
25. Hou, Z., Gong, X. & Li, Q. Al-induced reduction of the oxygen diffusion in hfo_2: an ab initio study. *J. Phys. Condens. Matter* **20**, 135206 (2008).
26. Öttking, R. et al. Defect generation and activation processes in hfo_2 thin films: contributions to stress-induced leakage currents. *Phys. Status Solidi* **212**, 547–553 (2015).
27. Foster, A. S., Gejo, F. L., Shluger, A. & Nieminen, R. M. Vacancy and interstitial defects in hafnia. *Phys. Rev. B* **65**, 174117 (2002).
28. Waser, R. & Aono, M. Nanoionics-based resistive switching memories. *Nat. Mater.* **6**, 833–840 (2007).
29. Calka, P. et al. Engineering of the chemical reactivity of the Ti/HFO_2 interface for rram: Experiment and theory. *ACS Appl. Mater. interfaces* **6**, 5056–5060 (2014).
30. Tahini, H. A., Chroneos, A., Middleburgh, S. C., Schwingenschlögl, U. & Grimes, R. W. Ultrafast palladium diffusion in germanium. *J. Mater. Chem. A* **3**, 3832–3838 (2015).
31. Chang, K.-C. et al. Physical and chemical mechanisms in oxide-based resistance random access memory. *Nanoscale Res. Lett.* **10**, 120 (2015).
32. Goux, L. et al. Coexistence of the bipolar and unipolar resistive-switching modes in NiO cells made by thermal oxidation of Ni layers. *J. Appl. Phys.* **107**, 024512 (2010).
33. Wang, C. et al. Conduction mechanisms, dynamics and stability in rerams. *Microelectron. Eng.* **187**, 121–133 (2018).
34. Funck, C. & Menzel, S. Comprehensive model of electron conduction in oxide-based memristive devices. *ACS Appl. Electron. Mater.* **3**, 3674–3692 (2021).
35. Bengel, C. et al. Variability-aware modeling of filamentary oxide-based bipolar resistive switching cells using spice level compact models. *IEEE Trans. Circuits Syst. I Regul. Pap.* **67**, 4618–4630 (2020).
36. Orchard, G., Jayawant, A., Cohen, G. K. & Thakor, N. Converting static image datasets to spiking neuromorphic datasets using saccades. *Front. Neurosci.* **9** Article 437 (2015).
37. Amir, A. et al. A low power, fully event-based gesture recognition system. In *IEEE Conference on Computer Vision and Pattern Recognition* (2017).

38. Shrestha, S. B. & Orchard, G. SLAYER: spike layer error reassignment in time. In *Advances in Neural Information Processing Systems*, 1412–1421 (2018).
39. Paszke, A. et al. Pytorch: an imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, 8024–8035 (2019).
40. Gerstner, W. Time structure of the activity in neural network models. *Phys. Rev. E* **51**, 738–758 (1995).
41. Shen, X. et al. Mechanisms of oxidase and superoxide dismutation-like activities of gold, silver, platinum, and palladium, and their alloys: a general way to the activation of molecular oxygen. *J. Am. Chem. Soc.* **137**, 15882–15891 (2015).
42. Tse, K.-Y. & Robertson, J. Control of Schottky barrier heights on high-k gate dielectrics for future complementary metal-oxide semiconductor devices. *Phys. Rev. Lett.* **99**, 086805 (2007).
43. Tung, R. T. Chemical bonding and Fermi level pinning at metal-semiconductor interfaces. *Phys. Rev. Lett.* **84**, 6078 (2000).
44. Abunahla, H. et al. Switching characteristics of microscale unipolar PD/hf/hfo$_2$/pd memristors. *Microelectron. Eng.* **185**, 35–42 (2018).
45. Puglisi, F. M., Pavan, P., Padovani, A. & Larcher, L. Perimeter and area current components in HfO_2 and HfO_{2-x} metal-insulator-metal capacitors. *J. Vac. Sci. Technol. B* **31**, 01A101 (2013).
46. Kaniselvan, M. et al. Electroforming kinetics in hfo x/ti rram: mechanisms behind compositional and thermal engineering. *ACS Nano* **19**, 27455–27466 (2025).
47. Napari, M., Stathopoulos, S., Prodromakis, T. & Simanjuntak, F. Forming-free and non-linear resistive switching in bilayer HfO_x/TaO_x memory devices by interface-induced internal resistance. *Electron. Mater. Lett.* **20**, 363–371 (2024).
48. Ranjan, A. et al. Probing resistive switching in hfo$_2$/al2o3 bilayer oxides using in-situ transmission electron microscopy. *Appl. Mater. Today* **31**, 101739 (2023).
49. Ju, D., Kim, S., Jang, J. & Kim, S. Improved uniformity of taox-based resistive switching memory device by inserting thin sio2 layer for neuromorphic system. *Materials* **16**, 6136 (2023).
50. Athena, F. F., West, M. P., Hah, J., Graham, S. & Vogel, E. M. Trade-off between gradual set and on/off ratio in hfo x-based analog memory with a thin sio x barrier layer. *ACS Appl. Electron. Mater.* **5**, 3048–3058 (2023).
51. Urquiza, M. L., Islam, M. M., Van Duin, A. C., Cartoixà, X. & Strachan, A. Atomistic insights on the full operation cycle of a hfo$_2$-based resistive random access memory cell from molecular dynamics. *ACS Nano* **15**, 12945–12954 (2021).
52. Yang, Z. et al. Neural univariate activity and multivariate pattern in the posterior superior temporal sulcus differentially encode facial expression and identity. *Sci. Rep.* **6**, 23427 (2016).
53. Luo, Q. et al. Self-rectifying and forming-free resistive-switching device for embedded memory application. *IEEE Electron Device Lett.* **39**, 664–667 (2018).
54. Kim, W. et al. Lowering forming voltage and forming-free behavior of ta 2 o 5 reram devices. In *2016 46th European Solid-State Device Research Conference (ESSDERC)*, 164–167 (IEEE, 2016).
55. Wang, Z. et al. Electroforming-free artificial synapses based on proton conduction in α-moo3 films. *Adv. Electron. Mater.* **6**, 1901290 (2020).
56. Sun, Y. et al. Hybrid channel induced forming-free performance in nanocrystalline-si: H/a-sinx: H resistive switching memory. *Nanotechnology* **30**, 365701 (2019).
57. Chakrabarti, B., Galatage, R. V. & Vogel, E. M. Multilevel switching in forming-free resistive memory devices with atomic layer deposited $HfTiO_x$ nanolaminate. *IEEE Electron Device Lett.* **34**, 867–869 (2013).
58. Kumar, S. & Rath, C. Oxygen vacancy mediated stabilization of cubic phase at room temperature and resistive switching effect in sm- and dy-doped HfO_2 thin film. *Phys. Status Solidi* **217**, 1900756 (2020).

Acknowledgements

This work was supported by Delft University of Technology. B.N., M.A., and E.v.d.V. acknowledge the financial support of the CogniGron research center and the Ubbo Emmius Funds (Univ. of Groningen). We acknowledge the kind support of Ali Kaichouhi for the measurement setup, and we also thank the Delft Kavli Nanofabrication cleanroom, especially Marinus Fischer, Roald van der Kolk, Charles de Boer, and Eugene Straver, for their unconditional assistance. We also specially thank Lennart P.L. Landsmeer and Emmanouil Arapidis for their kind support and discussion.

Author contributions

E.H., H.A., and R.I. conceived the study. E.H. performed device fabrication and electrical measurements. E.H., H.A., R.I., S.H., and G.G. performed formal analysis and data interpretation. M.A., B.N., E.H., E.v.d.V., and H.A. conducted STEM, EELS, and iDPC characterization and analysis. E.H., B.N., and H.A. performed AFM measurements and activation energy analysis. E.H. and H.X. developed and calibrated the device compact model. T.S., E.H., H.A., and A.G. implemented the SNN simulations and performed system-level energy analysis. A.M.S., E.H., H.A., L.B., and N.E. carried out RBS characterization and analysis. N.W. performed SIMS measurements. E.H., B.N., T.S., and H.A. wrote the original manuscript draft. All authors reviewed, edited, and approved the final manuscript. H.A. and R.I. jointly supervised this work.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44172-026-00650-3.

Correspondence and requests for materials should be addressed to Erbing Hua or Heba Abunahla.

Peer review information *Communications Engineering* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: [Wenjie Wang, Rosamund Daw]. A peer review file is available.

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

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026