ACS NANO

www.acsnano.org

ARTICLE

Leveraging Tunability of Localized-Interfacial Memristors for Efficient Handling of Complex Neural Networks

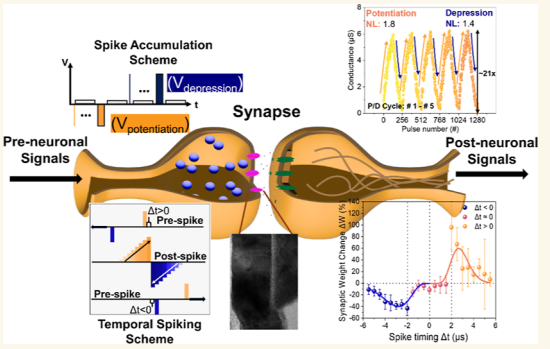
Eng Kang Koh, Putu Andhita Dananjaya, Lingli Liu, Calvin Xiu Xian Lee, Gerard Joseph Lim, Young Seon You, and Wen Siang Lew*

Cite This: *ACS Nano* 2024, 18, 29602–29617

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: **A scalable (<130 nm) resistive switching memristor that features both filamentary and interfacial switching aimed at neuromorphic computing is developed in this study. The typically perceived noise or volatility was effectively harnessed as a controlled mechanism for interfacial switching. The multilayer structure for the proposed memristor enhances switching stability by curbing ionic overmigration and mitigating leakage paths. Furthermore, the memristors showcased their reliability by demonstrating more than 15 M cycles in the filamentary mode and 1 M pulses in the interfacial mode. Additionally, retention tests at 85 °C for 10^4 s confirmed the stability across different states, affirming its reliability as a nonvolatile CMOS-compatible element. While many studies validate performance solely on the MNIST data set, this work also evaluates more complex data sets, demonstrating the robustness of the demonstrated memristor in supervised learning. Specifically, supervised learning simulations on MNIST and fashion MNIST data sets indicated a high learning rate with <4% deviations from numerical training, while offline inference trained on CIFAR-10 and CIFAR-100 data sets revealed <2.5% and <7% deviations caused by programing error accumulation, even with increased memristor counts for these highly complex data sets. Unsupervised learning via spike-timing-dependent plasticity further highlights the potential of the developed memristor in bridging artificial and biological paradigms, offering a significant advance toward efficient and biologically inspired computing architectures.**



KEYWORDS: *memristor, filamentary switching, interfacial switching, artificial neural networks, spiking neural networks*

In the ever-evolving pursuit of emerging memory technologies, the resistive switching memory, or memristors, stands out for its ability to store data by altering its resistance based on the applied external electric field.[1,2] This makes it suitable for a myriad of applications, from data storage to neuromorphic applications, positioning it as a viable successor to existing computing architectures.[3−5] The conventional computing systems, structured around separate processing and memory units, often face challenges in efficiently handling complex data sets typical of artificial intelligence (AI) tasks. These tasks range from supervised learning, where models learn from labeled data, to unsupervised learning, which involves identifying patterns in unlabeled data.[6−8] Memristors facilitate in-memory computing and promise a significant reduction in energy consumption and processing time, offering a better fit for modern AI demands.[9−11] There is also potential for this efficiency to be further enhanced by recent advancements that allow memristors to exhibit both filamentary and interfacial characteristics within a single device, which was previously reported as mutually exclusive.[12−17] The coexistence of these switching modes allows the memristors to broaden their applications to accommodate more of the aforementioned computational tasks. Specifically, the dual mode functionality of these devices not only accommodates binary logic essential for general computational tasks, as seen in filamentary devices, but also facilitates analogue computation akin to the biological brain's synaptic activities, characteristic of the interfacial mode. However, most of such research often framed interfacial switching as a noise disruption or found it to be volatile, making it unsuitable for specific AI applications such as edge

Received: June 5, 2024
Revised: October 9, 2024
Accepted: October 11, 2024
Published: October 19, 2024

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

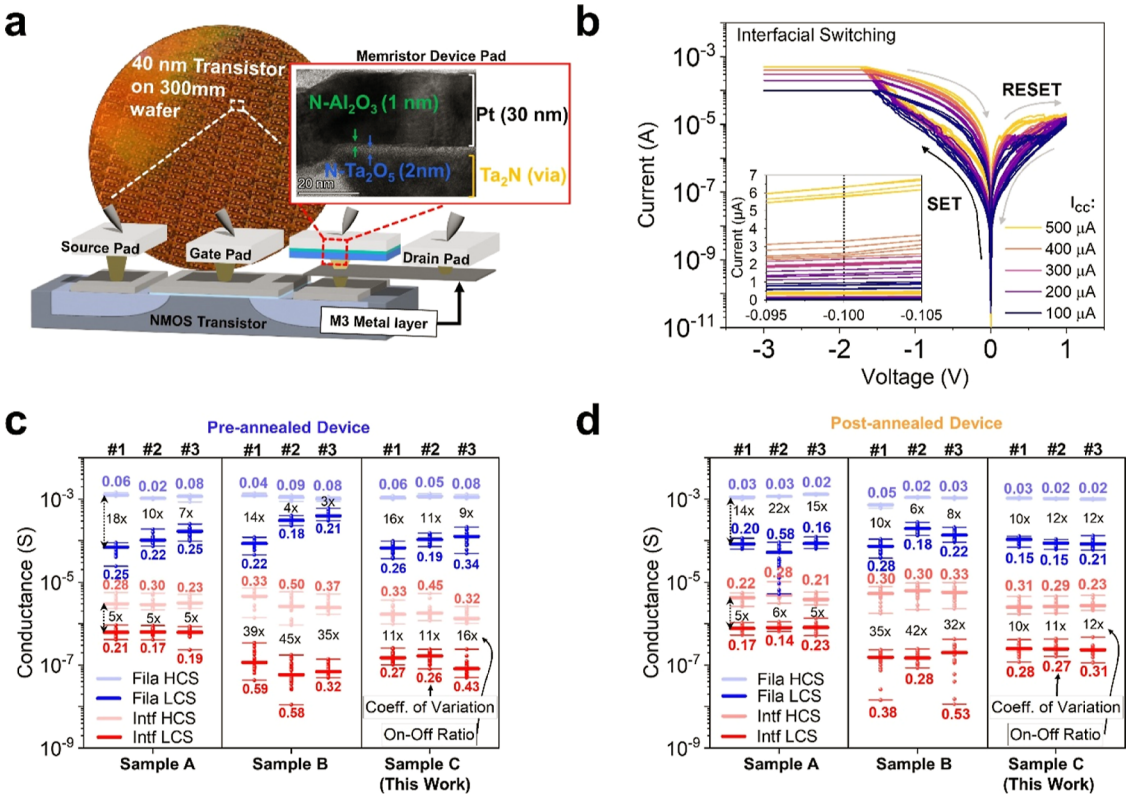


Figure 1. (a) Schematic of the 1T−1R device, not drawn to scale. The magnified area highlights the cross-sectional TEM image of the memristive device structure. Multilevel switching capabilities of the device through (b) interfacial switching compliance current variation (100−500 μA). (c,d) Pre- and postannealed device-to-device (3 devices) and cycle-to-cycle variation comparing resistance levels for control samples showing the non-nitrogen-doped devices [sample A − Pt (30 nm)/Al_2O_3 (1 nm)/Ta_2O_5 (2 nm)/Ta_2N (VIA)], single switching layer devices [sample B − Pt (30 nm)/N−Ta_2O_5 (3 nm)/Ta_2N (VIA)], and optimized devices in this work [sample C − Pt (30 nm)/N−Al_2O_3 (1 nm)/N−Ta_2O_5 (2 nm)/Ta_2N (VIA)].

computing and neural net processing where data persistence is required.[18−21] Furthermore, numerous studies have yet to validate the performance of their memristors against the specific applications for which they are purportedly suitable.[22−28] Additionally, the efficacy of many other works is asserted based on the high accuracy rates achieved by training their simulated memristors on only the MNIST data set.[29−32] The MNIST data set, with its relatively simple and well-understood set of grayscale images of handwritten digits, offers a convenient starting point for demonstrating the capability of developed computational architectures, including those based on memristors, to perform supervised pattern recognition tasks.[33] However, relying solely on MNIST for benchmarking presents several limitations that could hinder advancement and understanding of the true potential of memristive devices for more complex AI applications. This concern is due to the simplicity of the data set, which may not accurately reflect the performance of memristive devices on more challenging tasks.

This study aims to develop a memristor combining both filamentary and interfacial switching with a focus on state stability. Additionally, it seeks to meet the elevated performance standards required by such a dual-switching device compared with its single-switching mode counterparts. The key focuses include reliability factors, such as state variations, endurance, and retention. There is also an evaluation of how the presented memristor meets these criteria, which includes a supervised learning simulation for performance comparison against an ideal memristor array using more complicated data sets like CIFAR-10 and CIFAR-100. Finally, it investigates the memristor's versatility in potentially enabling unsupervised learning. This not only demonstrates the feasibility of dual-mode memristive devices for complex data processing and learning tasks but also sets a foundation for future research in biologically inspired computing architectures.

RESULTS AND DISCUSSION

Electrical Characterizations. In resistive switching, bringing the state of the device from a low conductive state (LCS) to a high conductive state (HCS) is also known as the set operation. Conversely, the reset operation returns the device from an HCS to an LCS.[34−36] Typically, these devices are characterized electrically using a DC double-sweep; the first sweep in one polarity acts to set the device, and the opposite polarity sweep resets the device. This essentially enables the elucidation of the resistive switching characteristics of the device through a pinched hysteresis loop that is commonly observed due to the change in device conductance during the sweep-forward and sweep-back state. The memristive structures presented in this work exhibit different switching characteristics depending on the operation scheme applied to the memristor. Figure 1a presents a schematic showing the integration of a Pt/N−Al_2O_3 (1 nm)/N−Ta_2O_5 (2 nm)/Ta_2N memristor structure connected to a 40 nm NMOS transistor at its drain. The magnified image depicts a transmission electron microscopy (TEM) image, confirming the deposited layer thicknesses. Moreover, Figure S1a shows the scanning electron

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

microscopy (SEM) image taken after the focused ion beam cut, revealing the entire cross section of the device. The integrated device consists of the transistor fabricated using the back-end-of-line process and the memristor integrated post-back-end-of-line. This memristive device, highlighted in Figure S1a, has a diameter of approximately 130 nm. Further magnification of this region, presented in Figure S1b, reveals the detailed EDX mapping of key elements, including O, Al, Ta, Pt, and N, which confirms the presence and distribution of the deposited layers. The EDX line scan, corresponding to the direction indicated in Figure S1b, is shown in Figure 1c, verifying the layer thicknesses of N−Ta_2O_5 and N−Al_2O_3 to be approximately 2 and 1 nm, respectively. These TEM and EDX analyses validate the structural composition of the device. During operation, the voltage was constantly applied with respect to the Ta_2N vertical interconnect access (VIA) while the memristor drain pad was always grounded. In other words, for 1-resistor (1R) electrical characterizations, the probes were only on the memristor device pad (ground) and the drain pad. In contrast, for 1 transistor−1 resistor (1T−1R) characterizations, the source, gate, and memristor device pad (grounded) were probed. All unused pads remained floating. In this work, all DC measurements were done using 1R while all pulsed measurements were completed by utilizing the 1T−1R device. In the integrated 1T−1R device, the transistor gate was mainly leveraged as a current limiter for the memristor. From a pristine state, applying a negative voltage on the memristor switches the device interfacially and sets the device to an interfacial HCS. This interfacial switching can be seen in Figure 1b, by which the current increases gradually as the voltage applied increases negatively (interfacial set process), and subsequent positive DC sweep induces a current decrease (interfacial reset process). Simultaneously, on the same device, filamentary switching is observed (see Figure S2a,b) when a positive forming voltage is applied to soft break down the device. This forming process is utilized to initiate switching in filamentary devices in which a higher voltage is generally required. In this work, forming is employed to switch the device to a filamentary HCS. This typically only happens during the first filamentary cycle and can be applied either in its pristine or interfacial state. Filamentary switching is typically characterized by the observation of the device's abrupt change in current as the positive voltage applied is increased. The origin of the abrupt current change in filamentary switching is known to be driven by the localized ionic migration, where oxygen vacancies are left behind, and they facilitate electron hopping. This reduces the effective resistance of the device. The localized region of accumulated ions effectively creates a "virtual electrode" within the switching layer, enhancing the local electric field. As this happens, the migration of ions accelerates, leading to a positive feedback loop until a fully conductive path connects both electrodes. This positive feedback loop is a direct consequence of the abrupt current change observed in localized filamentary switching. Conversely, interfacial switching does not exhibit this abrupt current change because its mechanism involves more uniform ionic modulation across the interfaces rather than the localized process described in filamentary switching. Hence, the self-reinforcing positive feedback effect that causes abrupt current changes in filamentary switching above does not occur in interfacial switching. This distinction is quantitatively reflected in the switching windows portrayed in Figure S2c, where the filamentary switching exhibited a median 0.439 mA current change over five voltage sweeps of 0.08 V. Comparatively, the interfacial switching exhibited a similar current increase of 0.453 mA over a much larger voltage sweep of 0.5 V.

The operating parameters of the selected structure were evaluated by 1R characterization. The parameter analyzer was also configured to investigate the multilevel cell (MLC) switching capabilities of the memristor. Generally, two methods can be utilized to achieve the MLC properties, either by applying different I_{CC} or by varying the reset voltage. Both methods work by manipulating the change of the ion dynamics within the switching layer/s to alter the resistance of the device, and the effects on filamentary MLC have been shown in Figures S1a,b, respectively.[37−39] The MLC capabilities of the device under the interfacial switching scheme are depicted in Figure 1b. This interfacial switching is observed on the device preforming, and the I_{CC} was limited from 100 to 500 μA with an increment of 100 μA. The different states in Figure 1b were stored as a single interfacial LCS and five different interfacial HCS corresponding to the different applied I_{CC}. Figure S2a depicts the device under ten different I_{CC}, limited from 0.1 to 1.0 mA in increments of 0.1 mA. This resulted in 11 different states, corresponding to those HCS produced by the I_{CC} and one LCS. Figure S2b depicts seven different MLCs (one HCS and six LCS) corresponding to the different reset voltages, varying from −0.5 to −1.0 V, with an increment of −0.1 V. It is also observed at the filamentary reset region that the device seems to undergo a reset voltage saturation past −1.0 V by which the current started to increase again instead of decreasing. Under all three schemes discussed in Figures 1b and S1a,b, ten continuous cycles of double sweeps were performed for each I_{CC} or reset voltages were applied. It should be noted that in all these figures, the inset depicts the MLC that could be obtained at a read voltage of 0.1 V and that none of those states overlapped with each other. This is crucial because nonoverlapping states are required for optimal neuromorphic computing operations. Moreover, it was also noticed that increasing the thickness of oxide layers of the memristor, specifically by maintaining the N-doped Ta_2O_5 (N−Ta_2O_5) and N−Al_2O_3 (N-doped Al_2O_3) ratios at 2:1, can significantly reduce the operating current, leading to better energy efficiency. In devices where the oxide layers' thicknesses were doubled, there was a reduction of current by more than 100 times, as illustrated in Figure S2d.

In our previous study, a device featuring Pt/Ta_2O_5/Ta_2N memristor structure, in which the focus was on unraveling the switching mechanism in both switching modes.[40] Building on the underlying physics outlined in the previous research, this work endeavors to advance the performance across various applications, incorporating several pivotal processes. These enhancements are critical for realizing the full potential of applications from memory devices to neuromorphic chips. First, from the left to right column, Figure 1c depicts the preannealed non-nitrogen-doped control devices [sample A − Pt (30 nm)/Al_2O_3 (1 nm)/Ta_2O_5 (2 nm)/Ta_2N (VIA)], single switching layer control devices [sample B − Pt (30 nm)/N−Ta_2O_5 (3 nm)/Ta_2N (VIA)], and the optimized devices presented in this work, which consists of both nitrogen-doped and bilayer devices [sample C − Pt (30 nm)/N−Al_2O_3 (1 nm)/N−Ta_2O_5 (2 nm)/Ta_2N (VIA)] read at 0.1 V after 25 cycles in both filamentary (blue) and interfacial (red) LCS and HCS. The interfacial states were achieved through DC cycling, utilizing −3.0 V for the set process and 1.0 V for the reset process with an I_{CC} of 100 μA. The filamentary states were also

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

obtained via DC cycling, utilizing 2.0 V for the set process and −1.25 V for the reset process with an I_{CC} of 500 μA. These DC cycles are depicted in Figures S3−S8. Figure 1d shows the same devices from Figure 1c after annealing at 400 °C for 1 h under vacuum conditions at approximately 5 × 10^{-8} Torr. In general, devices in sample C showed an increase in the on−off ratio compared to sample A in the interfacial mode, with the preannealed devices increasing the ratio from ∼5× to >11× and postannealed devices increasing from ∼5× to >10× (measured from median to median). Next, the devices in sample B, although it shows a higher on−off ratio in the interfacial mode compared to sample C (>35× compared to >11× for preannealed devices and >32× to >10× for postannealed devices), it is worse off in the filamentary mode (>3× compared to >9× for preannealed devices and >6× to >10× for postannealed devices). Moreover, annealing further improves the cycle-to-cycle and device-to-device variation of the devices, which is compared by utilizing the coefficient of variation (CV, $\sigma/\overline{x}$) for each mode and their respective LCS and HCS. Furthermore, the forming voltage of devices in samples A−C was also measured, and the results are presented in Figure S9. An observation could be made that the forming voltage decreased across all samples after annealing. Thermal treatment under vacuum accelerates the reduction of N−Ta_2O_5 and promotes the oxidation of Ta_2N, resulting in a thicker TaO_xN_y layer.[41,42] This process reduces forming voltage and has been shown to enhance dopant activation in nitrogen-doped TaO_x-based firms, contributing to the improved switching stability in both LCS and HCS, as mentioned previously.[43,44] The range of forming voltages is shown in Figure S9 with a median of ∼1.6 V and a standard deviation of 0.04 V. Sample C (the optimized structure presented in this work) is also observed to be the structure with the smallest forming voltage distribution compared with samples A and B, with an acceptably forming voltage magnitude. A uniform forming voltage is arguably more important than the magnitude, as the former leads to more consistent switching behavior across devices, which is critical for applications like memory and neuromorphic computing. This also ensures that all devices in the array can be activated without excessively high voltages, which could damage some devices or fail to form them. This is the key to producing high-density arrays or neuromorphic networks. Hence, the specific structure in sample C was chosen for this work after considering the trade-offs between the on−off ratio, variability, and forming voltage. Comparing sample A and sample C, N-doping has been well-documented to improve the variability of filamentary memristors. Misha et al. highlighted that optimum doping of nitrogen could mitigate leakage paths by occupying more oxygen vacancies while Sedghi et al. supported this claim through ab initio calculations, demonstrating that nitrogen doping eliminates midgap defect states caused by vacancies, thereby lowering the density of vacancy-induced defect states.[45,46] From the results in Figure 1c,d, it is hence hypothesized in this work that N-doping also aids in the mitigation of the leakage path in the interfacial mode. Moreover, the state stability enhancement suggests that the implementation of the bilayer leads to an interfacial layer that not only confines filament formation but also acts as a barrier to preventing the overmigration of ions.

Revelation of Switching Mechanism by Utilizing XPS Depth Profiling. By comparing control samples, insights were garnered on the adjustable parameters during the memristor optimization, specifically those with bimodal switching characteristics. Hence, it is paramount to achieve a comprehensive understanding of the switching mechanisms in the optimized structure. The X-ray photoelectron spectroscopy (XPS) depth profile analysis was conducted to ascertain the concentration of elements at specific depths within the device. Owing to the inherent constraints of XPS, enlarged devices were fabricated with dimensions of 100 μm in diameter and switching layers of 16 and 8 nm in thickness for N−Ta_2O_5 and N−Al_2O_3, respectively. The selection of these specific device sizes and layer thicknesses was critical to ensure consistency in resistance levels during XPS measurements compared to the device used in this study. As smaller devices do not generate sufficient signal for effective XPS measurements, the device diameter had to be scaled up (130 nm to 100 μm). To accommodate this larger device diameter needed for XPS measurements, the oxide thickness was also scaled up together. This was essential because increasing only the diameter would have increased the current flowing through the device, which would have diminished the switching window between different states. Increasing the device thickness [N−Al_2O_3(1 nm)/N−Ta_2O_5(2 nm) to N−Al_2O_3(8 nm)/N−Ta_2O_5(16 nm)], conversely, as shown previously in Figure S2d, increases the current across the device. This adjustment of both variables was optimized to effectively regulate the switching resistance, regardless of the device size. This technique is commonly employed in studies on scalable devices, particularly when the devices are too small for effective XPS signal detection.[15,47] The DC $I-V$ curves of both the XPS sample and the device used in this study, as shown in Figure S10 (with an inset depicting their dimensions—the XPS sample with a diameter of 100 μm and the device at 130 nm, drawn not to scale), corroborate this approach. They exhibit similar electrical characteristics despite the scaling, validating the method of simultaneously adjusting the device's thickness and diameter to maintain its switching behavior. Specifically, the process of filamentary switching is primarily governed by localized defect sites and conductive filament forming within a confined switching depth. Additionally, after forming, the device's switching depth is restricted to regions where the oxygen vacancies are generated and recombined. Yu et al. described this as the "active layer" of their TiN/metal oxide/Pt device.[48] They attributed the active layer to having a smaller oxygen migration barrier than that of the other switching layers. As a result, the filamentary switching mechanism should not fundamentally alter due to device scaling since the critical processes occur within a localized region that remains consistent despite changes in overall device dimensions. Similarly, in the case of interfacial switching, scaling both the oxide thickness and the device diameter ensures that the switching mechanism is preserved. This careful scaling approach confirms that the switching behavior observed in devices with different thicknesses is consistent, validating the use of the XPS sample for mechanistic insights. Three different devices of different states on the same sample were prepared and measured for this experiment: a pristine (doubled as the interfacial LCS, depicted in black), an interfacial HCS (in red), and a filamentary HCS device (in blue). The fitting of the XPS spectra for O 1s and Ta 4f are presented in Figures S11 and S12, respectively. The deconvolution of O 1s peaks resulted in lattice and nonlattice oxygen peaks (528−536 eV), and the Ta 4f peaks yielded five doublet peaks corresponding to different states of tantalum oxide and its compounds (20−32 eV).[49−53]

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

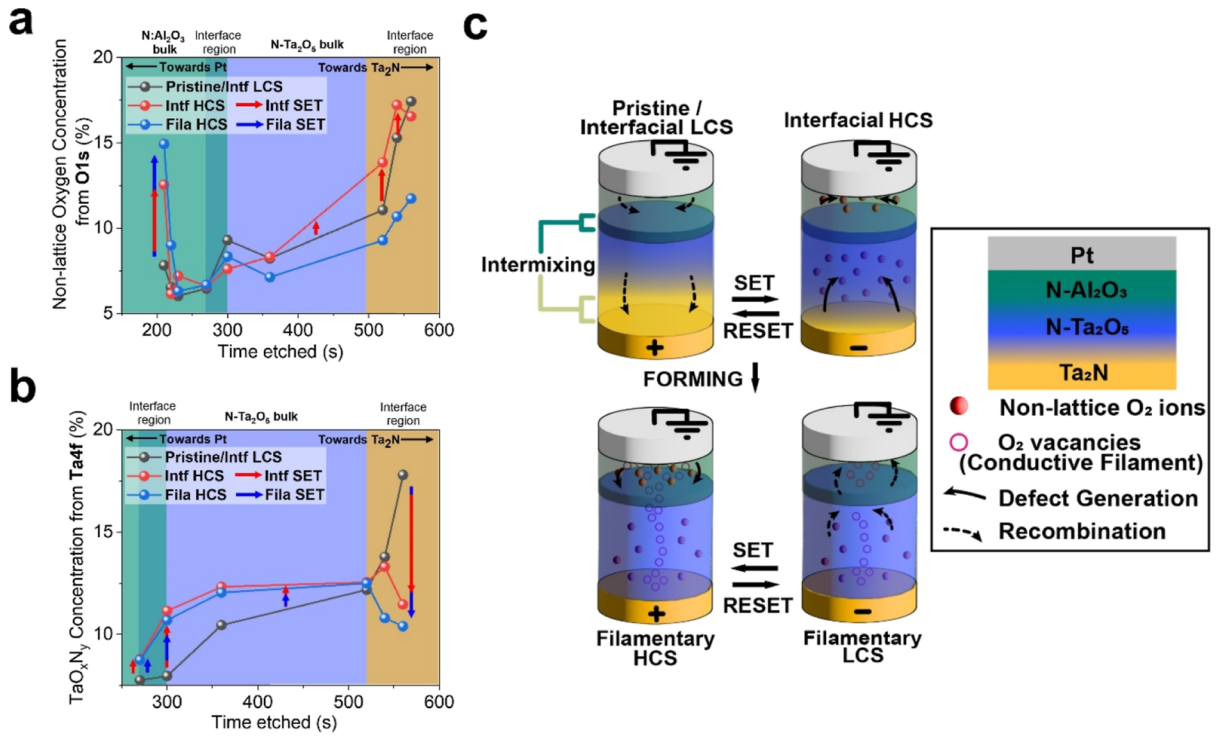


Figure 2. Percentage concentration of (a) nonlattice oxygen from O 1s, and (b) TaO_xN_y from Ta 4f as a function of etching time on an upscaled device obtained by XPS depth profiling, raw data presented in Figures S11 and S12. (c) Schematic of Pt/N−Al_2O_3/N−Ta_2O_5/Ta_2N memristor during both interfacial and filamentary switching.

Figure 2a,b illustrates the percentage concentrations of nonlattice oxygen and TaO_xN_y as they vary with the etching duration of the devices, respectively. The analysis of the TaO_xN_y profile in Figure 2b showed a gradually increasing trend from the N−Al_2O_3/N−Ta_2O_5 interface to the N−Ta_2O_5/Ta_2N interface. Initially, from the pristine state, a positive forming voltage switched the device into its filamentary HCS. As seen in Figure 2a, this transition led to an increase in nonlattice oxygen by around 7.5% within the N−Al_2O_3 bulk. Concurrently, Figure 2b indicates a decrease in TaO_xN_y bonds at the Ta_2N and N−Ta_2O_5 interface, mirroring the increase in nonlattice oxygen. Figure 2c shows the device switching mechanisms based on the XPS results. After forming, the device transitions to a filamentary HCS. From the pristine to the filamentary HCS, the highest increase of the nonlattice oxygen concentration was located within the N−Al_2O_3 bulk, indicating that majority of the defect generation during the filamentary set process occurred in that region. Furthermore, there was an increase in the nonlattice oxygen concentration between the N−Al_2O_3 bulk and N−Al_2O_3/N−Ta_2O_5 interface, which decreased below the interface, highlighting the interface's role in the switching process. This dynamic proves the hypothesis that the interfacial region acts as a critical barrier, modulating the nonlattice oxygen movement and influencing the device's stability during switching events. During filamentary reset, the filament tends to rupture at the weakest point, near the site with the least nonlattice oxygen concentration. This area is also where the nonlattice oxygen ions and oxygen vacancies are recombined.

To explain the interfacial mechanism, from Figure 2a, applying a negative voltage during the set process led to the generation of nonlattice oxygen in the interfacial region between Ta_2N and N−Ta_2O_5 bulk (around 2.5%), as well as in the N−Al_2O_3 bulk (around 5%). Conversely, applying a positive interfacial reset voltage promotes recombination. Figure 2b shows a 7% decrease in the TaO_xN_y concentration in the N−Ta_2O_5/TaO_xN_y interfacial region, while an increase then diminishing decrease in the concentration within the N−Ta_2O_5 bulk during the interfacial set. Referring to Figure 2c, these changes in compound percentages can be interpreted as the modulation of the N−Ta_2O_5/TaO_xN_y interfacial layer thickness during switching. Additionally, when the negative voltage is applied during the interfacial set, nonlattice oxygen ions are expected to move toward the direction of the Pt electrode. However, the decrease of the disparity in TaO_xN_y content between the pristine and interfacial HCS devices again signals the existence of a barrier at the N−Al_2O_3/N−Ta_2O_5 interface, as it is difficult for nonlattice oxygen to migrate across this barrier. This is also evident in the nonlattice oxygen percentage of the interfacial HCS device dipping below that of the pristine device toward the N−Al_2O_3/N−Ta_2O_5 interface. The XPS depth profiling analysis can be used in conjunction with Figure 1c,d to correlate ionic mechanisms during filamentary and interfacial switching to resistance levels during device operations. Most importantly, in both figures, the LCS in both modes of the single-layer control devices consistently showed higher CV compared with optimized devices. This suggests a link to the observation from the XPS depth profiling results, where the oxide bilayer's interface acts as an ion migration barrier, potentially increasing the rate of return to initial states in successive cycles.

Pulse Operations of Localized-Interfacial Memristors and Reliability. Following the unraveling of the switching mechanism of the memristors, the devices in their filamentary and interfacial modes were investigated for their use in binary and analogue switching capabilities. This was done through the

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

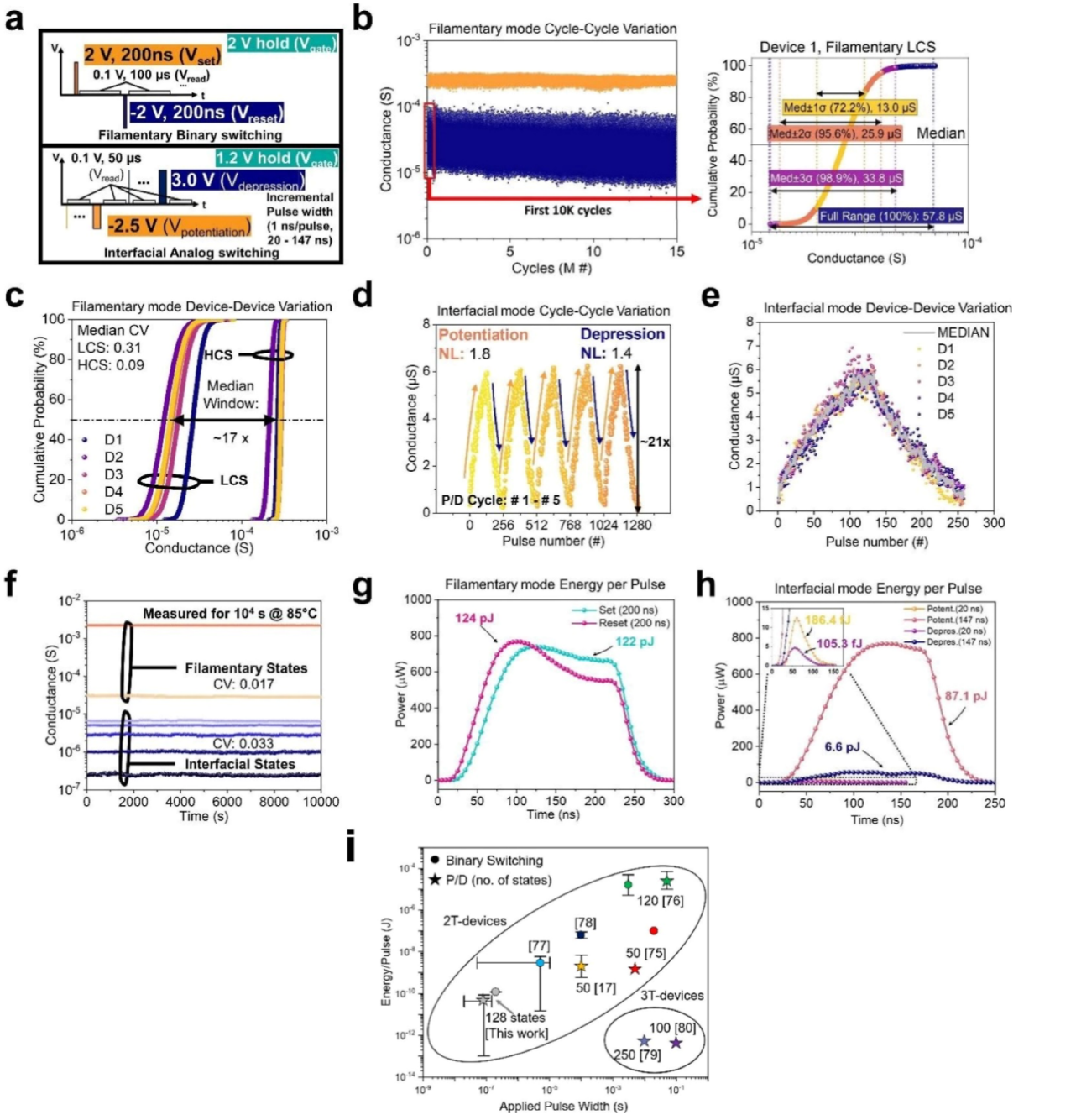


Figure 3. (a) Schematic of pulse schemes applied on the 1T−1R memristor during filamentary binary switching and interfacial analogue switching. (b) Operation of memristor in its filamentary mode showcased an endurance of over 15 million cycles shown in the form of cycle-to-cycle variation. The first 10k cycles were analyzed in a cumulative probability plot. (c) The device-to-device switching of five devices also shows a similar median CV, resulting in a median on−off window of ~17×. A different operating method was utilized for the interfacial mode, and cycle-to-cycle variation in the form of five full potentiation and depression cycles is shown in (d) and up to >3800 cycles with >1 million pulses is shown in Figure S15. The device-to-device variation of five devices is shown in (e). (f) Retention study of two and five filamentary and interfacial states were measured at 85 °C for a total of 10^4 s. Energy consumption per pulse of their respective applications in the (g) filamentary mode and (h) interfacial mode. (i) Benchmark of the energy consumption per pulse against applied pulse width against comparing the device performance of this work against the memristors developed in other works in terms of binary switching and P/D applications.

subjection of the 1T−1R devices under electrical pulses, which, unlike DC measurements presented previously, is what is needed to operate the device in real-world applications.[54] Nevertheless, while pulse measurements are essential for determining the practical performance, the DC $I-V$ curves depicted in Figure S13 still provide critical insights into the 1T−1R configuration. These insights help us understand how to best operate the device under pulsed conditions. Figure S13 illustrates the DC $I-V$ curves for the 1R and 1T−1R configurations and the output characteristics of the 1T, with the transistor operating gate voltages, V_G = 2.0 and 1.2 V highlighted for the filamentary and interfacial schemes, respectively. The black line represents the V_G applied to the transistor in the 1T−1R setup (shown as a blue line). The 1R

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

configuration provides insight into the basic resistive switching characteristics of the device with minimal external control (apart from the externally applied compliance current) while the 1T setup demonstrates the role of the transistor in controlling the current flow. The 1T−1R configuration, which integrates both a transistor and a memristive switching element, offers several significant advantages. First, the transistor allows for improved current control by adjusting the V_G to limit the current through the drain and the memristor element. However, a difference is observed between the compliance current set by the parameter analyzer and that of the transistor gate. As shown in Figure S13, the compliance current applied by the parameter analyzer precisely restricts the current across the integrated device to around 0.3 mA for filamentary switching (Figure S13a) and 0.2 mA for interfacial switching (Figure S13b). In contrast, in the 1T−1R configuration, the current continues to increase even after the device switches to the current level of the 1R, indicating that both V_G and V_{DS} are crucial parameters for controlling device switching. The integration of the 1T−1R also reduces overshoot current, particularly in filamentary switching, as seen in Figure S13a, though this is less critical for interfacial switching shown in Figure S13b. Notably, in the filamentary 1R configuration, the highest reset current consistently exceeds the highest set current, while in the 1T−1R configuration, these currents are approximately equal. This suggests more controlled filament formation in the filamentary mode, leading to improved reliability. Additionally, the 1T−1R setup enhances scalability due to the presence of the transistor, which allows for internal current control. Without this control, the 1R device could fail due to excessive current flow. By comparing these configurations, we gain a deeper understanding of the set and reset processes, the variability in switching behavior, and overall device reliability. The electrical pulse schemes used in this work are presented in Figure 3a for operating in both the filamentary mode to obtain binary switching and the interfacial mode to demonstrate analogue switching. Figure 3b,c shows the cycle-to-cycle and device-to-device variations of the devices in their filamentary mode. The filamentary pulsed operation scheme in Figure 3a shows the set and reset operations conducted through 2.0 and −2.0 V pulses for 200 ns each while holding a 2.0 V on the transistor gate. In comparison, the read operation was conducted by utilizing a long 100 μs pulse at 0.1 V to ensure accuracy. Filamentary switching contributes to the high endurance due to its localized conductive paths. The amount of ionic movement is kept to a minimum during switching, which leads to fewer switching failures. Figure 3b demonstrates the excellent endurance of the memristor, exceeding 15 million cycles in this mode. At first glance, the LCS variation in the endurance plot appears to be large, seemingly exceeding the on−off ratio (when calculated from the lowest HCS to the highest LCS). Moreover, this observation aligns with the existing literature, where the variation of the LCS is typically larger than that of the HCS.[55−60] This can primarily be attributed to two factors. First, the absence of current control in the LCS plays a significant role. During the set process, the transistor gate is utilized to halt any further increase in the current flowing through the device. This current control is absent in the reset process, where the only control is termination of the reset voltage. This results in the LCS becoming more susceptible to temporal or spatial fluctuations.[61,62] Second, noise is also a well-documented problem in scalable RRAM devices, with smaller devices leading to higher noise levels, which in turn leads to greater variability.[63] Furthermore, in the filamentary LCS, the noise from the various current paths within the broken filament accumulates, further contributing to the observed variability.[63,64] This observation might suggest significant cycle-to-cycle variation in the off-state during extended cycling. However, this perceived variability is largely due to the comprehensive presentation of data across all 15 million cycles rather than selectively displaying results, such as displaying every nth cycle.[60,65−68] This approach allows for the precise identification of failure points, with failure defined as the overlapping of HCS and LCS. This is significant because many studies do not present data for every cycle, potentially obscuring device failure moments.[69] By presenting every cycle, this study offers a realistic depiction of the long-term performance of the devices. Furthermore, the analysis focuses on device 1, which represents the worst-performing device among the five devices shown in the device-to-device analysis in Figure 3c. To further clarify the observed LCS variability, the cumulative distribution plot in Figure 3b provides a detailed analysis of the LCS in the first 10,000 cycles from the endurance plot. Given that the data do not follow a normal distribution, the 1, 2, and 3 sigmas from the median were analyzed. The results reveal that some of the data points skew toward the higher conductance states, while most remained at lower conductances. This indicates that it is not entirely accurate to base the small on−off ratio solely on the extreme values of the HCS and LCS. Moreover, by considering data within the 3σ (98.9 % of the data points), the variability in the LCS improves by 41.5%. Narrowing this to 2σ (95.6% of the data points) shows an even greater improvement of 55.2%. This skewness toward higher conductance suggests that a write-verify scheme could effectively filter out the less frequent, higher-conductance tail bits in practical applications, thereby mitigating the impact of off-state variability observed in the endurance plot. Next, in Figure 3c, five devices were utilized to showcase the device-to-device variation in their filamentary mode. Overall, the five devices displayed a median on−off window of ∼17× and an excellent median CV of 0.31 for LCS and 0.09 for HCS. Next, these five devices were also subjected to a separate scheme in the interfacial mode. As previously observed in DC measurements, the conductance change is gradual in the interfacial switching mode. Thus, by leveraging this characteristic, a sequence of pulses was applied to the device. This approach enables the memristor to function as an analogue memory, capable of storing a continuum of values. This is in contrast to the binary "0" or "1" states typically associated with the filamentary mode, as demonstrated in Figure 3b,c. An interfacial pulse operation scheme facilitating this capability is illustrated in Figure 3a, and the resulting data points are shown in Figure 3d,e. Starting from the lowest conductive state (approximately 0.2 μS), an incremental pulse width starting at 20 ns and increasing by 1 ns per pulse was applied at −2.5 V, culminating at 147 ns for the final pulse, thereby accessing 128 distinct states (indicated by the orange arrow). Subsequently, from the highest conductive state (approximately 6.0 μS), a similar incremental pulse width scheme was also deployed at 3.0 V to traverse the same 128 states (indicated by the blue arrow), resulting in the overall on/off ratio of the device in the interfacial mode to be approximately 21×. The V_G applied in this pulsed switching scheme was held at 1.2 V, corresponding to the lower operating current in the interfacial switching, as seen in Figure

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

1b. This gradual conductance change mirrors the variable strength of biological synapses, which is crucial for emulating synaptic plasticity, a fundamental mechanism of learning in the biological brain. In this context, increasing the strength of the synapse is known as potentiation, and the opposite is referred to as depression. Moreover, the results represent an enhancement over the Ta_2O_5-based memristor investigated in a previous work, which reached a saturation point after 49 potentiation pulses where the conductance level ceased to increase linearly with additional pulses. This saturation could affect the symmetry and linearity of the programing pulses, leading to increased accuracy degradation in neural network (NN) applications.[40] Specifically, in the framework of synaptic strength modulation, namely the potentiation and depression process, typically linear dynamics are desired.[70,71] Thus, a nonlinearity factor is introduced to describe this behavior quantitatively. This factor measures the deviation from a linear fitting when the change in synaptic strength (conductance in this case) is observed against the number of applied electrical stimuli. The factor is typically obtained from the following eqs 1−3[72]

$$G_{LTP} = G_{min} + B(1 - e^{-AP}) \quad (1)$$

$$G_{LTD} = G_{max} - B(1 - e^{A(P-1)}) \quad (2)$$

$$B = \frac{G_{max} - G_{min}}{1 - e^{-AP_{max}}} \quad (3)$$

where G_{max} and G_{min} are the maximum and minimum conductance levels and A represents the nonlinearity factor, which quantifies the linear deviation in the potentiation and depression process with respect to the pulse number, P. Lastly, B represents a fitting parameter. In an ideal memristor, where the conductance exhibits a perfectly linear behavior in response to changes in the number of pulses, the nonlinearity factor "A" would approach zero. This allows the change in the conductance between each state to be identical. Additionally, when the potentiation and depression curves become symmetrical, it implies a consistent traversal through identical states. This results in the uniform response of the memristor to varying stimuli and thereby ensures that the device conductance can be predictably modulated up or down without disproportionate deviation. Specifically, in this work, after fitting the first five potentiation and depression curves as shown in Figure 3d, an excellent median nonlinearity factor of 1.8 and 1.4 was obtained. Additionally, CV analysis was conducted for each of the 128 states to assess the deviations among them, and the results including the median and range of CVs are presented in Figure S14. As depicted in Figure S14a, the CV from all 128 potentiation and depression states of the device was determined to be in the range of 0.020−0.431 for potentiation and 0.026−0.576 for depression across the five different potentiation and depression (P/D) cycles. The low CV values demonstrate a uniform cycle-to-cycle response to the applied electrical stimuli. To delve deeper into the cycle-to-cycle variation of the memristor operating in the interfacial mode, the device was subjected to more than 10^6 pulses, encompassing a total of 3879 P/D cycles. Here, a P/D cycle is defined as a complete sequence comprising 128 potentiation pulses followed by 128 depression pulses. The analysis outcomes are presented in Figure S15, by which Figure S15a−c illustrates the first five P/D cycles for every order of pulses applied, extending up to the first million cycles. It was observed that most CV of the states remained <0.13, indicating minimal deviation throughout the extensive cycling. The nonlinearity factor saw a reduction to <1 across the span of the million pulses, showcasing an improvement in the linearity of the device response over prolonged cycling. Next, the five devices featured in Figure 3e were reused to assess the device-to-device variation when operating in the interfacial mode. By overlaying the first P/D cycle of each of the five devices, it was determined in Figure S14b that the CV from all 128 potentiation and depression states of the device was determined to be in the range of 0.016−0.396 for potentiation, and 0.033−0.507 for depression, indicating minimal variation across devices, underscoring their consistent performance and reliability in the interfacial mode. Moreover, across both cycle and device variations as shown in Figure S14, it is observed in the inset that most of the CVs across the states are <0.2, and the states in which CVs are >0.2 tend to be toward the lower conductance states. Figure 3f depicts the retention performance of the memristor in both the filamentary and interfacial modes. Two filamentary states and five interfacial states were tested for retention. They were read by utilizing a 0.1 V DC voltage in intervals of every ten s for 10^4 s. The sample was situated on a temperature stage set at 85 °C. Throughout this period, no significant fluctuations were observed, highlighting the nonvolatile nature of the memristor in both operational modes. This stability is noteworthy given the typically less stable retention characteristics of interfacial switching devices. It is noted that when comparing the CV of the conductance for the interfacial states with their filamentary counterparts, the former showed a median CV of 0.017, while the latter showed a relatively higher fluctuation of 0.033. While the interfacial states displayed minimal fluctuations, they still exhibited slightly higher variability compared to the filamentary states, which is still consistent with the general trend in the field where the retention of interfacial switching devices is often less stable than that of the filamentary switching devices. As previously mentioned, this improved retention performance can be attributed to the incorporation of nitrogen in the switching layer, which prevents diffusion of oxygen vacancies away from both the programed filamentary and interfacial states, thereby enhancing its retention. This nonvolatility is crucial in ensuring data integrity over time and energy efficiency by reducing the need for power to maintain data.[73] For neuromorphic computing applications, the stability of states enables the accurate emulation of biological synapses, essential for learning and memory tasks in artificial neural networks (ANNs). The energy consumption of the memristor was calculated by the time integral of instantaneous power, as depicted in Figure 3g,h for filamentary and interfacial switching pulses.[74] In this calculation, the entire waveform of the applied pulse was taken into account in the worst-case scenario, unlike other methods, which might take into account only the reading pulse or only the device switching time. For the filamentary switching, the energy per pulse for set and reset operations were 122 and 124 pJ, respectively. In the case of interfacial switching, the smallest (20 ns) and largest (147 ns) pulses for both potentiation and depression were analyzed, resulting in a range of energy consumption per pulse. For potentiation, the energy consumption ranged from 186 fJ for the smallest pulse to 87.1 pJ for the largest pulse, while for depression, the range was from 105 fJ to 6.6 pJ. All energy calculations included the 20 ns rise and fall time for all pulses. The switching energy of the memristive devices developed in this work is observed to

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

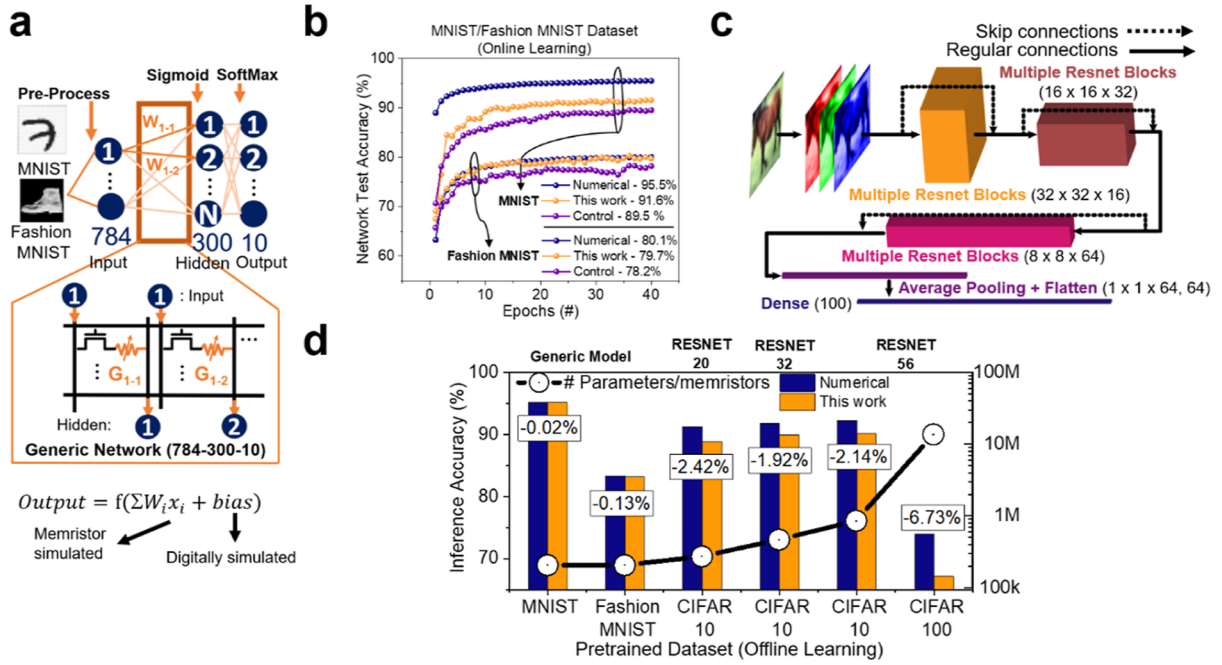


Figure 4. (a) Architecture of a fully connected MLP NN (784-300-10) trained on (b) MNIST and fashion MNIST data sets, showing test accuracy difference of 4% and 2.6% after 50 epochs in online learning. (c) Architecture of RESNET numerically trained on CIFAR-10 and CIFAR 100 data sets, with (d) offline learning conducted using test data sets. The study compares the impact of data set complexity and the number of simulated memristors on inference accuracy.

be relatively lower than reported devices in the literature, as shown in a benchmark plot in Figure 3i.[17,75−80] At lower conductance levels, these devices, when performing P/D, utilize less energy per pulse compared to certain three-terminal electrochemical devices, which are widely recognized for their low-energy operation.[79,80] However, those three-terminal devices often require longer switching times, whereas the devices in this study achieve comparable switching energies without sacrificing other critical switching parameters. Further details of the studies compared in Figure 3i are detailed in Table S1, where it is evident that the devices reported in this work offer superior scalability and endurance. This indicates that the combination of low-energy consumption, fast switching times, and high endurance makes these devices highly suitable for large-scale integration, surpassing the performance trade-offs commonly seen in other works. This establishes these memristive devices as strong candidates for next-generation energy-efficient memory and computing technologies.

Online and Offline Training with Progressively Complex Data Sets and ANNs. Beyond evaluating a memristor's intrinsic properties like variability, endurance, and retention is crucial, assessing its potential in real-world applications is equally important. Since the developed memristor in this work exhibits promising characteristics for mimicking synapses, the aim is to further analyze its performance through simulations of NNs. Generally, NN learning can be categorized into two types: online and offline learning. In online learning, the network continuously receives data and adapts to changing conditions.[81,82] Within the context of memristor networks, this means that the weights (or conductance levels) would continue to adjust as additional data is introduced. The significant challenge here is determining how device variations and the limited number of conductive states might hinder the accuracy of the training process. To determine this, a fully connected multilayer perceptron (MLP) network was established, as illustrated in Figure 4a.[83] The data sets utilized to train through this network were MNIST and fashion MNIST, each comprising grayscale handwritten and fashion product images of 28 × 28 pixels, respectively. These pixel values, ranging from 0 to 255, were normalized and fed through a test fully connected MLP network consisting of 784 input nodes, 300 nodes in the hidden layer, and 10 output nodes corresponding to the categories of the data sets. Each connection between the layers, known as synaptic weights, was modeled by using 1T−1R memristor devices. Furthermore, the activation functions, herein referred to as output neurons, were software implemented by using sigmoid and SoftMax functions, respectively. In summary, the *i*th row in the memristor array corresponds to the connection (synaptic weight) between the *i*th node from the preceding layer in the MLP simulation, and the *i*th column corresponds to the connection toward the *i*th node in the subsequent layer. The calculation for each layer can be summarized in eq 4 as follows

$$\text{output} = f(W_i x_i + \text{bias}) \tag{4}$$

where, each fully connected layer computes the summation of the *i*th weight, W_i, and the value from the previous node, x_i. A digital bias is then added to the result to offset the function. The total is then passed through a digital activation function *f*. This is then repeated for each layer until reaching the final layer, where the chosen output is determined by the highest percentage node. Figure 4b presents the test accuracy for classifying images from the MNIST and fashion MNIST data sets over 40 epochs using the network architecture outlined in Figure 4a. The cycle-to-cycle and device-to-device variations observed in the interfacial mode of the devices, as shown in Figure 3c,d, were incorporated as programing errors to simulate the device test conditions. Of the two data sets,

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

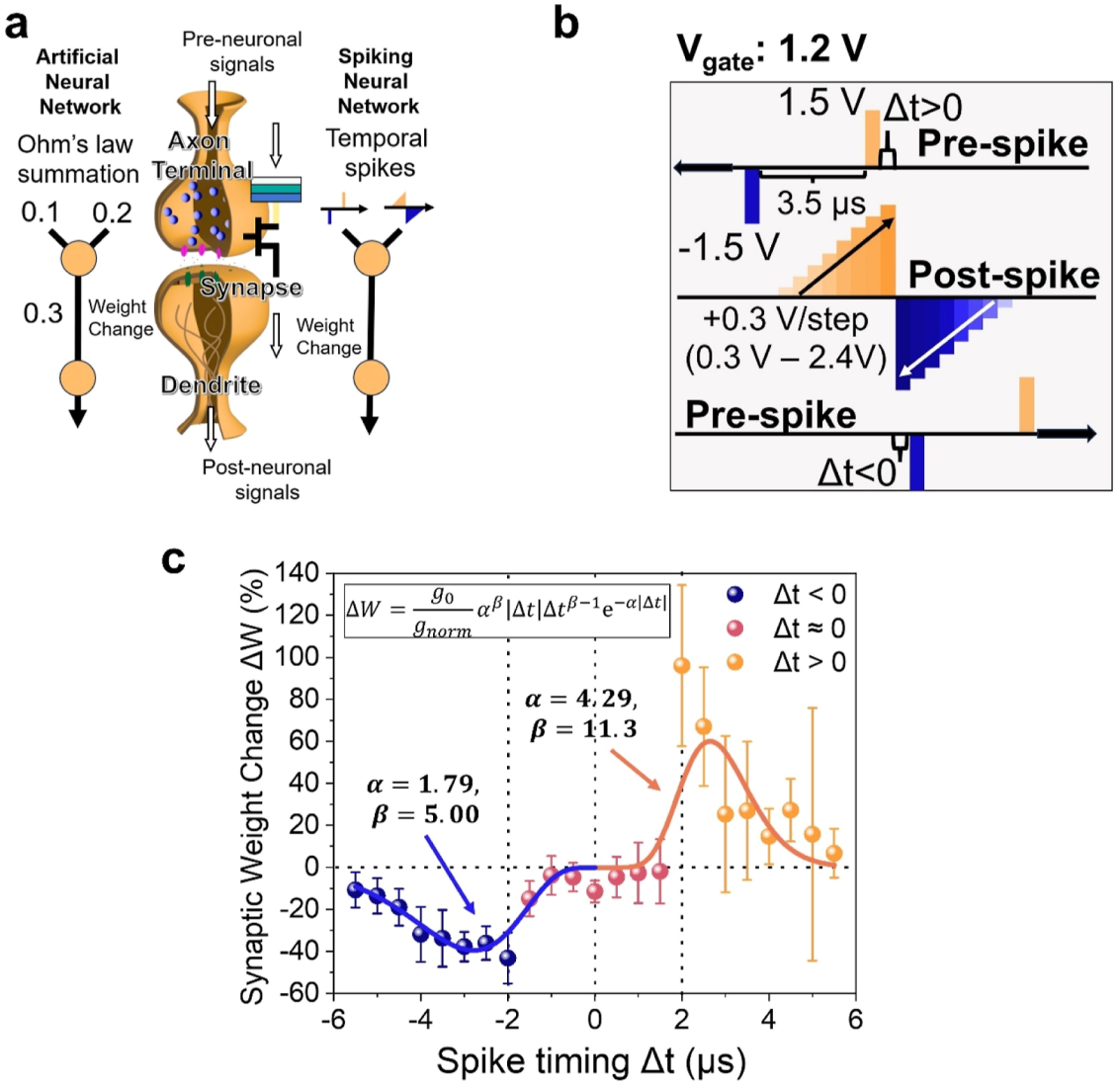


Figure 5. (a) Schematic of biological synapse element and how it inspires the ANN and SNN by leveraging on Kirchoff and Ohm's law and temporal spikes, respectively. (b) Pulse setup for the STDP demonstration and (c) experimental demonstration of STDP using pre- and postspiking pulses in the interfacial mode utilizing the 1T−1R memristor and synaptic weight change shown in response to the relative time delay between the two pulses.

fashion MNIST is recognized as the more challenging classification task compared to the standard MNIST data set.[84] Consequently, due to the relative ease of achieving high accuracy with the MNIST data set, many studies focus extensively on this data set.[29−31] These studies often leverage the high accuracy attained on the MNIST data set to validate the performance of their devices, overlooking that there are additional significant factors to be considered when evaluating device efficacy in neuromorphic applications. Hence, the primary interest here is not merely achieving high accuracy with the memristors developed in this work. Instead, the emphasis is on evaluating the discrepancy between the device test and the numerical test accuracies. Therefore, a generic network was employed to underscore this comparison. In Figure 4b, the differences between the numerical test and device memristor's (dubbed "This work") test accuracies after 40 epochs remain minor, with a 3.9% and 0.4% reduction for the MNIST and fashion MNIST data sets, respectively. This suggests that the differences in accuracies are not solely attributed to the inherent complexities of the data set. Instead, it is likely due to the saturation point in the training process; i.e., the network reaches a juncture in which no additional significant learning was achieved despite continued incoming data. It was also observed that in the numerical tests, the accuracy predominantly decreases as the number of epochs increases. This is opposed to the device tests, where there is a significantly higher frequency of accuracy decrease between epochs. This discrepancy can be attributed to the limited number of states available in the memristors, which hampers the backpropagation algorithm's ability to adjust the weights to their optimum values. Conversely, during numerical training, the number of states can be considered virtually infinite, allowing the backpropagation process to fine-tune the weights more precisely with each learning epoch, moving them closer to the optimum values. Subsequently, the Ta_2O_5-based memristor with its 49 potentiation and depression pulses was also simulated. This time, with fewer states than the device presented in this work, it lags by 2.1% and 1.5% in MNIST and fashion MNIST data set training accuracy, respectively. Figure 4c,d illustrates the architecture and results for offline learning.

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

In offline learning, the weights are pretrained and stored numerically.[3] This generally lessens the memristors' requirements since they only have to be programed once to the most optimum states as determined through numerical training. Taking advantage of this, a deep neural network (DNN) was simulated to investigate the effects of increasing network depth.[83] DNN is defined as any network architecture with more than one hidden layer.[85] As the number of NN layers increases, so does the requisite number of memristors needed to simulate each weight connection, and consequently, the likelihood of compounded errors also increases due to error aggregation through the network. However, merely adding more layers to a network often leads to diminished training accuracy due to the vanishing gradient problem. Hence, a well-established residual network architecture (RESNET) to enable layer incrementation with minimum impact on training accuracy was leveraged to determine the error aggregation in DNN based on the memristor developed in this work.[86] As depicted in Figure 4c, the RESNET architecture incorporates multiple skip connections to circumvent less significant layers. Figure 4d then showcases the inference accuracy for various NN architectures trained on different data sets. Two more data sets are introduced here, CIFAR-10 and CIFAR-100, which consist of colored images (32 × 32 pixels × 3 channels) and are categorized into 10 and 100 output classes, respectively.[87] Two aspects were analyzed to determine the impact on the disparity between the numerical and device inference accuracy: the complexity of the model, which relates to the number of memristors required to simulate the mode, and the complexity of the data set, which is ranked as follows: MNIST < fashion MNIST < CIFAR-10 < CIFAR-100. This order is based on the number of input pixels and output classes. Initially, the trained generic model, as described in Figure 4a, was utilized for inference on the MNIST and fashion MNIST data sets. Since the model used input pixels and number output classes were identical, the number of required memristors for the training of MNIST and fashion MNIST remained the same. It was noted that the difference between numerical and device inference accuracies for MNIST and fashion MNIST data sets was almost negligible. Subsequently, inference was done on the CIFAR-10 data set after training on different RESNET architecture networks, mainly the RESNET-20, 32, and 56, which corresponds to the total number of convolution layers within the DNN. The observed differences between the numerical and device inference accuracies were −2.42, −1.92, and −2.14%, corresponding to the RESNET-20, 32, and 56, respectively. Although these differences are still relatively insignificant, they reveal that there is no significant trend in the discrepancy between numerical and device inference accuracies, even as the number of simulated memristors increased from 273 to 857k. However, it is still crucial to note that both inference accuracies still increase with the complexity of the RESNET DNN layers, which align with expectations. Finally, the analysis was extended by introducing CIFAR-100, which consists of 100 output classes. This was trained again on the RESNET-56 network. Adjustments to the size of the convolution filters led to an increase in the number of memristors required for simulation, reaching ∼13.6 M. Here, the device accuracy deviates by almost −7%, marking a relatively more significant discrepancy than previous simulations. It should be noted that the numerical inference accuracies of the above simulations are consistent with the literature with a margin of error.[86] Therefore, it can be concluded that the complexity of data sets exerts a more substantial influence on the inference accuracy disparity than the cumulative errors resulting from the use of more memristors during offline learning. This is due to an absence of a need for weight tuning during inference, as opposed to online learning, where adjusting to the most optimal weight (conductance) values is crucial.

Adjustments of Weights Using Relative Timings of Action Potentials. In Figure 4, the focus was on utilizing memristors proposed in this work to simulate the construction of an ANN. The backpropagation algorithm, which is crucial to training the NN, relies on known label outputs to adjust the weights of the network. This is also known as supervised learning. Conversely, in a spiking neural network (SNN), the spike-timing-dependent plasticity (STDP) process facilitates unsupervised learning by strengthening synaptic connections based on neuronal spikes. The differences in data processing in ANN and SNN are depicted in Figure 5a. While ANN utilizes Kirchoff's and Ohm's law to perform weight adjustments, SNN stands out in its event-based processing, where computations are triggered by temporal events (spikes), making them highly efficient for real-time data processing.[88,89] Nevertheless, both ANN and SNN seek to emulate the biological processing by allowing the connection strength adjustment between pre- and postneurons, and this work aims to utilize the memristor proposed to emulate the synaptic element in both the ANN and SNN. Figure 5b depicts the pulse scheme for STDP demonstration, by which the delay of the pre- and postspikes (Δt) are separated into three regions: $\Delta t < 0$ (−5.50 to 2.02 μs), $\Delta t \approx 0$ (−1.50 to 1.50 μs), and $\Delta t < 0$ (2.02 to 5.50 μs). Figure 5c then illustrates the memristor demonstrating STDP behavior with these two distinct pulse spikes applied sequentially to the 1T−1R memristor in its interfacial mode. The synaptic weight (conductance) of the memristor, measured before and after the pulses, reveals changes corresponding to the spike timings. It was observed that a positive synaptic weight change results when the presynaptic spike precedes the postsynaptic spike, suggesting a causal relationship. Conversely, a negative change occurs when the order of spikes is reversed, indicating a lack of causality. This phenomenon, termed potentiation (orange data points) and depression (blue data points), respectively, also shows that the greater the time delay between spikes, the lesser the absolute change in synaptic weight. Moreover, in the pulse scheme depicted in Figure 5b, there is a specific time frame when the pre- and postsynaptic spikes are sent too closely together, which can result in no change in conductance (red data points). This phenomenon, which is consistent with other studies using nonidentical pulses, is also observed in biological systems under asymmetric Hebbian learning rule.[90] Each data point in Figure 5c represents the median weight change over 5 different paired pulses, and the error bars represent the range of the weight changes. This is represented by eq 5[91]

$$\Delta W = \frac{g_0}{g_{\text{norm}}} \alpha^{\beta} |\Delta t| \Delta t^{\beta-1} e^{-\alpha|\Delta t|} \tag{5}$$

where g_0 is the scaling factor, $g_0 = \beta e^{-\beta}$ is the normalizing constant, α and β are parameter values, and ΔW is the synaptic weight change as a result of Δt. The memristor developed in this study effectively functions within both ANNs and SNNs, addressing distinct computational needs. This versatility is advantageous, allowing the device to utilize the structured

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

pattern recognition capabilities of ANNs, along with the dynamic temporal processing of SNNs, thus providing comprehensive computational solutions.

CONCLUSION

This study developed a resistive switching memristor that exhibits both filamentary and interfacial switching capabilities aimed at advancing neuromorphic hardware elements. The initial demonstration of a 1T−1R bimodal device integration unraveled the operational current levels across different states during the switching process, providing essential insights into the functional characteristics of the device. XPS depth profiling was employed to elucidate the underlying switching mechanism, revealing the capacity of the memristor to support two distinct modes within a single device. It was also proposed that the improvements to the performance in both modes can be attributed to three distinct fabrication processes: N-doping, introduction to bilayer structure, and annealing. As variation has been a known problem plaguing resistive memristors, the temporal and spatial variations have also been investigated. Temporally, the memristor exhibited commendable endurance, surpassing 15 million cycles in the filamentary mode and a million in the interfacial mode. Spatial variability was also scrutinized across five devices, with the median CV recorded at 0.31 and 0.09 for the LCS and HCS in the filamentary mode, and most states remained <0.12 in the interfacial mode. Furthermore, retention tests at elevated temperatures of 85 °C for 10^4 s confirmed the stability of two filamentary and five interfacial states, underscoring the memristors' reliability and potential as nonvolatile memory elements suitable for CMOS integrations. Subsequently, the memristors were evaluated through computer simulations for their capabilities in both supervised and unsupervised learning. In supervised learning scenarios, particularly for online learning, the memristors achieved a high learning rate, with a deviation of less than 4% from numerical training in MNIST and fashion MNIST data sets on a fully connected NN architecture. Additionally, offline learning simulations with more complex data sets like CIFAR-10 and CIFAR-100 indicated a deviation of ∼2% and 6.73%, respectively, despite the significant increase in simulated memristors for the latter. This suggests that the error accumulation from utilizing more memristors had a minimal impact on the inference accuracy. Finally, the ability of the memristors to perform unsupervised learning was demonstrated through implementation of the biological plasticity mechanism. The devices in this work demonstrated a broad spectrum of applications, from conventional data processing to the analysis of real-time sensory information. This versatility not only enhances computational efficiency but also bridges artificial and biological intelligence by mimicking biological rules by using artificial memristive elements. This allows for the facilitation of in-memory computing, more effectively circumventing the von Neumann bottleneck and enabling parallel processing capabilities, representing a significant stride toward more efficient and biologically inspired computing architectures.

EXPERIMENTAL SECTION

Device Fabrication. The N−Ta_2O_5 and N−Al_2O_3 switching layers, and the Pt inert electrode layer were grown above the Ta_2N VIA (acting as the active electrode), which was connected to the drain of the NMOS transistor. Since the surface of the chip was encapsulated by Si_3N_4, electron beam lithography and reactive ion etching steps were utilized to gain access to the VIA. The chip was then subsequently transferred to a magnetron sputtering tool for deposition. An Ar reverse sputtering step was conducted on the exposed surface to preclean its surface from atmospheric contaminants with 50 W DC power at 5 mTorr chamber pressure for 2 min. A Ta_2O_5 ceramic target was utilized to deposit the 2 nm N−Ta_2O_5 film with a 50 W RF power, with the chamber pressure kept constant at 2 mTorr, while 19 sccm of Ar and 1 sccm of N_2 gases were released during the deposition process. A similar process was utilized to obtain a 1 nm N−Al_2O_3 film, albeit with an Al_2O_3 ceramic target at a 100 W RF power. Finally, a 20 nm Pt top electrode was grown by 50 W DC power deposited in 20 sccm of Ar at 2 mTorr.

Device and Material Characterization. For bulk characterizations, XPS was conducted by utilizing the Kratos Analytical Axis Supra Plus XPS system. The electrical *I*−*V* characterizations of the devices were performed using the Keithley 4200A-SCS semiconductor parameter analyzer.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.4c07454.

SEM, TEM, and EDX images of the cross section of the device; DC *I*−*V* characteristics of the memristor in both filamentary and interfacial switching depicting their multilevel switching; DC *I*−*V* characteristics of memristor in both filamentary and interfacial switching depicting the device-to-device and cycle-to-cycle variability for pre- and postannealed samples during the device optimization process; forming voltage device-to-device variation across samples during the device optimization process; comparison between DC *I*−*V* curve of the large XPS sample and device utilized in the study; XPS spectra analysis of O 1s and Ta 4f peaks; comparison of the DC *I*−*V* curve measured on the 1T, 1T−1R, and 1R devices in their filamentary and interfacial modes; cycle-to-cycle and device-to-device state variation during potentiation and depression; endurance of memristor during potentiation and depression across a million pulses; and summary of other developed memristors for similar applications (PDF)

AUTHOR INFORMATION

Corresponding Author

Wen Siang Lew − *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore;* orcid.org/0000-0002-5161-741X; Email: WenSiang@ntu.edu.sg

Authors

Eng Kang Koh − *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore; GLOBALFOUNDRIES Singapore Pte Ltd, Singapore 738406, Singapore;* orcid.org/0000-0003-0733-7688

Putu Andhita Dananjaya − *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore;* orcid.org/0000-0001-9416-6455

Lingli Liu − *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore;* orcid.org/0000-0002-4360-2699

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

Calvin Xiu Xian Lee − *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore; GLOBALFOUNDRIES Singapore Pte Ltd, Singapore 738406, Singapore;* orcid.org/0000-0002-5426-2482

Gerard Joseph Lim − *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore;* orcid.org/0000-0003-2411-5841

Young Seon You − *GLOBALFOUNDRIES Singapore Pte Ltd, Singapore 738406, Singapore*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsnano.4c07454

Author Contributions

E. K. Koh and P. A. Dananjaya proposed and designed this work. E. K. Koh conducted all of the experiments, including the fabrication of devices, measurements, collection of data, and analysis of results. P. A. Dananjaya, L. Liu, C. X. X. Lee, and G. J. Lim assisted in the analysis of results. Y. S. You and W. S. Lew supervised the research. E. K. Koh drafted the manuscript. All authors reviewed and revised the manuscript.

Funding

This work was supported by a RIE2020 ASTAR AME IAF-ICP grant (no. I1801E0030) and EDP-IPP: Economic Development Board—Industrial Postgraduate Programme (no. REQ0265124).

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

AI, artificial intelligence; LCS, low conductive state; HCS, high conductive state; TEM, transmission electron microscopy; SEM, scanning electron microscopy; FIB, focused ion beam; BEOL, back-end-of-line; VIA, vertical interconnect access; 1R, 1-resistor; 1T−1R, 1 transistor−1 resistor; MLC, multilevel cell; CV, coefficient of variation; XPS, X-ray photoelectron spectroscopy; P/D, potentiation and depression; ANN, artificial neural network; MLP, multilayer perceptron; DNN, deep neural network; RESNET, residual network architecture; SNN, spiking neural network; STDP, spike-timing-dependent plasticity

REFERENCES

(1) Sihn, S.; Chambers, W. L.; Abedin, M.; Beckmann, K.; Cady, N.; Ganguli, S.; Roy, A. K. Enhanced Computational Study with Experimental Correlation on I−V Characteristics of Tantalum Oxide (TaOx) Memristor Devices in a 1T1R Configuration. *Small* **2024**, 2310542.

(2) Lee, C. X. X.; Dananjaya, P. A.; Chee, M. Y.; Poh, H. Y.; Tan, F.; Thong, J. R.; Liu, L.; Lim, G. J.; Du, Y.; Tan, J. B.; Lew, W. S. Enhanced Resistive Switching Characteristics of Conductive Bridging Memory Device by a Co−Cu Alloy Electrode. *Appl. Phys. Lett.* **2023**, *123* (13), 133504.

(3) Yu, S.; Shim, W.; Peng, X.; Luo, Y. RRAM for Compute-in-Memory: From Inference to Training. *IEEE Transactions on Circuits and Systems I: Regular Papers* **2021**, *68* (7), 2753−2765.

(4) Yang, J.-Q.; Zhou, Y.; Han, S.-T. Functional Applications of Future Data Storage Devices. *Adv. Electron. Mater.* **2021**, *7* (5), 2001181.

(5) Liu, L.; Dananjaya, P. A.; Ang, C. C. I.; Koh, E. K.; Lim, G. J.; Poh, H. Y.; Chee, M. Y.; Lee, C. X. X.; Lew, W. S. A Bi-Functional Three-Terminal Memristor Applicable as an Artificial Synapse and Neuron. *Nanoscale* **2023**, *15* (42), 17076−17084.

(6) Dalgaty, T.; Moro, F.; Demirağ, Y.; De Pra, A.; Indiveri, G.; Vianello, E.; Payvand, M. Mosaic: In-Memory Computing and Routing for Small-World Spike-Based Neuromorphic Systems. *Nat. Commun.* **2024**, *15* (1), 142.

(7) Wan, W.; Kubendran, R.; Schaefer, C.; Eryilmaz, S. B.; Zhang, W.; Wu, D.; Deiss, S.; Raina, P.; Qian, H.; Gao, B.; Joshi, S.; Wu, H.; Wong, H.-S. P.; Cauwenberghs, G. A Compute-in-Memory Chip Based on Resistive Random-Access Memory. *Nature* **2022**, *608* (7923), 504−512.

(8) Nikam, R. D.; Lee, J.; Lee, K.; Hwang, H. Exploring the Cutting-Edge Frontiers of Electrochemical Random Access Memories (ECRAMs) for Neuromorphic Computing: Revolutionary Advances in Material-to-Device Engineering. *Small* **2023**, *19* (40), 2302593.

(9) Yin, S.; Sun, X.; Yu, S.; Seo, J.-S. High-Throughput In-Memory Computing for Binary Deep Neural Networks With Monolithically Integrated RRAM and 90-Nm CMOS. *IEEE Trans. Electron Devices* **2020**, *67* (10), 4185−4192.

(10) Ielmini, D.; Wong, H. S. P. In-Memory Computing with Resistive Switching Devices. *Nat. Electron.* **2018**, *1* (6), 333−343.

(11) Meng, J.; Yang, L.; Peng, X.; Yu, S.; Fan, D.; Seo, J.-S. Structured Pruning of RRAM Crossbars for Efficient In-Memory Computing Acceleration of Deep Neural Networks. *IEEE Trans. Circuits Syst.* **2021**, *68* (5), 1576−1580.

(12) Woo, J.; Yu, S. Resistive Memory-Based Analog Synapse: The Pursuit for Linear and Symmetric Weight Update. *IEEE Nanotechnol. Mag.* **2018**, *12* (3), 36−44.

(13) Sun, Y.; Xu, H.; Wang, C.; Song, B.; Liu, H.; Liu, Q.; Liu, S.; Li, Q. A Ti/AlOx/TaOx/Pt Analog Synapse for Memristive Neural Network. *IEEE Electron Device Lett.* **2018**, *39* (9), 1298−1301.

(14) Chen, W.-J.; Cheng, C.-H.; Lin, P.-E.; Tseng, Y.-T.; Chang, T.-C.; Chen, J.-S. Analog Resistive Switching and Synaptic Functions in WO_x/TaO_x Bilayer through Redox-Induced Trap-Controlled Conduction. *ACS Appl. Electron. Mater.* **2019**, *1* (11), 2422−2430.

(15) Chow, S. C. W.; Dananjaya, P. A.; Ang, J. M.; Loy, D. J. J.; Thong, J. R.; Hoo, S. W.; Toh, E. H.; Lew, W. S. Impact of Interfacial Engineering on MgO-Based Resistive Switching Devices for Low-Power Applications. *Appl. Surf. Sci.* **2023**, *608*, 155233.

(16) Wu, W.; Wu, H.; Gao, B.; Deng, N.; Qian, H. Suppress Variations of Analog Resistive Memory for Neuromorphic Computing by Localizing Vo Formation. *J. Appl. Phys.* **2018**, *124* (15), 152108.

(17) Kim, M.; Rehman, M. A.; Lee, D.; Wang, Y.; Lim, D.-H.; Khan, M. F.; Choi, H.; Shao, Q. Y.; Suh, J.; Lee, H.-S.; Park, H.-H. Filamentary and Interface-Type Memristors Based on Tantalum Oxide for Energy-Efficient Neuromorphic Hardware. *ACS Appl. Mater. Interfaces* **2022**, *14* (39), 44561−44571.

(18) Ryu, H.; Kim, S. Synaptic Characteristics from Homogeneous Resistive Switching in Pt/Al2O3/TiN Stack. *Nanomaterials* **2020**, *10* (10), 2055.

(19) Schönhals, A.; Rosário, C. M. M.; Hoffmann-Eifert, S.; Waser, R.; Menzel, S.; Wouters, D. J. Role of the Electrode Material on the RESET Limitation in Oxide ReRAM Devices. *Adv. Electron. Mater.* **2018**, *4* (2), 1700243.

(20) Rahmani, M. K.; Yang, B.-D.; Kim, H.; Kim, H.; Kang, M. H. Coexistence of Volatile and Non-Volatile Resistive Switching in Ni/SiO2/Pt Memristor Device Controlled from Different Current Compliances. *Semicond. Sci. Technol.* **2021**, *36* (9), 095031.

(21) Yun, S.; Park, J.; Kang, M.; Kim, S. Coexistence of Non-Volatile and Volatile Characteristics of the Pt/TaOx/TiN Device. *Results Phys.* **2022**, *34*, 105307.

(22) Covi, E.; Brivio, S.; Fanciulli, M.; Spiga, S. Synaptic Potentiation and Depression in Al:HfO2-Based Memristor. *Microelectron. Eng.* **2015**, *147*, 41−44.

(23) Ilyas, N.; Li, D.; Li, C.; Jiang, X.; Jiang, Y.; Li, W. Analog Switching and Artificial Synaptic Behavior of Ag/SiOx:Ag/TiOx/P++-Si Memristor Device. *Nanoscale Res. Lett.* **2020**, *15* (1), 30.

(24) Rahmani, M. K.; Ismail, M.; Mahata, C.; Kim, S. Effect of Interlayer on Resistive Switching Properties of SnO2-Based Memristor for Synaptic Application. *Results Phys.* **2020**, *18*, 103325.

(25) Bong, J. H.; Grebenchuk, S.; Nikolaev, K. G.; Chee, C. P. T.; Yang, K.; Chen, S.; Baranov, D.; Woods, C. R.; Andreeva, D. V.; Novoselov, K. S. Graphene Oxide−DNA/Graphene Oxide−PDDA Sandwiched Membranes with Neuromorphic Function. *Nanoscale Horiz.* **2024**, *9*, 863−872.
(26) Zhang, H.; Ju, X.; Yew, K. S.; Ang, D. S. Implementation of Simple but Powerful Trilayer Oxide-Based Artificial Synapses with a Tailored Bio-Synapse-Like Structure. *ACS Appl. Mater. Interfaces* **2020**, *12* (1), 1036−1045.
(27) Gonzalez-Rosillo, J. C.; Balaish, M.; Hood, Z. D.; Nadkarni, N.; Fraggedakis, D.; Kim, K. J.; Mullin, K. M.; Pfenninger, R.; Bazant, M. Z.; Rupp, J. L. M. Lithium-Battery Anode Gains Additional Functionality for Neuromorphic Computing through Metal−Insulator Phase Separation. *Adv. Mater.* **2020**, *32* (9), 1907465.
(28) Wang, Z.; Yin, M.; Zhang, T.; Cai, Y.; Wang, Y.; Yang, Y.; Huang, R. Engineering Incremental Resistive Switching in TaOx Based Memristors for Brain-Inspired Computing. *Nanoscale* **2016**, *8* (29), 14015−14022.
(29) Pérez, E.; Pérez-Ávila, A. J.; Romero-Zaliz, R.; Mahadevaiah, M. K.; Pérez-Bosch Quesada, E.; Roldán, J. B.; Jiménez-Molinos, F.; Wenger, C. Optimization of Multi-Level Operation in RRAM Arrays for In-Memory Computing. *Electronics* **2021**, *10* (9), 1084.
(30) Kwak, M.; Choi, W.; Heo, S.; Lee, C.; Nikam, R.; Kim, S.; Hwang, H. Excellent Pattern Recognition Accuracy of Neural Networks Using Hybrid Synapses and Complementary Training. *IEEE Electron Device Lett.* **2021**, *42* (4), 609−612.
(31) Pan, W.-Q.; Chen, J.; Kuang, R.; Li, Y.; He, Y.-H.; Feng, G.-R.; Duan, N.; Chang, T.-C.; Miao, X.-S. Strategies to Improve the Accuracy of Memristor-Based Convolutional Neural Networks. *IEEE Trans. Electron Devices* **2020**, *67* (3), 895−901.
(32) Jiang, Y.; Yang, Y.; Zhou, Z.; Xiang, Y.; Huang, P.; Kang, J. RRAM-Based Binary Neural Networks Using Back-Propagation Learning. In *2018 14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*; IEEE: Qingdao, 2018; pp 1−3.
(33) mnist|TensorFlow Datasets, TensorFlow. https://www.tensorflow.org/datasets/catalog/mnist (accessed April 10, 2024).
(34) Garcia, H.; Duenas, S.; Ossorio, O. G.; Castan, H. Current Pulses to Control the Conductance in RRAM Devices. *IEEE J. Electron Devices Soc.* **2020**, *8*, 291−296.
(35) Koveshnikov, S.; Matthews, K.; Min, K.; Gilmer, D. C.; Sung, M. G.; Deora, S.; Li, H. F.; Gausepohl, S.; Kirsch, P. D.; Jammy, R. Real-Time Study of Switching Kinetics in Integrated 1T/HfOx 1R RRAM: Intrinsic Tunability of Set/Reset Voltage and Trade-off with Switching Time. In *2012 International Electron Devices Meeting*; IEEE: San Francisco, CA, USA, 2012; pp 20.4.1−20.4.3.
(36) Dueñas, S.; Castán, H.; Ossorio, O. G.; García, H. Dynamics of Set and Reset Processes on Resistive Switching Memories. *Microelectron. Eng.* **2019**, *216*, 111032.
(37) Larentis, S.; Nardi, F.; Balatti, S.; Gilmer, D. C.; Ielmini, D. Resistive Switching by Voltage-Driven Ion Migration in Bipolar RRAM-Part II: Modeling. *IEEE Trans. Electron Devices* **2012**, *59* (9), 2468−2475.
(38) Liang, Z. A Comprehensive Understanding of Conductive Mechanism of RRAM: From Electron Conduction to Ionic Dynamics. In *2020 International Conference on Electrical Engineering and Control Technologies (CEECT)*; IEEE: Melbourne, VIC, Australia, 2020; pp 1−6.
(39) Ielmini, D. Modeling the Universal Set/Reset Characteristics of Bipolar RRAM by Field- and Temperature-Driven Filament Growth. *IEEE Trans. Electron Devices* **2011**, *58* (12), 4309−4317.
(40) Koh, E. K.; Dananjaya, P. A.; Poh, H. Y.; Liu, L.; Lee, C. X. X.; Thong, J. R.; You, Y. S.; Lew, W. S. Unraveling the Origins of the Coexisting Localized-Interfacial Mechanism in Oxide-Based Memristors in CMOS-Integrated Synaptic Device Implementations. *Nanoscale Horiz.* **2024**, *9*, 828−842.
(41) Sun, J.; Tan, J. B.; Chen, T. HfOx-Based RRAM Device With Sandwich-Like Electrode for Thermal Budget Requirement. *IEEE Trans. Electron Devices* **2020**, *67* (10), 4193−4200.
(42) Li, Y.; Suyolcu, Y. E.; Sanna, S.; Christensen, D. V.; Traulsen, M. L.; Stamate, E.; Pedersen, C. S.; van Aken, P. A.; García Lastra, J. M.; Esposito, V.; Pryds, N. Tuning the Resistive Switching in Tantalum Oxide-Based Memristors by Annealing. *AIP Adv.* **2020**, *10* (6), 065112.
(43) Cristea, D.; Cunha, L.; Gabor, C.; Ghiuta, I.; Croitoru, C.; Marin, A.; Velicu, L.; Besleaga, A.; Vasile, B. Tantalum Oxynitride Thin Films: Assessment of the Photocatalytic Efficiency and Antimicrobial Capacity. *Nanomaterials* **2019**, *9* (3), 476.
(44) Ravindra, P.; Liehr, M.; Mathkari, R.; Beckmann, K.; Tokranova, N.; Cady, N. Optimization of the Position of TaOx:N-Based Barrier Layer in TaOx RRAM Devices. *Front. Mater.* **2024**, *11*, 1343076.
(45) Misha, S. H.; Tamanna, N.; Woo, J.; Lee, S.; Song, J.; Park, J.; Lim, S.; Park, J.; Hwang, H. Effect of Nitrogen Doping on Variability of TaOx -RRAM for Low-Power 3-Bit MLC Applications. *ECS Solid State Lett.* **2015**, *4* (3), P25−P28.
(46) Sedghi, N.; Li, H.; Brunell, I. F.; Dawson, K.; Potter, R. J.; Guo, Y.; Gibbon, J. T.; Dhanak, V. R.; Zhang, W. D.; Zhang, J. F.; Robertson, J.; Hall, S.; Chalker, P. R. The Role of Nitrogen Doping in ALD Ta2O5 and Its Influence on Multilevel Cell Switching in RRAM. *Appl. Phys. Lett.* **2017**, *110* (10), 102902.
(47) Grossi, A.; Walczyk, D.; Zambelli, C.; Miranda, E.; Olivo, P.; Stikanov, V.; Feriani, A.; Sune, J.; Schoof, G.; Kraemer, R.; Tillack, B.; Fox, A.; Schroeder, T.; Wenger, C.; Walczyk, C. Impact of Intercell and Intracell Variability on Forming and Switching Parameters in RRAM Arrays. *IEEE Trans. Electron Devices* **2015**, *62* (8), 2502−2509.
(48) Yu, S.; Guan, X.; Wong, H.-S. P. On the Switching Parameter Variation of Metal Oxide RRAM-Part II: Model Corroboration and Device Design Strategy. *IEEE Trans. Electron Devices* **2012**, *59* (4), 1183−1188.
(49) Ismail, M.; Abbas, H.; Sokolov, A.; Mahata, C.; Choi, C.; Kim, S. Emulating Synaptic Plasticity and Resistive Switching Characteristics through Amorphous Ta2O5 Embedded Layer for Neuromorphic Computing. *Ceram. Int.* **2021**, *47* (21), 30764−30776.
(50) Shi, K. X.; Xu, H. Y.; Wang, Z. Q.; Zhao, X. N.; Liu, W. Z.; Ma, J. G.; Liu, Y. C. Improved Performance of Ta2O5−x Resistive Switching Memory by Gd-Doping: Ultralow Power Operation, Good Data Retention, and Multilevel Storage. *Appl. Phys. Lett.* **2017**, *111* (22), 223505.
(51) Simpson, R.; White, R. G.; Watts, J. F.; Baker, M. A. XPS Investigation of Monatomic and Cluster Argon Ion Sputtering of Tantalum Pentoxide. *Appl. Surf. Sci.* **2017**, *405*, 79−87.
(52) Khan, S.; Zapata, M. J. M.; Pereira, M. B.; Gonçalves, R. V.; Strizik, L.; Dupont, J.; Santos, M. J. L.; Teixeira, S. R. Structural, Optical and Photoelectrochemical Characterizations of Monoclinic Ta_3N_5 Thin Films. *Phys. Chem. Chem. Phys.* **2015**, *17* (37), 23952−23962.
(53) Cristea, D.; Crisan, A.; Cretu, N.; Borges, J.; Lopes, C.; Cunha, L.; Ion, V.; Dinescu, M.; Barradas, N. P.; Alves, E.; Apreutesei, M.; Munteanu, D. Structure Dependent Resistivity and Dielectric Characteristics of Tantalum Oxynitride Thin Films Produced by Magnetron Sputtering. *Appl. Surf. Sci.* **2015**, *354*, 298−305.
(54) Lanza, M.; Wong, H. S. P.; Pop, E.; Ielmini, D.; Strukov, D.; Regan, B. C.; Larcher, L.; Villena, M. A.; Yang, J. J.; Goux, L.; Belmonte, A.; Yang, Y.; Puglisi, F. M.; Kang, J.; Magyari-Köpe, B.; Yalon, E.; Kenyon, A.; Buckwell, M.; Mehonic, A.; Shluger, A.; Li, H.; Hou, T. H.; Hudec, B.; Akinwande, D.; Ge, R.; Ambrogio, S.; Roldan, J. B.; Miranda, E.; Suñe, J.; Pey, K. L.; Wu, X.; Raghavan, N.; Wu, E.; Lu, W. D.; Navarro, G.; Zhang, W.; Wu, H.; Li, R.; Holleitner, A.; Wurstbauer, U.; Lemme, M. C.; Liu, M.; Long, S.; Liu, Q.; Lv, H.; Padovani, A.; Pavan, P.; Valov, I.; Jing, X.; Han, T.; Zhu, K.; Chen, S.; Hui, F.; Shi, Y. Recommended Methods to Study Resistive Switching Devices. *Adv. Electron. Mater.* **2019**, *5* (1), 1800143.
(55) Prabhu, N.; Loy Jia Jun, D.; Dananjaya, P. A.; Lew, W. S.; Toh, E. H.; Raghavan, N. Exploring the Impact of Variability in Resistance Distributions of RRAM on the Prediction Accuracy of Deep Learning Neural Networks. *Electronics* **2020**, *9* (3), 414.

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617

(56) Chen, A.; Lin, M.-R. Variability of Resistive Switching Memories and Its Impact on Crossbar Array Performance. In *2011 International Reliability Physics Symposium*, 2011; pp MY.7.1−MY.7.4.
(57) Hwan Kim, G.; Ho Lee, J.; Yeong Seok, J.; Ji Song, S.; Ho Yoon, J.; Jean Yoon, K.; Hwan Lee, M.; Min Kim, K.; Dong Lee, H.; Wook Ryu, S.; Joo Park, T.; Seong Hwang, C. Improved Endurance of Resistive Switching TiO2 Thin Film by Hourglass Shaped Magnéli Filaments. *Appl. Phys. Lett.* **2011**, *98* (26), 262901.
(58) Balatti, S.; Ambrogio, S.; Wang, Z.; Sills, S.; Calderoni, A.; Ramaswamy, N.; Ielmini, D. Voltage-Controlled Cycling Endurance of HfO_x-Based Resistive-Switching Memory. *IEEE Trans. Electron Devices* **2015**, *62* (10), 3365−3372.
(59) Kim, K. M.; Yang, J. J.; Strachan, J. P.; Grafals, E. M.; Ge, N.; Melendez, N. D.; Li, Z.; Williams, R. S. Voltage Divider Effect for the Improvement of Variability and Endurance of TaOx Memristor. *Sci. Rep.* **2016**, *6* (1), 20085.
(60) Liu, S.; Lu, N.; Zhao, X.; Xu, H.; Banerjee, W.; Lv, H.; Long, S.; Li, Q.; Liu, Q.; Liu, M. Eliminating Negative-SET Behavior by Suppressing Nanofilament Overgrowth in Cation-Based Memory. *Adv. Mater.* **2016**, *28* (48), 10623−10629.
(61) Aziza, H.; Coulié, K.; Rahajandraibe, W. Design Considerations Towards Zero-Variability Resistive RAMs in HRS State. In *2021 IEEE 22nd Latin American Test Symposium (LATS)*, 2021; pp 1−5.
(62) Hajri, B.; Aziza, H.; Mansour, M. M.; Chehab, A. RRAM Device Models: A Comparative Analysis With Experimental Validation. *IEEE Access* **2019**, *7*, 168963−168980.
(63) Lee, J.-K.; Kim, S. Comparative Analysis of Low-Frequency Noise Based Resistive Switching Phenomenon for Filamentary and Interfacial RRAM Devices. *Chaos, Solit. Fractals* **2023**, *173*, 113633.
(64) Veksler, D.; Bersuker, G.; Vandelli, L.; Padovani, A.; Larcher, L.; Muraviev, A.; Chakrabarti, B.; Vogel, E.; Gilmer, D. C.; Kirsch, P. D. Random Telegraph Noise (RTN) in Scaled RRAM Devices. In *2013 IEEE International Reliability Physics Symposium (IRPS)*, 2013; pp MY.10.1−MY.10.4.
(65) Zhao, X.; Ma, J.; Xiao, X.; Liu, Q.; Shao, L.; Chen, D.; Liu, S.; Niu, J.; Zhang, X.; Wang, Y.; Cao, R.; Wang, W.; Di, Z.; Lv, H.; Long, S.; Liu, M. Breaking the Current-Retention Dilemma in Cation-Based Resistive Switching Devices Utilizing Graphene with Controlled Defects. *Adv. Mater.* **2018**, *30* (14), 1705193.
(66) Nikam, R. D.; Rajput, K. G.; Hwang, H. Single-Atom Quantum-Point Contact Switch Using Atomically Thin Hexagonal Boron Nitride. *Small* **2021**, *17* (7), 2006760.
(67) Lee, M.-J.; Lee, D.; Cho, S.-H.; Hur, J.-H.; Lee, S.-M.; Seo, D. H.; Kim, D.-S.; Yang, M.-S.; Lee, S.; Hwang, E.; Uddin, M. R.; Kim, H.; Chung, U.-I.; Park, Y.; Yoo, I.-K. A Plasma-Treated Chalcogenide Switch Device for Stackable Scalable 3D Nanoscale Memory. *Nat. Commun.* **2013**, *4* (1), 2629.
(68) Lee, M.-J.; Lee, C. B.; Lee, D.; Lee, S. R.; Chang, M.; Hur, J. H.; Kim, Y.-B.; Kim, C.-J.; Seo, D. H.; Seo, S.; Chung, U.-I.; Yoo, I.-K.; Kim, K. A Fast, High-Endurance and Scalable Non-Volatile Memory Device Made from Asymmetric Ta2O5−x/TaO2−x Bilayer Structures. *Nat. Mater.* **2011**, *10* (8), 625−630.
(69) Lanza, M.; Waser, R.; Ielmini, D.; Yang, J. J.; Goux, L.; Suñe, J.; Kenyon, A. J.; Mehonic, A.; Spiga, S.; Rana, V.; Wiefels, S.; Menzel, S.; Valov, I.; Villena, M. A.; Miranda, E.; Jing, X.; Campabadal, F.; Gonzalez, M. B.; Aguirre, F.; Palumbo, F.; Zhu, K.; Roldan, J. B.; Puglisi, F. M.; Larcher, L.; Hou, T.-H.; Prodromakis, T.; Yang, Y.; Huang, P.; Wan, T.; Chai, Y.; Pey, K. L.; Raghavan, N.; Dueñas, S.; Wang, T.; Xia, Q.; Pazos, S. Standards for the Characterization of Endurance in Resistive Switching Devices. *ACS Nano* **2021**, *15* (11), 17214−17231.
(70) Zhang, W.; Gao, B.; Tang, J.; Li, X.; Wu, W.; Qian, H.; Wu, H. Analog-Type Resistive Switching Devices for Neuromorphic Computing. *Phys. Status Solidi RRL* **2019**, *13* (10), 1−13.
(71) Zhao, M.; Gao, B.; Tang, J.; Qian, H.; Wu, H. Reliability of Analog Resistive Switching Memory for Neuromorphic Computing. *Applied Physics Reviews* **2020**, *7* (1), 011301.
(72) Ielmini, D. Brain-Inspired Computing with Resistive Switching Memory (RRAM): Devices, Synapses and Neural Networks. *Microelectron. Eng.* **2018**, *190*, 44−53.
(73) Zahoor, F.; Azni Zulkifli, T. Z.; Khanday, F. A. Resistive Random Access Memory (RRAM): An Overview of Materials, Switching Mechanism, Performance, Multilevel Cell (Mlc) Storage, Modeling, and Applications. *Nanoscale Res. Lett.* **2020**, *15* (1), 90.
(74) Yu, S.; Wu, Y.; Wong, H.-S. P. Investigating the Switching Dynamics and Multilevel Capability of Bipolar Metal Oxide Resistive Switching Memory. *Appl. Phys. Lett.* **2011**, *98* (10), 103514.
(75) Li, Q.; Yu, S.; Yang, P.; Wang, Q.; Liu, S. A Large Memory Window and Low Power Consumption Self-Rectifying Memristor for Electronic Synapse. *Electron. Lett.* **2023**, *59* (2), No. e12717.
(76) Ilyas, N.; Li, C.; Wang, J.; Jiang, X.; Fu, H.; Liu, F.; Gu, D.; Jiang, Y.; Li, W. A Modified SiO2-Based Memristor with Reliable Switching and Multifunctional Synaptic Behaviors. *J. Phys. Chem. Lett.* **2022**, *13* (3), 884−893.
(77) Pérez, E.; González Ossorio, Ó.; Dueñas, S.; Castán, H.; García, H.; Wenger, C. Programming Pulse Width Assessment for Reliable and Low-Energy Endurance Performance in Al:HfO2-Based RRAM Arrays. *Electronics* **2020**, *9* (5), 864.
(78) Shih, C.-C.; Chen, W.-J.; Chang, K.-C.; Chang, T.-C.; Tsai, T.-M.; Chu, T.-J.; Tseng, Y.-T.; Wu, C.-H.; Su, W.-C.; Chen, M.-C.; Huang, H.-C.; Wang, M.-H.; Chen, J.-H.; Zheng, J.-C.; Sze, S. M. Ultra-Low Switching Voltage Induced by Inserting SiO2 Layer in Indium−Tin−Oxide-Based Resistance Random Access Memory. *IEEE Electron Device Lett.* **2016**, *37* (10), 1276−1279.
(79) Sharbati, M. T.; Du, Y.; Torres, J.; Ardolino, N. D.; Yun, M.; Xiong, F. Low-Power, Electrochemically Tunable Graphene Synapses for Neuromorphic Computing. *Adv. Mater.* **2018**, *30* (36), 1802353.
(80) Li, X.; Yu, B.; Wang, B.; Bao, L.; Zhang, B.; Li, H.; Yu, Z.; Zhang, T.; Yang, Y.; Huang, R.; Wu, Y.; Li, M. Multi-Terminal Ionic-Gated Low-Power Silicon Nanowire Synaptic Transistors with Dendritic Functions for Neuromorphic Systems. *Nanoscale* **2020**, *12* (30), 16348−16358.
(81) Kim, M.-H.; Lee, S.-H.; Kim, S.; Park, B.-G. A Fast Weight Transfer Method for Real-Time Online Learning in RRAM-Based Neuromorphic System. *IEEE Access* **2022**, *10*, 37030−37038.
(82) Dong, Z.; Zhou, Z.; Li, Z.; Liu, C.; Huang, P.; Liu, L.; Liu, X.; Kang, J. Convolutional Neural Networks Based on RRAM Devices for Image Recognition and Online Learning Tasks. *IEEE Trans. Electron Devices* **2019**, *66* (1), 793−801.
(83) Xiao, T. P.; Bennett, C. H.; Feinberg, B.; Marinella, M. J.; Agarwal, S. CrossSim: Accuracy Simulation of Analog in-Memory Computing. https://github.com/sandialabs/cross-sim (accessed April 2, 2024).
(84) Xiao, H.; Rasul, K.; Vollgraf, R. Fashion-MNIST: A Novel Image Dataset for Benchmarking Machine Learning Algorithms. **2017**, arXiv:1708.07747
(85) Samek, W.; Montavon, G.; Lapuschkin, S.; Anders, C. J.; Muller, K.-R. Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications. *Proc. IEEE* **2021**, *109* (3), 247−278.
(86) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016; pp 770−778.
(87) CIFAR-10 and CIFAR-100 datasets. https://www.cs.toronto.edu/~kriz/cifar.html (accessed April 11, 2024).
(88) Tavanaei, A.; Ghodrati, M.; Kheradpisheh, S. R.; Masquelier, T.; Maida, A. Deep Learning in Spiking Neural Networks. *Neural Network.* **2019**, *111*, 47−63.
(89) Kasabov, N. K. *Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence*; Springer Nature, 2019.
(90) Talathi, S. S.; Hwang, D.-U.; Ditto, W. L. Spike Timing Dependent Plasticity Promotes Synchrony of Inhibitory Networks in the Presence of Heterogeneity. *J. Comput. Neurosci.* **2008**, *25* (2), 262−281.
(91) Yang, R.; Huang, H.-M.; Hong, Q.-H.; Yin, X.-B.; Tan, Z.-H.; Shi, T.; Zhou, Y.-X.; Miao, X.-S.; Wang, X.-P.; Mi, S.-B.; Jia, C.-L.;

Guo, X. Synaptic Suppression Triplet-STDP Learning Rule Realized in Second-Order Memristors. *Adv. Funct. Mater.* **2018**, *28* (5), 1704455.

https://doi.org/10.1021/acsnano.4c07454
ACS Nano 2024, 18, 29602−29617