

Cite this: *Mater. Horiz.*, 2025,  
12, 7509Received 22nd February 2025,  
Accepted 3rd June 2025

DOI: 10.1039/d5mh00324e

rsc.li/materials-horizons

## Highly-efficient and scalable TrioN (3N0C) synaptic cell for analog process-in-memory†

Junyoung Choi,<sup>id</sup> Byoungwoo Lee,<sup>id</sup> Jinho Byun,<sup>id</sup> Hyejin Kim,<sup>id</sup>  
Seungkun Kim,<sup>id</sup> Junyong Lee,<sup>id</sup> Hyunjeong Kwak,<sup>id</sup> Jeonghoon Son<sup>id</sup> and  
Seyoung Kim<sup>id</sup>\*

The development of non-volatile memory (NVM)-based cross-point arrays has paved the way for the realization of neuromorphic architectures toward analog process-in-memory (aPIM). However, the inherent non-ideal characteristics of NVM devices such as asymmetry necessitate capacitor-based synaptic cells. We introduce a capacitorless 3-NMOS 0-capacitor (TrioN, 3N0C) cross-point device, a novel synaptic cell leveraging amorphous indium gallium zinc oxide (a-IGZO), and a promising oxide semiconductor with low off-current. Using only NMOS transistors, TrioN enables high-density arrays with a simplified fabrication process without the need for PMOS or external capacitors. It exhibits excellent switching characteristics, including perfect symmetry, a high on/off ratio, and ultra-fast 10 ns switching. Fabricated hardware demonstrates precise selective updates, supporting 2-cycle updates for improved speed and energy efficiency over conventional 4-cycle methods. Neural network simulations using stochastic gradient descent (SGD) and Tiki-Taka algorithm version 1 (TTv1) on a multi-layer perceptron (MLP) for the MNIST dataset achieve high accuracy of 96.89% and 97.19%, respectively. These results highlight TrioN's potential as a compact, energy-efficient, and scalable solution for neuromorphic computing.

### New concepts

Our work introduces TrioN (3N0C), the first capacitorless cross-point synaptic cell utilizing IGZO-based NMOS transistors. Unlike conventional capacitor-based synaptic cells such as 3T1C (S. Kim *et al.*, 2017 MWSCAS<sup>16</sup>), which require large capacitors and PMOS transistors, our TrioN device without external capacitors and with only NMOS transistors leverages the intrinsic overlap capacitance of oxide semiconductor transistors, significantly reducing device footprint and fabrication complexity. This breakthrough enables high-density integration and enhanced energy efficiency in analog process-in-memory (aPIM) architectures. TrioN differentiates itself through its perfect symmetry, ultra-fast 10 ns switching, large on/off ratio and a novel 2-cycle update mechanism, which improves computational speed and energy efficiency compared to traditional 4-cycle updates. Its selective update mechanism also enhances large-scale array performance, offering a compact and scalable alternative to existing architectures. From a materials science perspective, this research highlights the potential of IGZO thin-film transistors in high-performance neuromorphic applications. The low leakage current of IGZO ensures excellent retention, while its high stability enables rapid and precise synaptic weight updates. By demonstrating a capacitorless architecture with a high integration potential, our work provides new insights into oxide semiconductor applications, paving the way for next generation ultra-efficient, scalable neuromorphic computing systems.

## 1. Introduction

The rapid advancement of artificial intelligence (AI) has been driving innovation across various fields, from healthcare to autonomous systems.<sup>1</sup> Central to this progress is the development of computing architectures that can meet ever-increasing demands for computational power and energy efficiency.<sup>2</sup> However, the traditional von Neumann architecture has approached its inherent limitations, particularly in terms of memory bottlenecks and energy inefficiency.<sup>3</sup> This has led to the exploration of alternative architectures, including near-memory computing

and process-in-memory (PIM).<sup>4</sup> Among these architectures, PIM has gained prominence for its ability to address these limitations effectively.

PIM systems aim to integrate data storage and computation within the same physical location, thereby reducing data transfer overhead and improving energy efficiency.<sup>5</sup> Furthermore, for accelerating neural network training by processing the storage and computation of large-scale data in parallel, cross-point array-based neuromorphic architectures have been proposed.<sup>6,7</sup>

Analog process-in-memory (aPIM) leverages the unique properties of analog memory elements, such as resistive random access memory, phase-change memory, ferroelectric memory, magnetoresistive memory, and electrochemical memory, to enable high-density and energy-efficient computing.<sup>8–11</sup> In previous research addressing the device specifications of resistive

Department of Materials Science and Engineering, Postech, Pohang, 37673, Republic of Korea. E-mail: kimseyoung@postech.ac.kr; Tel: +010 9797 0779

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d5mh00324e>



processing units for aPIM, the non-ideal characteristics of these non-volatile memories (NVMs), such as non-linearity, asymmetry, and switching speed, degrade neural network training performance.<sup>12–15</sup> Meanwhile, capacitor-based cells have recently emerged as a promising approach.<sup>16–19</sup> Capacitor-based cells operate by programming data through the charging and discharging of a capacitor, with the stored charge being read *via* devices like transistors. Since its initial proposal in 2017,<sup>16</sup> capacitor-based cells have undergone significant advancements, including design strategies for optimization, successful fabrication, and demonstration of array-level operations.<sup>17,20,21</sup> Despite these achievements, the technology faces notable challenges. For example, the reliance on conventional silicon transistors leads to high leakage currents, necessitating the use of large capacitors to store charge.<sup>22</sup> This requirement increases the overall device size and limits scalability. Additionally, the fabrication of PMOS transistors is complex, compounded by the challenges in asymmetry due to the difference in carrier mobility between the utilized PMOS and NMOS transistors.

To address these limitations, extensive research has been conducted on oxide semiconductor-based devices, such as indium–gallium–zinc oxide (IGZO), for memory and AI applications, including capacitorless 2-Transistor 0-Capacitor (2T0C),<sup>23–29</sup> 2T1C,<sup>30</sup> and 6T1C<sup>31</sup> configurations. Oxide transistors are known for their exceptionally low off-state leakage current, often as low as  $10^{-22}$  A  $\mu\text{m}^{-1}$ ,<sup>32</sup> which makes them highly suitable for charge storage applications.<sup>33</sup> This research has focused on device architectures composed solely of NMOS transistors with oxide channels instead of Si, and alternative approaches utilizing the gate capacitance of the read transistor without external capacitors have been actively explored. In this work, we propose a capacitorless-3N0C synaptic cell with only 3 NMOSs leveraging oxide transistors, specifically indium-gallium-zinc oxide (IGZO) transistors,<sup>34–38</sup> to develop a cross-point array architecture for an analog AI accelerator. A key strength of our proposed TrioN cell is its capacitorless structure, which enables fully parallel updates in a cross-point array. This feature is critical for analog process-in-memory applications, allowing individual devices to be updated in parallel without interference. While the conventional 6T1C structure supports parallel updates, its reliance on multiple transistors and a capacitor results in a large device area and significant scaling challenges. On the other hand, the capacitorless 2T0C structure is more favorable for scaling and high-density integration, but its single write transistor prevents fully-parallel updates, as an entire row must be activated for writing operations. To overcome these limitations, our TrioN cell incorporates two write NMOSs connected in series, each operated with independent gate control. This configuration not only prevents interference between neighboring devices but also enables fully parallel updates across the array. Furthermore, our design introduces a highly efficient 2-cycle update method, a significant improvement over the conventional 4-cycle update. This advantage has been validated through successful fully parallel cross-point array operations, demonstrating its potential for energy-efficient and scalable aPIM systems.

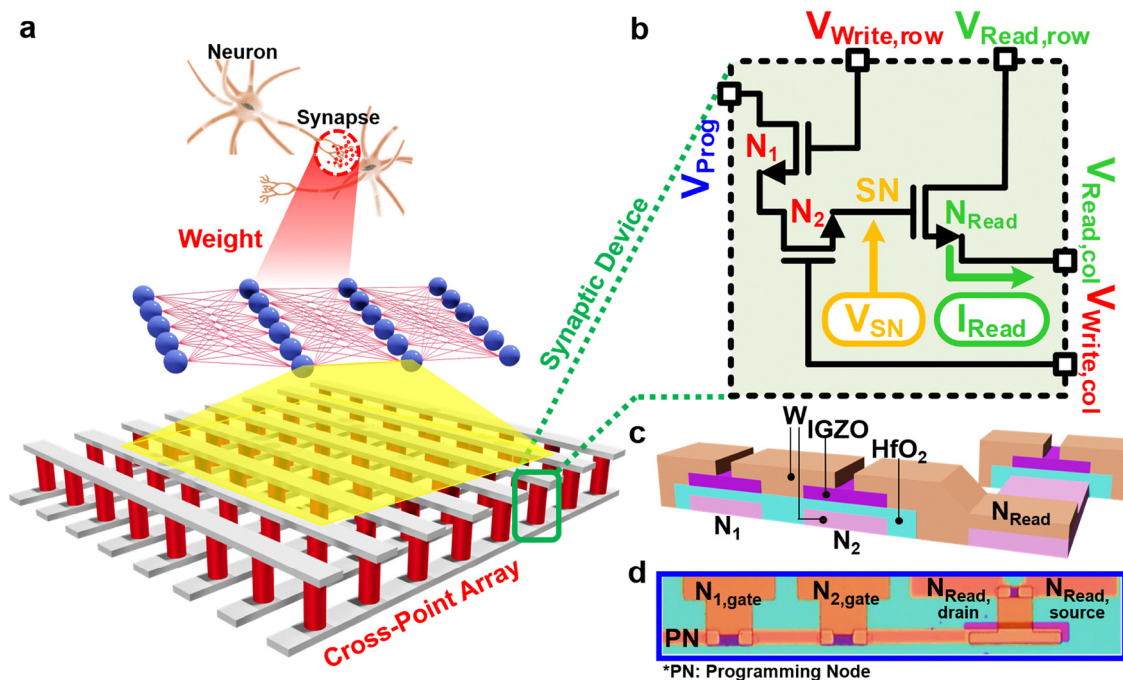
## 2. Results and discussion

In this study, we investigated the development and performance of a novel synaptic device: a TrioN cell, specifically designed for application in cross-point arrays for neural network acceleration. Fig. 1a illustrates the architecture of the cross-point array, which consists of neurons, synapses, and synaptic devices. The conductance of each cross-point element corresponds to the weight of a neural network. The synaptic device used as the cross-point element is a crucial component enabling large-scale parallel processing, a foundational requirement for analog AI computations. The TrioN cell, highlighted in green in Fig. 1b, serves as the focus of this study.

### 2.1. Concept of the TrioN cell

The TrioN device is a synaptic cell composed of only three NMOS transistors ( $N_1$ ,  $N_2$ , and  $N_{\text{Read}}$ ) based on IGZO, utilizing the overlap capacitance between transistors as a storage node instead of a dedicated capacitor, thereby achieving high integration and efficiency with a minimal transistor count. Two NMOSs,  $N_1$  and  $N_2$ , are connected in series to regulate the charging/discharging currents and serve as an AND gate for an array-level update, while the gate of  $N_{\text{Read}}$ , which serves as the storage node, is connected to the drain of  $N_2$ . The synaptic weight is encoded by the changes in the current flowing through  $N_{\text{Read}}$ , which is influenced by the amount of charge on the storage node determining the storage node voltage ( $V_{\text{SN}}$ ). The TrioN consists of 5 terminals: the programming node (PN), the gate of  $N_1$  ( $N_{1,\text{gate}}$ ), the gate of  $N_2$  ( $N_{2,\text{gate}}$ ), the drain of  $N_{\text{Read}}$  ( $N_{\text{Read,drain}}$ ), and the source of  $N_{\text{Read}}$  ( $N_{\text{Read,source}}$ ) where PN serves as a global terminal shared across multiple cells, while the remaining terminals are individually assigned to each cell within the cross-point array. These terminals are separated across the different row and column lines in the cross-point array and are shared with neighboring devices. The PN is a terminal that generates current flow when charging and discharging storage nodes, and  $N_{\text{Read,drain}}$  and  $N_{\text{Read,source}}$  are needed to read the conductance of  $N_{\text{Read}}$ , which can be represented by  $I = G \cdot V$ . Fig. 1c shows the schematic for the TrioN, and fabrication details are provided in Fig. S1a (ESI<sup>†</sup>). All three transistors consisting of the TrioN device are bottom gate transistors, and the bottom gate is formed by depositing W metal, followed by the deposition of the  $\text{HfO}_2$  gate dielectric using ALD. To form the storage node by connecting the drain of the second transistor  $N_2$  and the gate of the read transistor, hole etching is carried out. After that, the TrioN device is fabricated by depositing and patterning the IGZO channel layer, followed by the deposition of the source and drain metals. Specifically, as shown in Fig. S1b and c (ESI<sup>†</sup>), the storage node capacitance is based on the overlap capacitance composed of W,  $\text{HfO}_2$ , and IGZO, generated during the fabrication of  $N_{\text{Read}}$ . The calculated capacitance value is approximately 1.77 pF, as detailed in Fig. S1c (ESI<sup>†</sup>). Fig. 1d presents an optical microscopy (OM) image of the fabricated TrioN. Further material characterization of the  $N_{\text{Read}}$  active area was conducted using scanning electron microscopy (SEM) and energy dispersive spectroscopy (EDS), as





**Fig. 1** Architecture of the TrioN cell and its integration within a cross-point array. (a) Neural network weights represented as both synapses in a biological system and the conductance values of synaptic devices within a cross-point array. (b) Single TrioN cell schematics. A single cell as the cross-point element has 5 terminals:  $V_{\text{Prog}}$ ,  $V_{\text{Write,row}}$ ,  $V_{\text{Write,col}}$ ,  $V_{\text{Read,row}}$ , and  $V_{\text{Read,col}}$ . (c) Device structure of the fabricated TrioN cell. (d) Optical microscopy (OM) image of the fabricated TrioN cell. The five terminals of a single TrioN cell shown in Fig. 1b are connected to PN (programming node),  $N_{1,\text{gate}}$ ,  $N_{2,\text{gate}}$ ,  $N_{\text{Read,drain}}$ , and  $N_{\text{Read,source}}$  on each transistor.

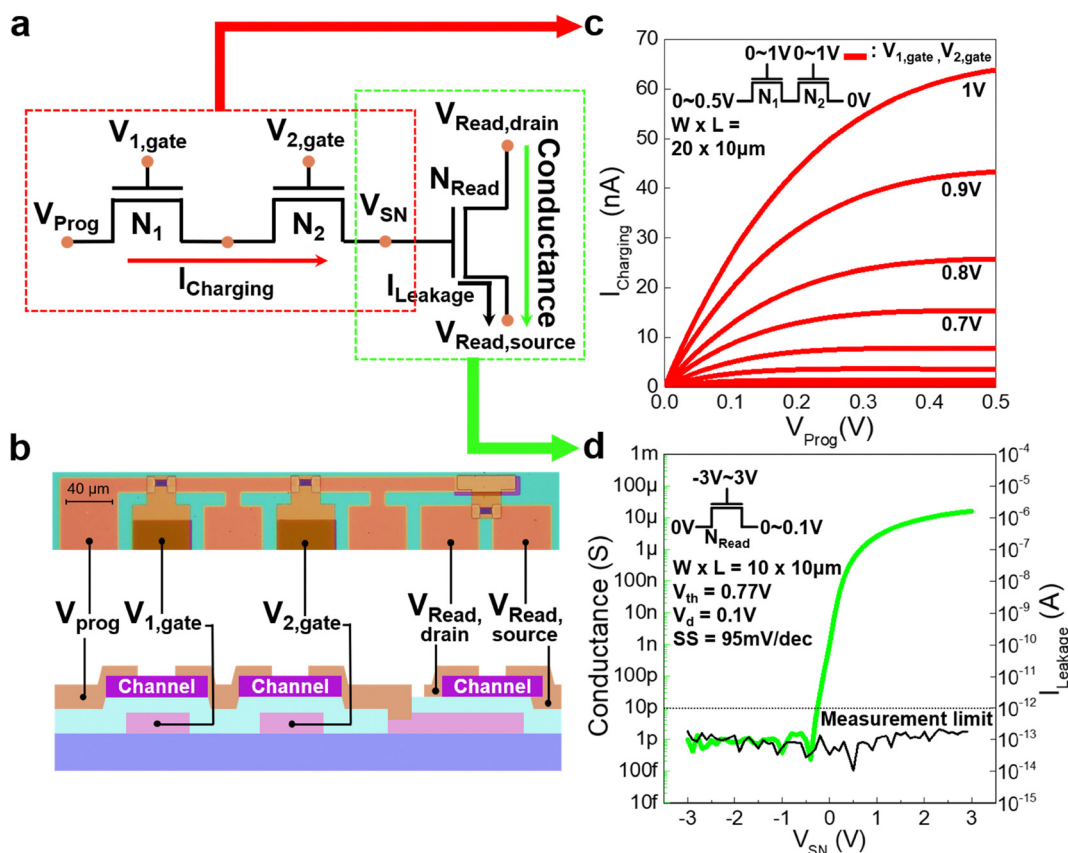
shown in Fig. S2 (ESI<sup>†</sup>), confirming the device structure with W electrodes, HfO<sub>2</sub> gate dielectric, and IGZO channel material. The IGZO transistor shown in the SEM image in Fig. S2 (ESI<sup>†</sup>) is a neighboring device fabricated on the same chip with a larger area of  $30 \mu\text{m} \times 40 \mu\text{m}$  ( $W \times L$ ) for a clearer EDS profile, compared to the smaller device shown in Fig. 2b, which is measured in this work. In addition, the vertical structure and thickness of the stacked layers were directly verified through cross-sectional TEM imaging, as shown in Fig. S3 (ESI<sup>†</sup>), which further confirms the well-defined gate stack consisting of W/HfO<sub>2</sub>/IGZO layers.

## 2.2. Transistor characteristics for TrioN cell operation

We analyzed the electrical characteristics of the NMOSs comprising TrioN to determine the optimal operating voltages and currents. We measured the current flowing through the two transistors ( $N_1, N_2$ ) connected in series as a function of the  $V_{\text{Prog}}$ ,  $V_{1,\text{gate}}$ , and  $V_{2,\text{gate}}$  voltages to investigate the charging and discharging currents, and the results are shown in Fig. 2c. We applied the voltage from 0 V to 1 V in 0.1 V steps to the  $N_{1,\text{gate}}$  and  $N_{2,\text{gate}}$  simultaneously, and from 0 V to 0.5 V in 0.01 V steps to the PN. Additionally, we measured the conductance of  $N_{\text{Read}}$  as a function of  $V_{\text{SN}}$  as depicted in Fig. 2d. We conducted measurements by sweeping the voltage across the gate of the  $N_{\text{Read}}$  from  $-3 \text{ V}$  to  $3 \text{ V}$  in 0.1 V steps, and read the conductance of the  $N_{\text{Read}}$  by applying 0.1 V to  $V_{\text{Read,drain}}$  and 0 V to  $V_{\text{Read,source}}$ . The fabricated  $N_{\text{Read}}$  transistor, with dimensions of  $10 \mu\text{m} \times 10 \mu\text{m}$ , a threshold voltage ( $V_{\text{th}}$ ) of 0.77 V and a

subthreshold swing (SS) value of  $95 \text{ mV dec}^{-1}$ , exhibits a low SS value below  $0.1 \text{ V dec}^{-1}$  comparable to those reported in recent studies on amorphous indium-gallium-zinc oxide (a-IGZO) thin film transistors that utilize HfO<sub>2</sub> as the gate dielectric with a similar thickness and  $k$ -value.<sup>39–41</sup> A low subthreshold swing value implies that conductance can increase by an order of magnitude with only a small change in the storage node voltage, meaning that a sufficient on/off ratio can be achieved even within a small voltage range. Fig. S4 (ESI<sup>†</sup>) presents a comprehensive analysis of the device-level reliability and electrostatic integrity of the TrioN cell. Fig. S4a (ESI<sup>†</sup>) illustrates the drain-induced barrier lowering (DIBL) and gate-induced drain leakage (GIDL) characteristics of the  $N_{\text{Read}}$ . To quantify DIBL, threshold voltages ( $V_{\text{th}}$ ) were extracted at multiple drain voltages ( $V_{\text{D}}$ ). The results show a  $V_{\text{th}}$  shift from 1.05 V at  $V_{\text{D}} = 0.1 \text{ V}$  to 0.99 V at  $V_{\text{D}} = 3.0 \text{ V}$ , corresponding to a DIBL value of approximately  $20.0 \text{ mV V}^{-1}$ . This value is significantly lower than those reported for conventional oxide semiconductor thin-film transistors, indicating excellent short-channel control.<sup>42–44</sup> Although the minimum measurement limit is about  $1 \times 10^{-12} \text{ A}$ , the drain current remained close to this lower bound even when relatively high gate voltages were applied in the off-state. This suggests that gate-induced leakage is effectively suppressed in the TrioN architecture, supporting its potential for low-power and stable retention operation. Fig. S4b (ESI<sup>†</sup>) shows the device-to-device (D2D) variation characteristics extracted from 17 fabricated TrioN cells. The threshold voltage and subthreshold swing values exhibit minimal deviation, demonstrating excellent uniformity





**Fig. 2** (a) A schematic and (b) optical microscopy and cross-sectional view image of the TrioN device structure for evaluating the operating characteristics of each transistor comprising the TrioN cell. (c)  $V_{\text{Prog}}-I_{\text{Charging}}$  graph for two transistors ( $N_1, N_2$ ) in series.  $V_{1,\text{gate}}$  and  $V_{2,\text{gate}}$  applied to the gate of  $N_1$  and  $N_2$  were from 0 V to 1 V in 0.1 V steps. (d)  $V_{\text{SN}}-G$  (conductance) graph for  $N_{\text{Read}}$ . The  $V_{\text{Read,drain}}$  was 0.1 V and the  $V_{\text{Read,source}}$  was 0 V. The threshold voltage was measured to be 0.77 V and the subthreshold swing was 95 mV dec<sup>-1</sup>. Furthermore, the off-current and gate leakage current of the IGZO transistor were measured below the measurement limit.

and stability of multiple devices which is an essential requirement for array-level integration. Lastly, Fig. S4c (ESI<sup>†</sup>) provides the transfer characteristics of the TrioN read transistor under positive gate bias stress conditions (1 V for 10 s, 100 s, and 1000 s). To enable multi-level updates in the TrioN device, fine-tuning of the current flowing through the write transistor is required. This means that the voltage and duration of the pulses applied to the gates of the two write transistors must be precisely controlled. Since a switching characteristic in Fig. 3d is conducted by applying a voltage of 0.7 V, a larger voltage of 1 V was used to perform a positive bias stress test, the results of which are presented in Fig. S4c (ESI<sup>†</sup>). As a result, no significant shift in  $V_{\text{th}}$  or degradation in the subthreshold swing is observed, indicating strong temporal reliability and bias-stress resilience. Research on a-IGZO transistors has primarily focused on achieving a high on/off current ratio, low subthreshold swing, and high mobility, thereby maximizing the on-current while minimizing the off-current. To enhance the on-current and mobility, various techniques such as reducing the contact resistance between the source/drain metal and oxide channel, utilizing thin high- $k$  gate dielectrics, and applying post-deposition annealing have been extensively investigated.<sup>45-47</sup> However, in the case of our TrioN device, such optimization for a high on-current or mobility is not

necessary. To minimize gate leakage current which could negatively affect the characteristics of the TrioN device, we employed a 20 nm-thick HfO<sub>2</sub> gate dielectric. This dielectric thickness significantly influences the carrier concentration, mobility, and on-current of the channel. We extracted the field-effect mobility of 8.34 cm<sup>2</sup> V<sup>-1</sup> s<sup>-1</sup>, using the extrapolation in the linear region (ELR) method, which falls within the typical range of reported a-IGZO devices. For comparison under the same bias condition with 0.1 V drain voltage and 1 V gate voltage, our device showed a drain current of 26 nA μm<sup>-1</sup>, while other works reported approximately in a range of 10–20 nA μm<sup>-1</sup>.<sup>39,40</sup> The low off-current reduces leakage during the hold state, which improves retention performance, and the low on-current which is a characteristic of a-IGZO transistors even without employing mobility or on-current boosting strategies helps to minimize charging currents to the storage node, enabling multi-level updates with fine granularity. For instance, when an on-current of 70 nA is supplied to the storage node for a duration of 10 μs, the resulting charge is calculated to be 7 × 10<sup>-14</sup> C. Given the 1.77 pF capacitance of our device's storage node, this corresponds to a voltage increase of 0.04 V. As depicted in Fig. 2c, the TrioN cell operates within a voltage range where the on-current does not exceed 70 nA, facilitating finer control over the



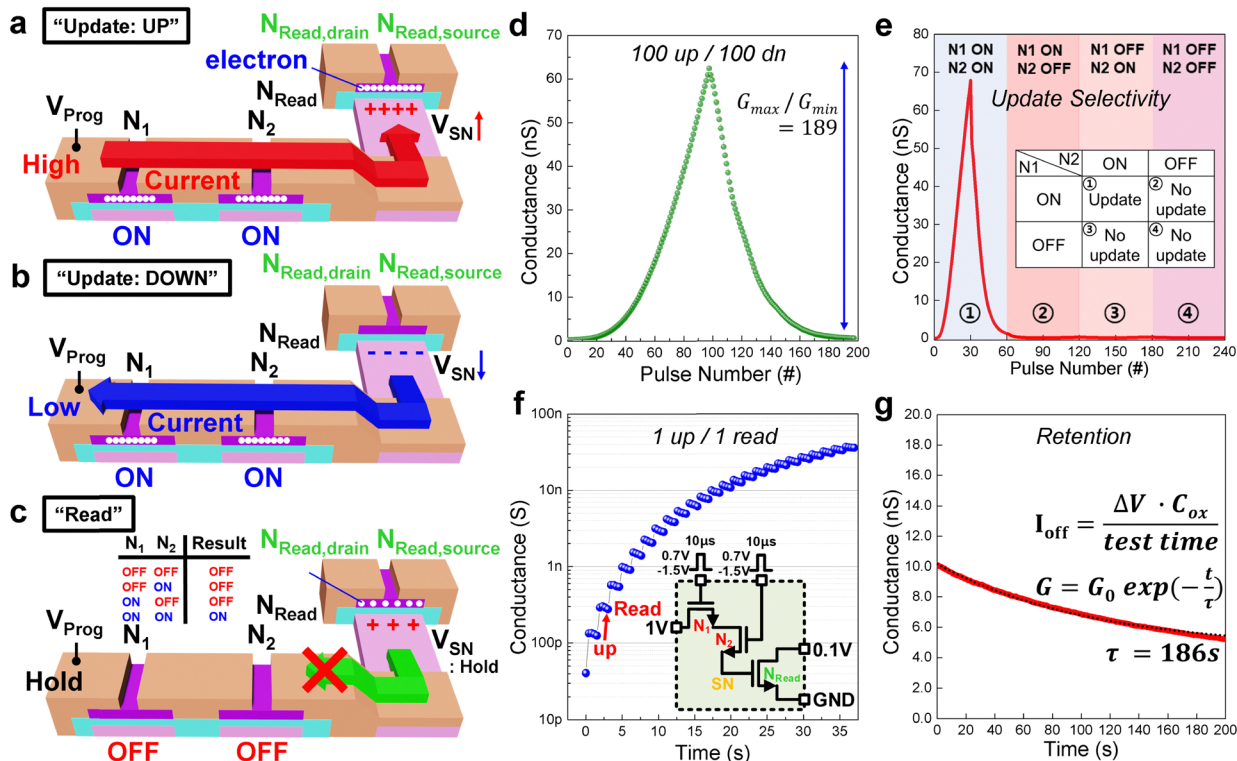


Fig. 3 Operation modes and device characteristics of a TrioN cell. (a) Update: UP operation of TrioN.  $V_{\text{Prog}}$  is 'High',  $N_1$  and  $N_2$  are turned ON. (b) Update: DOWN operation of TrioN.  $V_{\text{Prog}}$  is 'Low',  $N_1$  and  $N_2$  are turned ON. (c) Hold operation of TrioN.  $V_{\text{Prog}}$  is 'Hold', either  $N_1$  or  $N_2$  are OFF. (d) The switching characteristics for 100 up/100 down switching.  $N_1$  and  $N_2$  are ON at  $V_{1,\text{gate}}$ ,  $V_{2,\text{gate}} = 0.7\text{ V}$  in 100 ns and  $V_{\text{Prog}}$  is set to 1 V. The  $G_{\text{max}}/G_{\text{min}}$  ratio for 100 up/100 down is 189. (e) Updated selectivity of the TrioN cell. The TrioN cell only updates when both  $N_1$  and  $N_2$  are ON, and does not update when either  $N_1$  or  $N_2$  is OFF. (f) Storage node charging and hold operation using 1 up/1 read pulse schematics. (g) Retention characteristics of the TrioN cell.

voltage stored in the storage node. Furthermore, as the storage node voltage increases, the required charge supply decreases, which benefits the TrioN cell by supporting a higher number of states.

### 2.3. Operation mechanism and switching characteristics of the TrioN cell

For a synaptic device-based cross-point array to function effectively, three operations are critical: forward propagation, backward propagation, and weight update. In the context of TrioN, the update operation modifies the conductance value representing the synaptic weight. The potentiation (storage node charging), depression (storage node discharging), and retention (storage node holding) working principles are demonstrated in Fig. 3a–c. The potentiation is achieved by setting  $V_{\text{Prog}}$  to 'High' and applying short voltage pulses simultaneously to the gates of  $N_1$  and  $N_2$  via the  $N_{1,\text{gate}}$  and  $N_{2,\text{gate}}$ , turning both transistors ON.

When a high voltage is applied to  $V_{\text{Prog}}$ , charging current flows through  $N_1$  and  $N_2$  to charge the storage node, causing the storage node voltage to increase. This, in turn, activates the channel of  $N_{\text{Read}}$ , resulting in an increase in conductance. In contrast, the depression is performed by setting  $V_{\text{Prog}}$  to 'Low' and applying short voltage pulses to  $N_{1,\text{gate}}$  and  $N_{2,\text{gate}}$  to turn them ON. This causes discharging current flow from the storage node to PN, causing the storage node voltage to decrease and reducing the conductance of  $N_{\text{Read}}$ 's channel. The storage node

behavior follows the relationship  $Q = I \cdot \Delta t$ , where the amount of charge ( $Q$ ) depends on the charging/discharging current ( $I$ ) and the pulse duration ( $\Delta t$ ) during the ON state of  $N_1$  and  $N_2$ . By controlling these parameters, the TrioN exhibits the properties of an analog synaptic device. Using the mechanisms described above, 100 up/100 down voltage pulses were applied to the TrioN, as shown in Fig. 3d, demonstrating its analog switching characteristics. High and low voltages of 1 V and 0 V were applied to  $V_{\text{Prog}}$  and 0.7 V pulses with a duration of 100 ns were applied to the gates of  $N_1$  and  $N_2$  to turn them ON while applying a voltage of  $-1.5\text{ V}$  to both transistors to turn them OFF normally. The storage node voltage, which corresponds to the gate voltage of  $N_{\text{Read}}$ , determines the conductance value and exhibits exponential increases and decreases during charging and discharging operations. Key parameters influencing the storage node voltage include the  $V_{\text{Prog}}$ ,  $V_{1,\text{gate}}$  and  $V_{2,\text{gate}}$  pulse voltages and pulse widths. A higher  $V_{\text{Prog}}$  voltage increases the current through  $N_1$  and  $N_2$  as shown in Fig. 2c, while larger gate pulse amplitudes or longer durations also increase the total charge stored in the storage node. However, by optimizing various parameters including the  $V_{\text{Prog}}$  voltage, the gate pulse amplitude and the pulse width, the TrioN device in this work achieves consistent and precise charge increments at the storage node, enabling multi-state conductance levels beyond binary ON/OFF states. Experimental results reveal a  $G_{\text{max}}/G_{\text{min}}$  ratio of 189 with 100 ns pulse widths, and Fig. S5 (ESI<sup>†</sup>) shows conductance asymmetry under



13%, indicating near-ideal symmetry. Furthermore, Fig. S6 (ESI<sup>†</sup>) demonstrates that the TrioN cell can respond to pulses as short as 10 ns, which is even shorter than 100 ns, enabling updates. Additionally, Fig. S7 (ESI<sup>†</sup>) shows that the cell consistently maintains stable conductance values while switching under 100 up and 100 down pulses across 10 cycles. Fig. 3e demonstrates AND logic-based selective updates through independent control of  $N_1$  and  $N_2$  gates, where updates occur only when both  $N_1$  and  $N_2$  are ON. No updates occur if either transistor is OFF. This property ensures that the TrioN can achieve fully parallel updates in a cross-point array. To maintain conductance values over time, a storage node hold operation is essential. Fig. 3c illustrates this operation, where turning OFF either  $N_1$  or  $N_2$  preserves the storage node voltage. To confirm the update and hold operations of the TrioN device, we conducted a 1 update and 1 read operation test. A 0.7 V voltage pulse with a pulse width of 10  $\mu$ s was simultaneously applied to the gates of  $N_1$  and  $N_2$ , allowing a charging current to flow. At the same time, 1 V was applied to the programming node to initiate storage node charging. As shown in Fig. 3f, we observed that the conductance increased following the update operation and was successfully maintained during the read operation. The retention characteristics were evaluated by holding the storage node voltage for 200 s, with results presented in Fig. 3g, and the time constant ( $\tau$ ) was calculated as 186 s. Due to the operational mechanism of our TrioN device, higher programming voltages applied to the programming node (PN) result in greater charge accumulation at the storage node. Consequently, the electrostatic tendency for the stored charge to discharge at the storage node becomes stronger, inherently leading to larger conductance decay in the higher conductance regime compared to the lower conductance regime. Despite this characteristic, additional retention measurements are shown in Fig. S8 (ESI<sup>†</sup>), confirming that the conductance remains stable for 50 s even in a significantly higher range than that presented in Fig. 3b.

To elucidate the origin of conductance decay in TrioN, we first note that the synaptic conductance is determined by the current flowing through the channel of the  $N_{\text{Read}}$ . The loss of storage node charge arises from leakage currents through both the  $N_2$  and the  $N_{\text{Read}}$ . In  $N_2$ , leakage can occur *via* the  $\text{HfO}_2$  gate dielectric and the off-state channel, and in  $N_{\text{Read}}$ , it can occur *via* the gate dielectric. However, in Fig. 2d, since the 20 nm  $\text{HfO}_2$  thickness in all three NMOS devices yields a gate leakage current that is effectively independent of gate bias, and the extracted off-state current in Fig. 3g is negligible, both gate dielectric leakage and off-state channel leakage were excluded from our retention analysis. Instead, we attribute the dominant charge-loss mechanism to trap-assisted tunneling (TAT) through oxide traps in the  $\text{HfO}_2$  gate dielectric of  $N_2$  and  $N_{\text{Read}}$ , as well as the oxygen vacancies in the a-IGZO channel of  $N_2$ . Both the a-IGZO channel and  $\text{HfO}_2$  dielectric are known to host oxygen vacancy and other deep traps, which capture and subsequently release electrons. This TAT pathway provides a leakage path from the storage node into the electrodes of each transistor, thereby producing the retention decay.<sup>48</sup> Similar TAT-dominated retention loss has been widely reported in

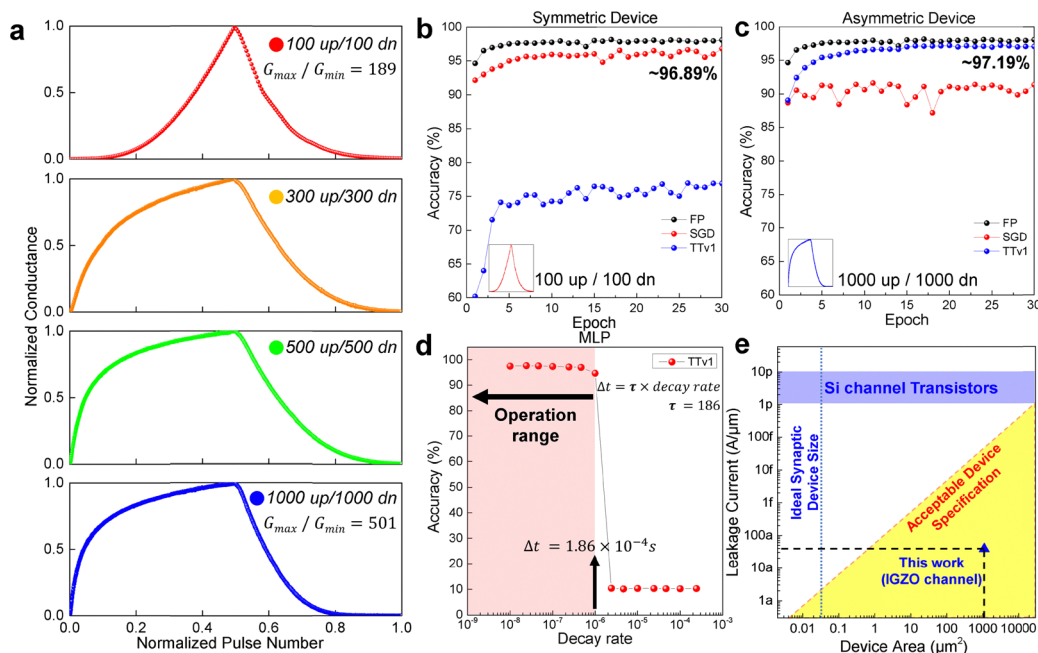
1T1C and 2T0C DRAM cells,<sup>49</sup> as well as in modern charge-trap NAND Flash memories,<sup>50</sup> underscoring the generality of oxide trap-based decay in memory devices.

We also extracted the time constant at different states to analyze the retention behavior of TrioN. In the higher conductance range, the time constant was measured to be in the range of 122 to 186, indicating faster conductance decay due to a stronger tendency for electron discharge. In addition, based on the retention characteristics shown in Fig. 3g and using the  $I_{\text{off}}$  extraction criterion, the calculated  $I_{\text{off}}$  value was approximately 39 aA  $\mu\text{m}^{-1}$ .<sup>51</sup> There have been several studies reporting extremely low off-current and the mechanisms behind it, particularly focusing on the role of various defects in the oxide channel that act as traps and significantly influence the off-leakage of oxide FETs.<sup>52,53</sup> The off-leakage of the device fabricated in this study can be further reduced by passivating the device or adjusting the amount and behavior of defects in the active layer through optimization in process conditions.

#### 2.4. Controllable asymmetry characteristics of the TrioN cell and neural network simulation

Unlike the previously proposed 3T1C cell, which uses PMOS and NMOS for charging and discharging, respectively,<sup>16</sup> the TrioN employs two series-connected NMOSs to handle both charging and discharging operations. As shown in Fig. 3a, during storage node charging, when  $V_{\text{Prog}}$  is set to 'High', the storage node voltage charges. As the storage node voltage increases, the charging current through  $N_1$  and  $N_2$  gradually decreases as the storage node voltage approaches the 'High' voltage. Conversely, during depression, as shown in Fig. 3b, when  $V_{\text{Prog}}$  is set to 'Low', the voltage difference between the initially 'High' storage node voltage and the 'Low'  $V_{\text{Prog}}$  results in a large amount of discharging current through  $N_1$  and  $N_2$  at the beginning of the depression. This discharging current decreases as the storage node voltage approaches the 'Low' voltage. Consequently, the TrioN inherently exhibits asymmetry, as evidenced by the switching results shown in Fig. 4a, where 300 up/300 down (orange), 500 up/500 down (green), and 1000 up/1000 down (blue) pulses demonstrate asymmetric behavior. However, if the storage node voltage operates within a much narrower range instead of charging and discharging across the full range from 0 V to 'High', the device can exhibit symmetric operation. As observed with 100 up/100 down pulses (red) in Fig. 4a, the operation of the device is confined to a narrow range rather than spanning the full range, yet it exhibits greater symmetry compared to devices subjected to a larger number of pulses. This narrow range can be attributed to the conductance maximum-to-minimum ratio observed in the 100 up/100 down case ( $G_{\text{max}}/G_{\text{min}} = 189$ ) being significantly smaller than the ratio in the 1000 up/1000 down case ( $G_{\text{max}}/G_{\text{min}} = 501$ ). Using the TrioN, which can display both asymmetric and symmetric characteristics depending on the operational range, we conducted simulations with various learning algorithms, including stochastic gradient descent (SGD) and Tiki-Taka algorithm version 1 (TTv1).<sup>54–56</sup> The TTv1 algorithm is specifically designed to account for and overcome the asymmetry inherent in the device. By leveraging the IBM Analog Hardware Acceleration





**Fig. 4** (a) The normalized conductance for 100, 300, 500, and 1000 up and down pulses. 100 up/100 down switching result shows perfect symmetry with on/off ratios of 189 and 300, 500, 1000 up and down switching results show an asymmetry, with an on/off ratio of 501 in 1000 up and down switching. (b) Simulation result with symmetric characteristic (100 up/100 down) and (c) Simulation result with asymmetric characteristic (1000 up/1000 down). All simulations are MLP learning simulations on the MNIST dataset, showing higher accuracy for symmetric characteristics using the SGD algorithm and asymmetric characteristics using the TTV1 algorithm. (d) Decay rate vs. inference accuracy showing nearly 100% accuracy for decay rates below  $10^{-6}$ . (e) Device specifications for array-level operation, with the fabricated IGZO based device ( $1100 \mu\text{m}^2$ ,  $39 \text{ aA } \mu\text{m}^{-1}$ ) falling within the acceptable range. Scalability to  $0.04 \mu\text{m}^2$  with leakage currents below 100 aA is achievable.

KIT (AIHWKIT),<sup>57</sup> parameters such as linearity and asymmetry were extracted from the operational characteristics of the TrioN. These parameters were then used to perform multi-layer perceptron (MLP) learning simulations on the Modified National Institute of Standards and Technology (MNIST) dataset. The neural network architecture consisted of two hidden layers with a structure of 784-256-128-10.

Fig. 4b and c present the MNIST pattern recognition simulation results based on the switching characteristics of symmetric and asymmetric devices, respectively, using the SGD and TTV1 algorithms. The simulation results indicate that a symmetric device, as characterized by 100 up/100 down pulses, achieved over 95% accuracy within 5 epochs using the SGD algorithm and reached 96.89% accuracy after 30 epochs, as shown in Fig. 4b. In contrast, an asymmetric device, characterized by 1000 up/1000 down pulses, achieved a maximum accuracy of only 92% with the SGD algorithm, even after 30 epochs. However, with the TTV1 algorithm, the asymmetric device achieved a high accuracy of 97.19%, nearly equivalent to that of a floating-point (FP) device. These results emphasize that the TrioN cell demonstrates near-perfect symmetry, enabling it to achieve high neural network training accuracy even with the SGD algorithm. Moreover, the TrioN cell exhibits a high on/off ratio and supports sufficient multi-states with small voltage amplitudes and short voltage pulses. Additionally, the number of states can be adjusted to control asymmetry, and even when asymmetry occurs, the TrioN cell operates effectively to achieve high training accuracy when leveraging the TTV1 algorithm. In

neural network simulations, the retention characteristics of the device are critical to accuracy. The TrioN cell leverages the intrinsic gate capacitance of its read transistor formed *in situ* during device fabrication, thereby obviating the need for an external capacitor for charge storage. As the retention convergence point (RCP) of the TrioN device is zero, the storage node tends to discharge regardless of its charge state. To quantify the impact of the retention characteristic on network performance, we extracted the conductance decay rate from the TrioN's retention profile in Fig. 3g and incorporated the resulting decay rate into our simulation model.

## 2.5. Learning results and acceptable device specification

Fig. 4d illustrates the relationship between the decay rate and inference accuracy, based on simulations conducted using a multi-layer perceptron (MLP) model and the MNIST dataset. The decay rate, defined as the ratio of switching speed ( $\Delta t$ ) to retention ( $\tau$ ), represents the relation between device switching speed and data retention. The simulation results indicate that smaller decay rates lead to higher inference accuracy. Specifically, decay rates below  $10^{-6}$  result in nearly 100% accuracy, as shown in Fig. 4d. This finding suggests that the switching speed of the device should be at least  $1.86 \times 10^{-4}$  s to achieve accurate inference. Experimental data confirm that the switching speed of the fabricated device is 10 ns, which is well within the operational range required for maintaining high accuracy. These results validate the feasibility of using the device for neuromorphic computing applications. Fig. 4e evaluates whether the



device satisfies the specifications required for array-level operation. The graph depicts the relationship between the device area and the leakage current required for synaptic transistors, derived from redefined capacitor specifications tailored to the characteristics of the device. According to previous studies, among the device specifications required for neural network applications, capacitor-based synaptic cells must exhibit a leakage current below a certain threshold depending on the device area. This boundary is defined and visualized as the yellow region in Fig. 4e, representing the acceptable device specification.<sup>12,21</sup> The device, which consists of both a read transistor and a write transistor, occupies an area of approximately  $1100 \mu\text{m}^2$  and exhibits a leakage current of  $39 \text{ aA } \mu\text{m}^{-1}$ , based on IGZO channel properties. The plotted data confirms that this configuration lies within the acceptable device specification region, represented by the yellow area. Furthermore, scaling down the device to the ideal synaptic device size ( $0.04 \mu\text{m}^2$ ) requires a leakage current below  $100 \text{ aA}$ . This level of leakage cannot be achieved with conventional Si transistor channels due to their relatively high off-state current,<sup>22</sup> highlighting the necessity of using oxide semiconductor-based transistors instead.<sup>21</sup> The intrinsic characteristics of IGZO channels meet this requirement, indicating that the device is capable of scaling to meet the stringent demands of large-scale synaptic arrays without compromising performance.

Although this work focuses on demonstrating the concept and operational validity of the TrioN cell, its potential advantages in power/performance/area/cost (PPAC) can be reasonably estimated based on device-level measurements and known properties of the IGZO-based transistors. The TrioN device exhibits a low off-current ( $39 \text{ aA } \mu\text{m}^{-1}$ ) and fast switching speed ( $10 \text{ ns}$ ), indicating excellent energy efficiency. Furthermore, its capacitorless architecture and use of only NMOS transistors enable simpler fabrication, reduced cell area, and cost-effective integration. Owing to its distinctive characteristics, the TrioN demonstrates the promising potential to outperform conventional capacitor-based synaptic cells across all PPAC metrics, particularly when optimized and scaled for integration into future system-level architectures.

## 2.6. 2-Cycle update: a new strategy for efficient update in cross-point arrays

Fully parallel computation of matrix-vector multiplication with vector-vector outer product operation requires a cross-point array architecture. In Fig. S9 (ESI<sup>†</sup>), we evaluated the scalability of the TrioN cell into a  $2 \times 2$  array through array simulation, showing the feasibility of fully parallel updates without interference between neighboring devices. Furthermore, by applying update pulses, we successfully showed that a consistent amount of current flowed through  $N_1$  and  $N_2$  to charge or discharge the storage node, effectively maintaining each conductance state with high stability. In Fig. 5a, we assume a TrioN based  $n \times m$  cross-point array architecture. The TrioN cells located in the same row share  $V_{1,\text{gate}}$  and PL through  $\text{Write}_{\text{row},i}$  and  $\text{PL}_i$  ( $i = 1, 2, \dots, n$ ), respectively. Similarly, TrioN cells in the same column share  $V_{2,\text{gate}}$  through  $\text{Write}_{\text{col},j}$  ( $j = 1, 2, \dots, m$ ) with each individual TrioN cell. In Fig. 5b, we illustrated the method for

updating  $\text{Cell}_{ij}$ , which is located at the  $i$ th row and  $j$ th column, using diagrams. Each cell requires simultaneous pulses at its connected  $\text{Write}_{\text{row}}$  and  $\text{Write}_{\text{col}}$  terminals for an update to occur. The value entering from the PL determines whether the TrioN cell's state will increase (UP) or decrease (DOWN) based on whether the voltage is 'High' or 'Low', respectively. This update mechanism enables parallel updates without interference from neighboring cells. The selective update mechanism has been validated through simulations on a  $2 \times 2$  cell array, as shown in Fig. S9 (ESI<sup>†</sup>), demonstrating selective updates. If the selective updates are successfully achieved in the array, a stochastic update scheme can be employed to enable parallel training of the array.<sup>58</sup> The stochastic update scheme is a method for parallel updates of analog arrays, enabling matrix multiplication by applying random pulses to rows and columns at frequencies proportional to their respective values. In the cross-point array architecture, the update process involves applying the outer product of the input vector  $x$  and the error vector  $\delta$  to the synaptic devices, thereby incrementally adjusting the synaptic weights stored in the array. Since both  $x$  and  $\delta$  are vectors of rank 1, the resulting weight update matrix  $\Delta W$  also has rank 1. This process can be expressed mathematically as:

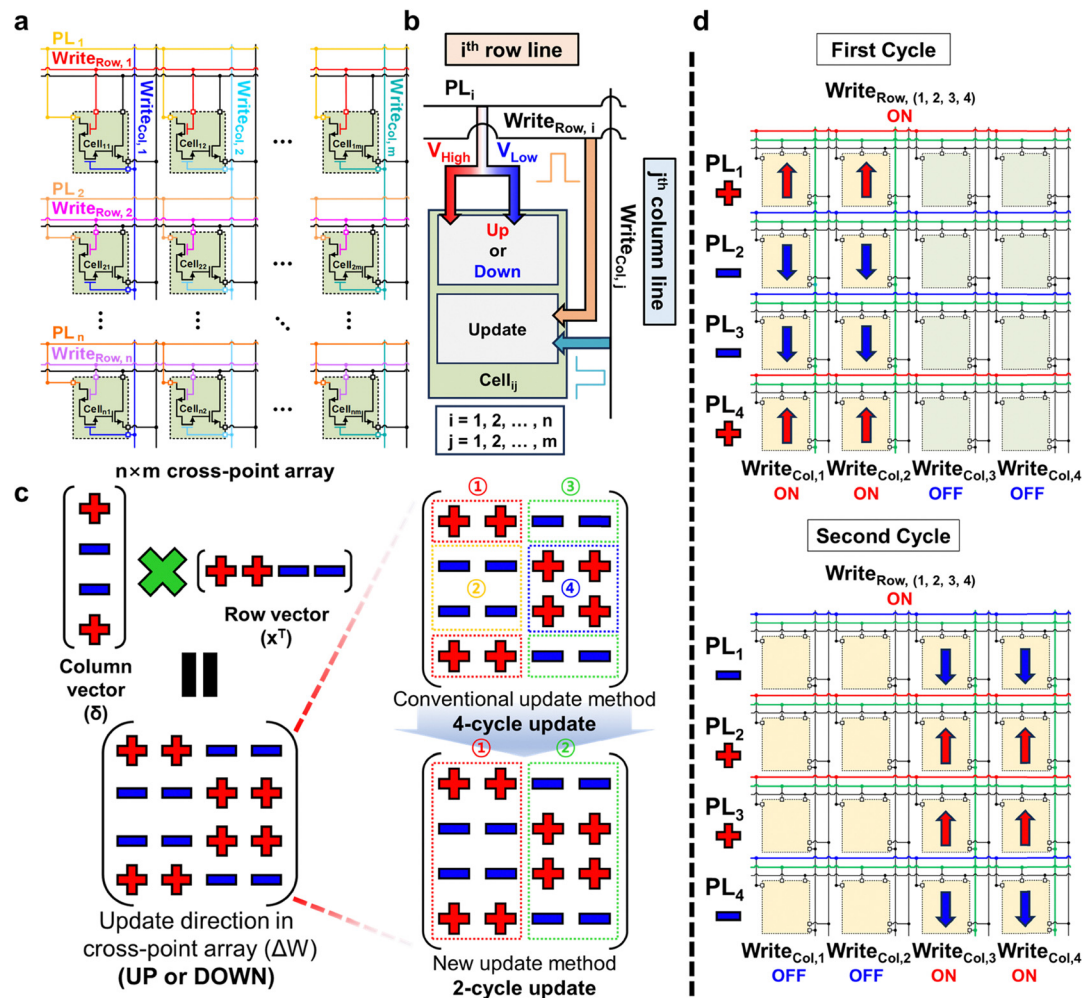
$$\Delta W = \eta \cdot \delta \cdot x^T \quad (1)$$

where  $\eta$  denotes the learning rate.

When all row and column values are positive, the entire array can be updated in a single pulse cycle. However, when both positive and negative values are present in rows or columns, four separate pulse cycles are required in arrays composed of devices with a conventional selective update scheme. In such cases, rows and columns are grouped based on their polarity combinations, specifically (+,+), (+,-), (-,+), and (-,-), and updated sequentially. This approach ensures that individual components within the array are updated without interference. Fig. 5c illustrates the sequence of updates in the conventional method, which requires four pulse cycles to complete. While effective, this method has the significant drawback of extended update times, limiting its overall efficiency.<sup>59</sup>

TrioN devices address this limitation through their unique architecture, where the switching mechanism is decoupled, which implies that the update direction and the selection of the device to be updated are handled independently. Specifically, the device separates the activation and deactivation functionality, controlled by the  $V_{1,\text{gate}}$  and  $V_{2,\text{gate}}$ , from the mechanisms that determine UP and DOWN, controlled by the PL. This decoupling enables a 2-cycle update scheme, significantly improving the efficiency of array updates. Fig. 5d demonstrates the implementation of this 2-cycle update process. In the first cycle, all  $\text{Write}_{\text{row}}$  are activated, while signals for the  $\text{Write}_{\text{col}}$  are enabled only for rows with positive values in the row vector. For the column vector, positive values are subjected to a 'High' voltage at PL, while negative values receive a 'Low' voltage. In the second cycle, all  $\text{Write}_{\text{row}}$  remain activated, while  $\text{Write}_{\text{col}}$  signals are enabled only for rows with negative values in the row vector. For the column vector, positive values receive a 'Low' voltage at PL, while negative values receive a





**Fig. 5** (a)  $n \times m$  cross-point array schematic based on TrioN cells. Programming lines  $PL_1, PL_2, \dots, PL_n$  are lines that connect the  $V_{Prog}$  terminals of individual cells row by row. (b) The method of applying voltage pulses to update a single TrioN cell (c) A comparison of the conventional 4-cycle update and the new 2-cycle update for effectively updating all 16 devices in a  $4 \times 4$  cell array. (d) Illustration of the 2-cycle  $n \times m$  update method implemented in the TrioN device.

'High' voltage. This 2-cycle method effectively reduces the update time by half compared to the conventional 4-cycle approach. Such a significant reduction in update time underscores the potential of TrioN devices in accelerating array operations and enhancing the overall efficiency of neural network computations.

To further highlight the advantages of our TrioN cell, Table 1 benchmarks its key metrics such as device count, fast switching speed, selective update capability, and retention against the conventional capacitor-based cells. As shown, our TrioN combines a capacitor-less architecture with the fastest switching

speed and competitive retention characteristic. These features make TrioN a promising synaptic cell candidate for high-density, highly integrated cross-point arrays in analog process-in-memory systems.

### 3. Conclusion

This study presents the TrioN architecture, an innovative capacitorless resistive processing unit utilizing IGZO thin-film transistors for analog process-in-memory (aPIM) applications. The TrioN addresses critical challenges in aPIM systems,

**Table 1** A quantitative comparison between the conventional capacitor-based cell and TrioN

	Y. Li <i>et al.</i> <sup>18</sup>	S. Park <i>et al.</i> <sup>28</sup>	J. Won <i>et al.</i> <sup>31</sup>	This work
Transistor	9 (9PMOS + 4NMOS)	2 (NMOS only)	6 (NMOS only)	3 (NMOS only)
Capacitor	1	0	1	0
Minimum update pulse width	50 ns	16–21 ns	300 ns	10 ns
Selective update	O	X	O	O
Retention	Few seconds	300 s [ $\Delta V_{SN} = 10\%$ ]	7200 s [ $\Delta V_{SN} = 0.1$ V]	200 s [ $\Delta V_{SN} = 0.03$ V]



including high leakage currents and scalability limitations, by utilizing the overlap capacitance of an IGZO NMOS as a storage node capacitance. Experimental validation demonstrates that the TrioN exhibits exceptional characteristics, meeting the requirements for synaptic devices, including perfect symmetry, fast switching and selective update mechanisms, which are critical for efficient and parallel operation in cross-point arrays. Furthermore, simulations on neural network acceleration with the MNIST dataset highlight the architecture's compatibility with advanced learning algorithms, achieving high accuracy. The retention and decay rate analyses also confirm that the TrioN meets the stringent requirements for neuromorphic computing, with scalability prospects validated by its low leakage current and compact footprint. These results establish the TrioN as a promising technology for scalable, energy-efficient analog AI systems, paving the way for future advancements in neuromorphic computing and artificial intelligence hardware.

## Author contributions

J. C. and B. L. contributed equally to the conceptualization, investigation, methodology, the development of software and validation of this work. J. C. took the lead in formal analysis, data curation and visualization with the support of B. L. and H. K. J. C. and B. L. performed the experiments, and measurements with the support of J. B., H. K., and S. K., and J. L., J. C. and B. L. wrote the original draft, while J. S., J. B., and Prof. S. K. contributed to reviewing and editing. Prof. S. K. directed the team and led the funding acquisition, project administration, resource management, and supervision. All authors have read and approved the final manuscript.

## Data availability

The authors declare that the data supporting the findings of this study are available within the paper and its ESI.†

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgements

This work is supported by K-CHIPS (Korea Collaborative & High-tech Initiative for Prospective Semiconductor Research) (2410010977, RS-2023-00235402, 23003-15FC) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea), and National R&D Program through the National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT (2410000283, RS-2024-00405960, 23008-45FC). J. C. gratefully acknowledges the financial support provided by the Hyundai Motor Chung Mong-Koo Foundation through a scholarship. The EDA tool was supported by the IC Design Education Center (IDEC), Korea.

## Notes and references

- 1 Y. LeCun, Y. Bengio and G. Hinton, *Nature*, 2015, **521**, 436–444.
- 2 L. Chang, C. Li, Z. Zhang, J. Xiao, Q. Liu, Z. Zhu, W. Li, Z. Zhu, S. Yang and J. Zhou, *Sci. China Inf. Sci.*, 2021, **64**, 160403.
- 3 M. Shaafiee, R. Logeswaran and A. Seddon, 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence, 2017, pp. 199–203.
- 4 G. Santoro, G. Turvani and M. Graziano, *Micromachines*, 2019, **10**, 368.
- 5 N. Verma, H. Jia, H. Valavi, Y. Tang, M. Ozatay, L.-Y. Chen, B. Zhang and P. Deaville, *IEEE Solid-State Circuits Mag.*, 2019, **11**, 43–55.
- 6 Q. Xia and J. J. Yang, *Nat. Mater.*, 2019, **18**, 309–323.
- 7 G. Pedretti and D. Ielmini, *Electronics*, 2021, **10**, 1063.
- 8 C.-X. Xue, T.-Y. Huang, J.-S. Liu, T.-W. Chang, H.-Y. Kao, J.-H. Wang, T.-W. Liu, S.-Y. Wei, S.-P. Huang and W.-C. Wei, *et al.*, 2020 IEEE International Solid-State Circuits Conference-(ISSCC), 2020, pp. 244–246.
- 9 M. Suri, O. Bichler, D. Querlioz, O. Cueto, L. Perniola, V. Sousa, D. Vuillaume, C. Gamrat and B. DeSalvo, 2011 International Electron Devices Meeting, 2011, pp. 4–4.
- 10 S. Jung, H. Lee, S. Myung, H. Kim, S. K. Yoon, S.-W. Kwon, Y. Ju, M. Kim, W. Yi and S. Han, *et al.*, *Nature*, 2022, **601**, 211–216.
- 11 M. Jerry, P.-Y. Chen, J. Zhang, P. Sharma, K. Ni, S. Yu and S. Datta, 2017 IEEE international electron devices meeting (IEDM), 2017, pp. 6–2.
- 12 T. Gokmen and Y. Vlasov, *Front. Neurosci.*, 2016, **10**, 333.
- 13 R. Islam, H. Li, P.-Y. Chen, W. Wan, H.-Y. Chen, B. Gao, H. Wu, S. Yu, K. Saraswat and H. P. Wong, *J. Phys. D: Appl. Phys.*, 2019, **52**, 113001.
- 14 Y. Xiang, P. Huang, Y. Zhao, M. Zhao, B. Gao, H. Wu, H. Qian, X. Liu and J. Kang, *IEEE Trans. Electron Devices*, 2019, **66**, 4517–4522.
- 15 S. Yu, *Proc. IEEE*, 2018, **106**, 260–285.
- 16 S. Kim, T. Gokmen, H.-M. Lee and W. E. Haensch, 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), 2017, pp. 422–425.
- 17 Y. Li, S. Kim, X. Sun, P. Solomon, T. Gokmen, H. Tsai, S. Koswatta, Z. Ren, R. Mo and C. C. Yeh, *et al.*, 2018 IEEE Symposium on VLSI Technology, 2018, pp. 25–26.
- 18 S. Ambrogio, P. Narayanan, H. Tsai, R. M. Shelby, I. Boybat, C. Di Nolfo, S. Sidler, M. Giordano, M. Bordini and N. C. Farinha, *et al.*, *Nature*, 2018, **558**, 60–67.
- 19 Y. Luo and S. Yu, *IEEE Trans. Comput.*, 2020, **69**, 1113–1127.
- 20 Y. Kohda, Y. Li, K. Hosokawa, S. Kim, R. Khaddam-Aljameh, Z. Ren, P. Solomon, T. Gokmen, S. Rajalingam and C. Baks, *et al.*, 2020 IEEE International Electron Devices Meeting (IEDM), 2020, pp. 36–2.
- 21 B. Lee, W. Ji, H. Kim, S. Han, G. Park, P. Hur, G. Jeon, H.-M. Lee, Y. Chung and J. Son, *et al.*, *Adv. Intell. Syst.*, 2024, 2400600.
- 22 K. Roy, S. Mukhopadhyay and H. Mahmoodi-Meimand, *Proc. IEEE*, 2003, **91**, 305–327.



- 23 Q. Hu, C. Gu, S. Zhu, Q. Li, A. Tong, J. Kang, R. Huang and Y. Wu, *IEEE Electron Device Lett.*, 2022, **44**, 60–63.
- 24 A. Belmonte, H. Oh, N. Rassoul, G. Donadio, J. Mitard, H. Dekkers, R. Delhougne, S. Subhechha, A. Chasin and M. Van Setten, *et al.*, 2020 IEEE International Electron Devices Meeting (IEDM), 2020, pp. 28–2.
- 25 A. Belmonte, H. Oh, S. Subhechha, N. Rassoul, H. Hody, H. Dekkers, R. Delhougne, L. Ricotti, K. Banerjee and A. Chasin, *et al.*, 2021 IEEE International Electron Devices Meeting (IEDM), 2021, pp. 10–6.
- 26 Y. Su, M. Shi, J. Tang, Y. Li, Y. Du, R. An, J. Li, Y. Li, J. Yao and R. Hu, *et al.*, *IEEE Trans. Electron Devices*, 2024, **71**, 3336–3342.
- 27 M. Shi, Y. Su, J. Tang, Y. Li, Y. Du, R. An, J. Li, Y. Li, J. Yao and R. Hu, *et al.*, 2023 International Electron Devices Meeting (IEDM), 2023, pp. 1–4.
- 28 S. Park, S. Seong, G. Jeon, W. Ji, K. Noh, S. Kim and Y. Chung, *Adv. Electron. Mater.*, 2023, **9**, 2200554.
- 29 S. Ryu, M. Kang, K. Cho and S. Kim, *Adv. Mater. Technol.*, 2024, **9**, 2302209.
- 30 M. Oota, R. Hodo, T. Ikeda, S. Yamazaki, Y. Ando, K. Tsuda, T. Koshida, S. Oshita, A. Suzuki, K. Fukushima, S. Nagatsuka and T. Onuki, 2019 IEEE International Electron Devices Meeting (IEDM), 2019, 3.2.1–3.2.4.
- 31 J. Won, J. Kang, S. Hong, N. Han, M. Kang, Y. Park, Y. Roh, H. J. Seo, C. Joe and U. Cho, *et al.*, *Adv. Sci.*, 2023, **10**, 2303018.
- 32 Z. Wang, N. Lu, J. Wang, D. Geng, L. Wang and G. Yang, *Materials*, 2023, **16**, 2282.
- 33 D. Saito, J. Doevenspeck, S. Cosemans, H. Oh, M. Perumkunnil, I. Papistas, A. Belmonte, N. Rassoul, R. Delhougne and G. Kar, *et al.*, *IEEE Trans. Electron Devices*, 2020, **67**, 4616–4620.
- 34 K. Nomura, H. Ohta, A. Takagi, T. Kamiya, M. Hirano and H. Hosono, *Nature*, 2004, **432**, 488–492.
- 35 Y. Sekine, K. Furutani, Y. Shionoiri, K. Kato, J. Koyama and S. Yamazaki, *ECS Trans.*, 2011, **37**, 77.
- 36 Y. Zhu, Y. He, S. Jiang, L. Zhu, C. Chen and Q. Wan, *J. Semicond.*, 2021, **42**, 031101.
- 37 J. S. Park, W.-J. Maeng, H.-S. Kim and J.-S. Park, *Thin Solid Films*, 2012, **520**, 1679–1693.
- 38 T. Kamiya, K. Nomura and H. Hosono, *Sci. Technol. Adv. Mater.*, 2010, **11**, 044305.
- 39 M.-C. Yu, D.-B. Ruan, P.-T. Liu, T.-C. Chien, Y.-C. Chiu, K.-J. Gan and S. M. Sze, *IEEE Trans. Nanotechnol.*, 2020, **19**, 481–485.
- 40 K.-M. Lee, B.-K. Ju and S.-H. Choi, *IEEE Trans. Electron Devices*, 2022, **70**, 127–134.
- 41 L.-Y. Su, H.-Y. Lin, H.-K. Lin, S.-L. Wang, L.-H. Peng and J. Huang, *IEEE Electron Device Lett.*, 2011, **32**, 1245–1247.
- 42 J. Wang, J. Bi, G. Xu and M. Liu, *Electronics*, 2024, **13**, 1427.
- 43 Y.-G. Kim and C.-E. Oh, *et al.*, *J. Semicond. Technol. Sci.*, 2024, **24**, 379–387.
- 44 J. Y. Lin, *et al.*, *IEEE Trans. Electron Devices*, 2023, **70**, 1234–1240.
- 45 G. Xu, L. Cai, Z. Wang, Q. Wu, C. Lu, Z. Zhao, Y. Zhao, D. Geng, L. Li and M. Liu, *et al.*, *IEEE Trans. Electron Devices*, 2019, **66**, 5166–5169.
- 46 Z. Pan, Y. Hu, J. Chen, F. Wang, Y. Jeong, D. P. Pham and J. Yi, *Trans. Electr. Electron. Mater.*, 2024, **25**, 371–379.
- 47 J.-S. Park, J. K. Jeong, Y.-G. Mo and S. Kim, *Appl. Phys. Lett.*, 2009, **94**, 042105.
- 48 L. S. Salomone, J. Lipovetzky, S. H. Carbonetto, M. G. Inza, E. G. Redin, F. Campabadal and A. Faigón, *Thin Solid Films*, 2016, **600**, 36–42.
- 49 J. Seok, J. E. Seo, D. K. Lee, J. Y. Kwak and J. Chang, *ACS Nano*, 2025, **19**, 2458–2467.
- 50 K. Lee and H. Shin, *IEEE Trans. Device Mater. Reliab.*, 2017, **17**, 758–762.
- 51 A. Belmonte, S. Kundu, S. Subhechha, A. Chasin, N. Rassoul, H. Dekkers, H. Puliyalil, F. Seidel, P. Carolan and R. Delhougne, *et al.*, 2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits), 2023, pp. 1–2.
- 52 S. Datta, E. Sarkar, K. Aabrar, S. Deng, J. Shin, A. Raychowdhury, S. Yu and A. Khan, 2024 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits), 2024, pp. 1–2.
- 53 G. Wakimura, Y. Yamauchi, T. Matsuoka and Y. Kamakura, 2014 IEEE International Meeting for Future of Electron Devices, Kansai (IMFEDK), 2014, pp. 1–2.
- 54 T. Gokmen and W. Haensch, *Front. Neurosci.*, 2020, **14**, 103.
- 55 C. Lee, K. Noh, W. Ji, T. Gokmen and S. Kim, *Front. Neurosci.*, 2022, **15**, 767953.
- 56 J. Byun, S. Kim, D. Kim, J. Lee, W. Ji and S. Kim, *Adv. Intell. Syst.*, 2024, 2400543.
- 57 M. Le Gallo, C. Lammie, J. Büchel, F. Carta, O. Fagbohunbe, C. Mackin, H. Tsai, V. Narayanan, A. Sebastian and K. El Maghraoui, *et al.*, *APL Mach. Learn.*, 2023, **1**, 041102.
- 58 M. V. Nair and P. Dudek, 2015 International Joint Conference on Neural Networks (IJCNN), 2015, pp. 1–7.
- 59 W. Haensch, T. Gokmen and R. Puri, *Proc. IEEE*, 2019, **107**, 108–122.

