Nanoscale

PAPER

View Article Online
View Journal



Cite this: DOI: 10.1039/d5nr04850h

# *Operando* thermal behaviour of transistor-integrated memristors and its implications on online and offline learning

Eng Kang Koh, [a,b] Putu Andhita Dananjaya, [a] Young Seon You[b] and Wen Siang Lew *[a]

The influence of *in situ* temperature on the electrical characteristics of memristors is a critical consideration for the reliable deployment of neuromorphic computing systems. This study investigates the *operando* thermal behaviour of a 1T–1R (1 Transistor–1 Resistor) $Ta_2O_5$-based memristive device operated at temperatures up to 150 °C, focusing on real-time read-margin changes rather than long-term retention degradation. The demonstrated device exhibited stable potentiation and depression (P/D) behaviour over 2M consecutive pulses without significant degradation, confirming its robustness for thermal stress analysis. Pulse-based measurements revealed distinct temperature-dependent behaviours of the transistor and memristor components. While the transistor current consistently decreased with increasing temperature due to enhanced carrier scattering, the integrated 1T–1R configuration exhibited a unique response: the high-conductance state (HCS) current decreased from metallic-like filamentary conduction, whereas the low-conductance state (LCS) current increased due to thermally activated hopping conduction. Consequently, the dynamic range of conductance compressed from ~17× at 25 °C to ~5× at 150 °C, highlighting the severity of thermal-induced read margin shrinkage. Neural network simulations demonstrated that this compression results in a degradation in accuracy of more than 6% in simple classification tasks, such as MNIST, and up to 2.7% and 4% degradation in offline training for the MNIST and Fashion-MNIST datasets, respectively, compared to the ideal memristor. A scaling-based compensation model was proposed to restore the effective conductance range, thereby recovering the inference accuracy at elevated temperatures. These findings highlight a universal thermal interaction challenge in 1T–1R RRAM architectures and establish a quantitative framework for evaluating and mitigating its impact on neuromorphic system reliability.

Received 17th November 2025,
Accepted 15th May 2026

DOI: 10.1039/d5nr04850h

rsc.li/nanoscale



## Introduction

In the era of the Internet of Things and edge Artificial Intelligence (AI), memristors have emerged as a promising technology due to their scalability and multi-bit storage capabilities.[1–5] When integrated into a crossbar architecture, these devices hold great potential for accelerating AI workloads.[6–8] However, a significant hurdle in adopting resistive memristors for practical systems lies in their relatively large spatial and temporal variations compared to other memory technologies.[9,10] This variability becomes exacerbated at elevated operational temperatures, which require reliability up to 150 °C.[11] Such thermal stresses undermine the stability and reliability of memristors, which are critical for neuromorphic implementations.[12,13] Choi *et al.* demonstrated that operational temperature stress on $WO_x$-based memristors at elevated temperatures up to 373 K (99.85 °C) resulted in severe degradation of memristor-based neural network (NN) performance, reducing it to near-zero levels.[14]

Despite these challenges, studies examining the *operando* thermal effects of 1T–1R memristors are scarce. Most studies focus solely on temperature stress on the 1R cell during retention tests to evaluate neuromorphic reliability, leaving operational stability at elevated temperatures less explored.[15–19] This is despite most functional memristor-based neuromorphic projects employing a 1T–1R architecture, as seen in projects such as NeuRRAM and DenRAM, which are representative examples of memristor-based neuromorphic hardware.[20,21] This configuration is favoured due to better control of the resistive switching characteristics, improving the overall performance in large-scale implementations. Hence, considering the prevalence of 1T–1R architecture in neuromorphic chips, the lack of systematic studies examining the temperature-dependent interaction between transistor and memristor components can be identified as a notable gap in the current literature.

[a]*School of Physical and Mathematical Sciences, Nanyang Technological University, 637371, Singapore. E-mail: WenSiang@ntu.edu.sg*
[b]*GLOBALFOUNDRIES Singapore Pte Ltd, 60 Woodlands Industrial Park D Street 2, Singapore 738406, Singapore*

To address these gaps, this study systematically investigates the impact of *in situ* thermal stress on the operational characteristics of the 1T–1R memristor devices utilised for neuromorphic computing. By capturing real-time dynamic interactions between the transistor and memristor components during pulsed operation, the present study provides direct insight into how temperature affects the coupled system. This differs from conventional retention tests on other scalable memristor platforms, which examine device degradation over time or device failure.[22,23] This work focuses on how synaptic behaviour (specifically P/D dynamics) changes at elevated ambient temperatures of up to 150 °C. By studying conductance fluctuations at different temperature points, the research reveals opposing trends in higher and lower conductance states. It links these shifts to underlying physical mechanisms, such as metallic conduction and thermally activated hopping. This *operando* approach enables the observation of temperature-induced read-margin compression, offering a practical framework for assessing the thermal reliability of memristor-based neuromorphic systems. More importantly, this study also incorporates these real-world device behaviours into neural network simulations, demonstrating how temperature-dependent variability in conductance directly impacts training and inference accuracy. To address this issue, a temperature-aware conductance rescaling scheme is proposed, which digitally compensates for temperature-induced variations based on real-time feedback from on-chip thermal sensors. By normalising the measured conductance to its equivalent room-temperature value through calibrated scaling coefficients and enforcing a conductance floor, this method effectively restores network accuracy under varying thermal conditions. The findings highlight the importance of thermal-aware design, *in situ* temperature monitoring, and adaptive learning algorithms in ensuring the robust deployment of memristive systems in AI hardware.

## Results and discussion

### Memristor cell and operation

In this investigation, as depicted in Fig. 1a, the applied DC voltage was referenced to the source line (SL), with the bit line (BL) grounded, unless otherwise stated. This study focuses on

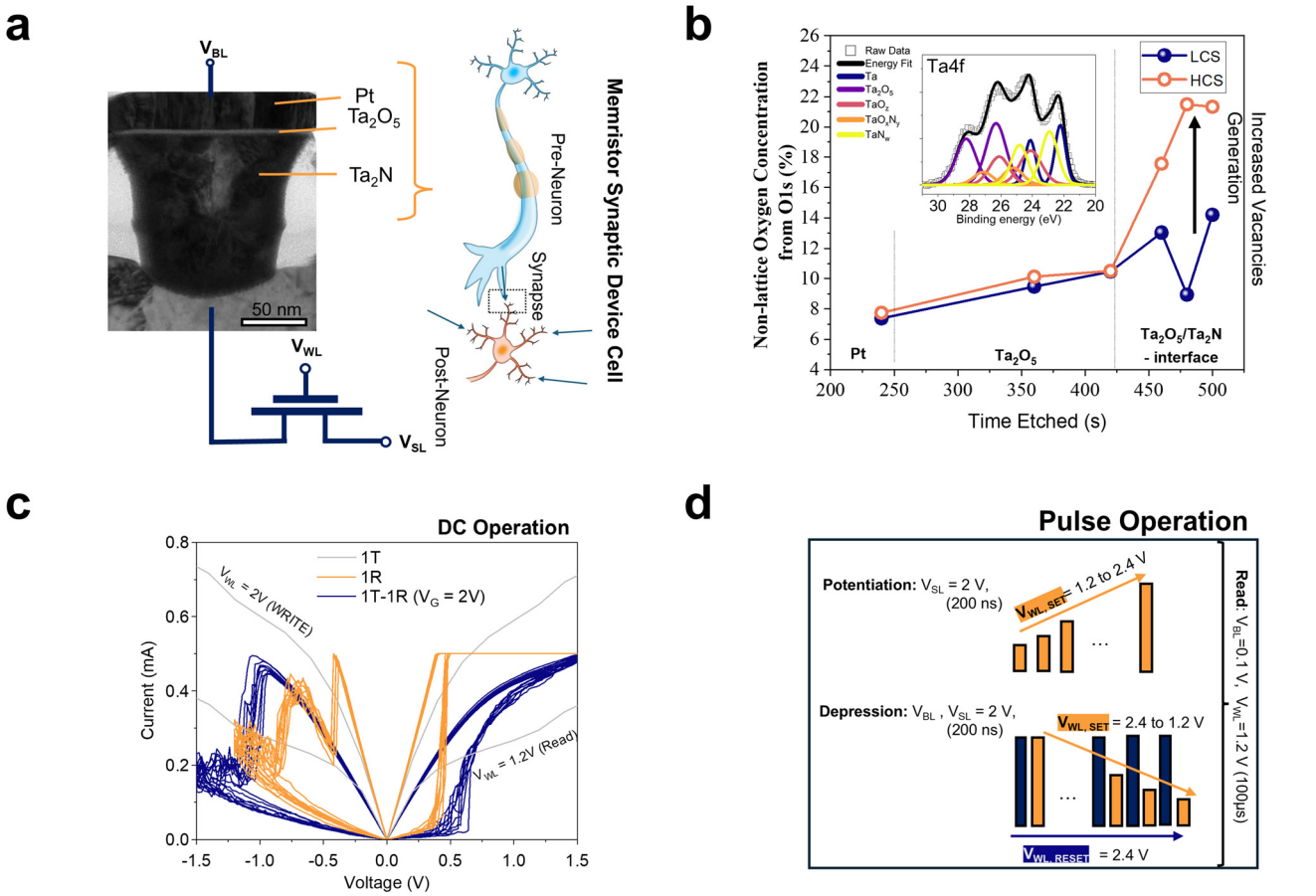


**Fig. 1** (a) TEM imaging of the 1R cell connected to a 40 nm NMOS and the corresponding schematic of the emulated biological synapses. (b) XPS depth profiling performed on devices in the LCS and HCS showing the evolution of the non-lattice oxygen concentration extracted from O 1s peaks as a function of etching time. The inset presents Ta 4f peaks at the approximate middle of the switching layer. (c) DC characteristics of 1T, 1R, and combined 1T–1R cell. (d) Pulse operation scheme deployed for synaptic weight updating.

exploring multilevel switching behaviour of the 1T–1R cell, aiming to emulate synaptic functions under various external temperature conditions. As illustrated in Fig. 1a, the synapse serves as a fundamental component in neuromorphic computing; the memristor stores electrical input from a pre-neuron and relays it to a post-neuron, mimicking biological neural pathways and enabling network-based processing of complex tasks. Fig. 1b presents the XPS depth profiling results obtained from two devices, one in the LCS and another in the HCS. Within the $Ta_2O_5/Ta_2N$ interface, the non-lattice oxygen concentration (as a percentage using the O 1s peak, from 528–535 eV, deconvoluted into two peaks) was found to increase from approximately 12% to 22% upon switching from the LCS to the HCS, indicating the generation of oxygen vacancies during the switching process.[24–28] Additionally, the Ta 4f spectra (from 20–31 eV, deconvoluted into five doublet peaks) depicted in the inset were acquired from approximately the mid-region of the switching layer, revealing the coexistence of multiple Ta oxidation states within the switching film.[24–28] Fig. 1c presents the DC *I–V* characteristics of three configurations: transistor-only (1T – grey), memristor-only (1R – orange), and the combined 1T–1R cell (blue). The cells were swept to 1.5 V for set and ∼−1.25 V to −1.5 V for reset. The differences in current levels before, during, and after switching highlight the essential role of the transistor in modulating current flow and maintaining stable operation in the 1T–1R configuration during weight updates. Fig. 1d illustrates the schematic of the non-identical pulse conditions employed for the emulation of P/D throughout the remaining study. This allows for the evaluation of the memristor performance under realistic operation conditions. This scheme mimics the conductance increase (potentiation) and decrease (depression) analogous to the strengthening and weakening of biological synapses. During potentiation, 128 $V_{WL}$ pulses were incrementally increased from 1.2 to 2.4 V to obtain consecutive states. For depression, a constant complete reset step ($V_{WL}$ = 2.4) was applied, followed by a decremental $V_{WL}$ set condition from 2.4 to 1.2 V. This reset-and-set depression scheme is similar to an existing scheme presented by Li *et al.*[29] The source line ($V_{SL}$) and bit line ($V_{BL}$) were fixed at 2.0 V during both potentiation and depression steps. All write pulses were 200 ns in duration, and read operations were conducted at 0.1 V with a delay of 100 μs after each state transition using a fixed $V_{WL}$ of 1.2 V.

### Potentiation and depression under external temperature stress

Fig. 2a illustrates the P/D behaviour of the memristor at different ambient temperatures. This ambient temperature was applied by a temperature stage on which the sample was placed. Five devices were evaluated at room temperature (RT, ∼23 °C or 296.15 K) with their variations shown as error bars. Additionally, the devices were also evaluated at elevated temperatures of 85 °C (358.15 K), 105 °C (378.15 K), 125 °C (398.15 K), and 150 °C (423.15 K), which correspond to the automotive grades defined by the AEC-Q100 standards.[11] This range represents the highest temperature at which the NMOS transistor in the 1T–1R configuration was qualified to operate reliably. Beyond this limit, the transistor may experience dielectric stress or leakage-related failure, which would obscure the intrinsic thermal behaviour of the memristive element. As the purpose of this work is to investigate *operando* temperature effects on device conductance states rather than extreme-environment qualification or reliability testing, temperatures exceeding 150 °C were not explored. Additionally, the same five devices were subjected to a complete heating–cooling cycle (RT → 85 °C → 105 °C → 125 °C → 150 °C → 125 °C → 105 °C → 85 °C → RT) while applying potentiation and depression pulses. As the external temperature increased, the higher conductance states decreased while the lower conductance states increased, effectively narrowing the dynamic range. This degradation in dynamic range is attributed to the devices' sensitivity to thermal variations, suggesting potential temperature-induced changes in the conduction mechanism. The heating–cooling cycle verified that this effect was intrinsic and reversible, since the conductance dynamic range fully recovered once the devices were cooled back to room temperature. Overlaid results from the five devices confirm that this temperature-dependent behaviour was consistent and repeatable. It is also worth noting that this present work employs a discrete 1T–1R configuration, in which the memristor layers are connected to the post-metal 4 layer, with the probe connected directly to the top electrode. This arrangement enables relatively efficient heat dissipation compared to actual monolithically integrated arrays, which tend to trap more heat within the dielectric and interconnect layers. This would further intensify local temperature rise. Hence, the thermal-induced conductance shrinkage observed in this study represents a baseline scenario, and the effect is anticipated to be even more pronounced in fully integrated systems.

Fig. 2b presents the endurance characteristics (at room temperature) of 1T–1R cells after being subjected to the pulse scheme highlighted in Fig. 1d. Each subfigure in Fig. 2b presents overlapped full P/D cycles from different devices (portrayed with an error bar at each state), corresponding to weight updates at 1, 10k, 100k, 1M, and 2M iterations. The endurance test demonstrated stable conductance ranges even after more than 2 million weight updates, far exceeding the approximately 2500 pulses applied during a single heating–cooling cycle. This indicates that normal operational fatigue does not contribute to the observed temperature effects, supporting the interpretation that the conductance variations are thermally induced and reversible.

### Traversal of states

To further elucidate the observed temperature dependency of the P/D behaviour observed at different temperatures, an analysis was conducted to determine how the underlying conduction mechanisms influenced the states. Our prior studies on an identical structure established that, under low electric fields, the devices exhibit ohmic conduction in both the high-conductance state (HCS) and the low-conductance state (LCS).[30] To understand the observed temperature-dependent behaviour, the current variations across the highest and lowest

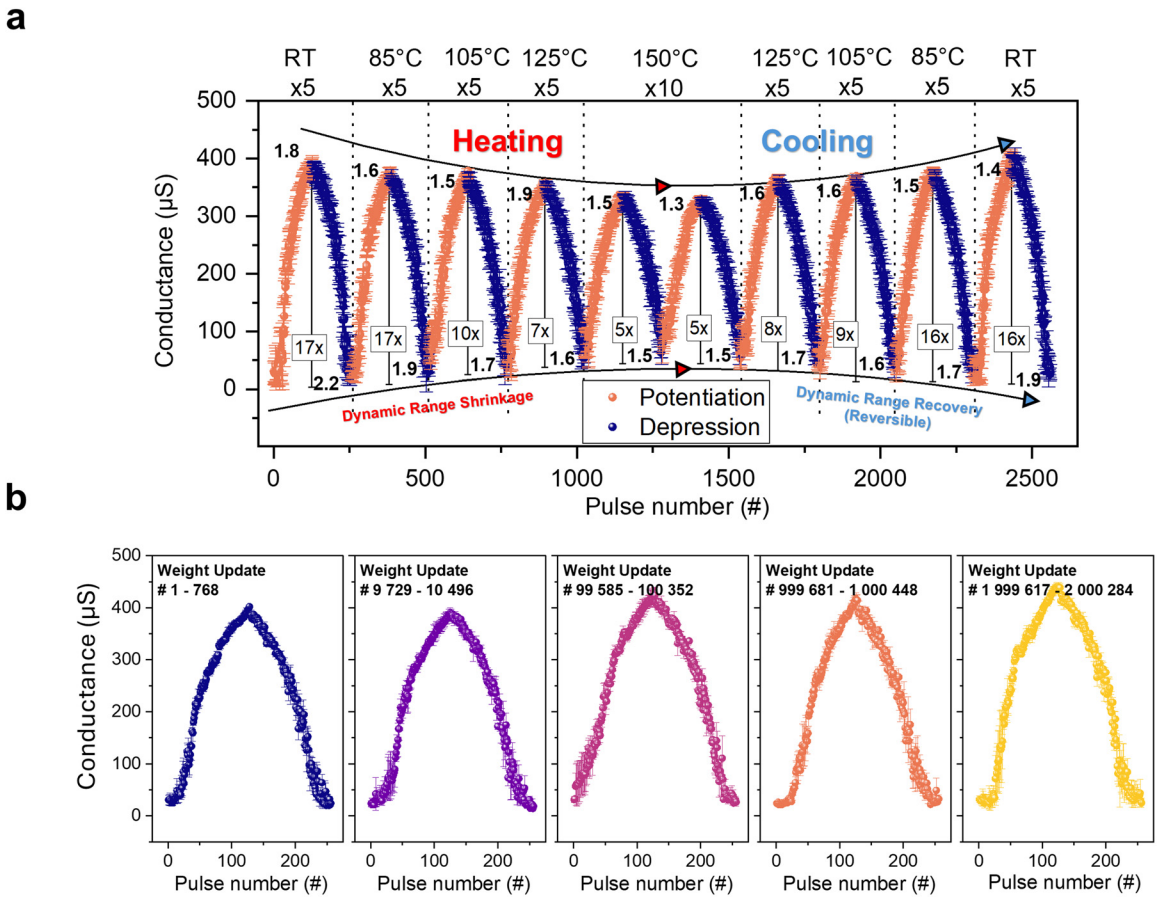


**Fig. 2** (a) Behaviour of five devices and their P/D curves under different external temperature stress conditions. Conductance states show temperature-dependent reversible shifts – while higher conductance states decrease, lower conductance states increase, highlighting the thermally driven conduction effects. (b) Overlay of 3 potentiation and depression cycles up to >2M total weight updates.

conductance states were examined. The devices were switched to their respective states at room temperature before being read at elevated temperatures. Fig. 3a illustrates the behaviour of the 1T–1R cells when subjected to a low-field pulsed sampling double sweep (0.1 V–0.35 V, 100 μs). The double sweep was performed to ensure that any change in read current was solely due to the external temperature influence, rather than an unintended state transition resulting from higher voltage exposure (>0.1 V). The results confirm the observed P/D trends: at the HCS, the current decreases with increasing temperature, while at the LCS, it increases with increasing temperature. This inverse behaviour confirms that the conductance variations are attributable to readout conditions rather than intrinsic shifts in the device states. Furthermore, to isolate the contribution of the transistor, Fig. 3b shows the 1T cell measured separately under the same pulse sampling conditions as those in Fig. 3a, confirming that the drain current decreases as the temperature rises. Fig. 4 depicts a schematic representation of the temperature-dependent conduction behaviour of the 1T, 1R, and 1T–1R devices. Within the NMOS, it is expected that as the temperature increases, the mobility of charge carriers in an NMOS decreases due to enhanced scattering from lattice vibrations (phonons), which in turn causes a decrease in the channel current.[31]

Therefore, although the transistor alone exhibits a decrease in current with rising temperature, the impact of this behaviour is not significant at the HCS since both 1T and 1R experience a similar current reduction with increased external temperature. However, at the LCS, the opposing effects of the 1T and 1R lead to a prevailing increase in current, as the current control is dominated mostly across the memristor. To further explain this opposing trend, the ohmic conduction mechanism as presented in eqn (1) must be examined:[32]

$$J_{\text{ohmic}} = q\mu nE = q\mu N_C E \exp\left(-\frac{E_C - E_F}{kT}\right). \quad (1)$$

Referring back to Fig. 4, in the 1R cell at the HCS, the CF bridges the top and bottom electrodes *via* a continuous path of oxygen vacancies.[33] This continuous path of oxygen vacancies exhibits metallic-like conduction, where the carrier density ($n$) remains nearly constant as the temperature changes. This is similar to metals, where free electrons are

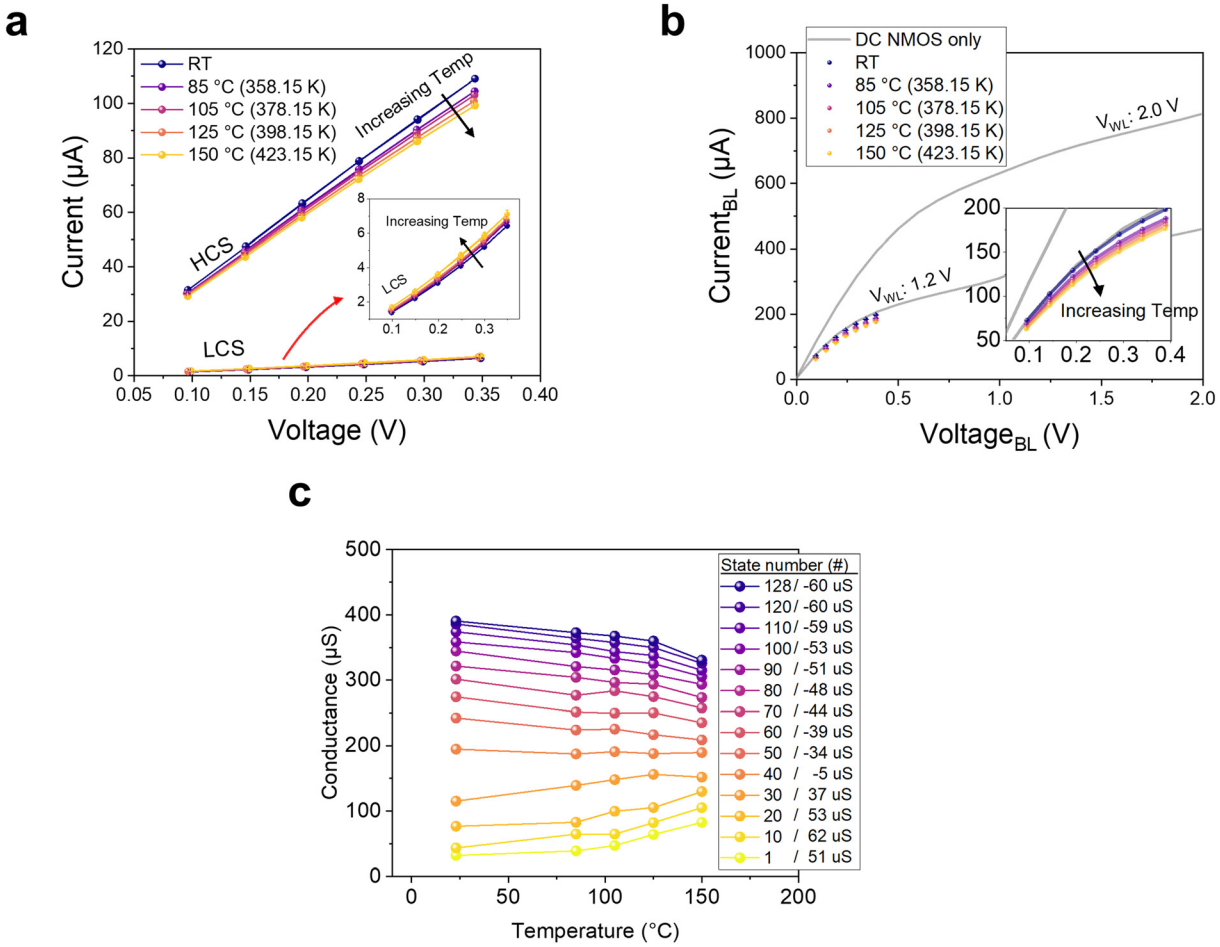


**Fig. 3** Pulse *I*–*V* curves of the 1T–1R cell at various temperatures, read at the (a) highest and lowest conductance under a low field (<0.35 V). (b) The same measurement performed on the 1T cell compared to a DC voltage reference in grey. (c) Analysis of various conductance states as a function of the external temperature, with the conductance shift (150 °C–RT) of the state denoted in the legend.

present at all times and are not generated thermally. Consequently, the temperature dependence of conductivity is primarily governed by carrier mobility ($\mu$). As the temperature increases, enhanced electron–phonon scattering reduces $\mu$, leading to a decrease in current density $J_{\text{ohmic}}$.[34] Conversely, in the LCS, the CF is partially ruptured, introducing a semiconducting gap in the conduction path. In this state, the carrier density follows a thermally activated relationship, $n = N_C E \exp[-(E_C - E_F)/kT]$, resulting in a strong temperature dependence due to thermal excitation of carriers across the localised potential barrier. Although the carrier mobility still decreases with the temperature due to phonon scattering, the rapid increase in carrier concentration outweighs this effect, resulting in an overall increase in current density. Increasing temperature provides sufficient energy for carriers to overcome potential barriers, resulting in a significant increase in carrier concentration. In the combined 1T–1R configuration, the overall temperature dependence arises from the interplay between these two components. In the HCS, both the NMOS and the memristor exhibit reduced current with increasing temperature, resulting in a decrease in the overall current. In contrast, in the LCS, although the NMOS contribution decreases, the thermally enhanced hopping conduction in the memristor dominates, as it is the main current limiter, resulting in an overall increase in current. This competing behaviour explains the asymmetric temperature response observed in the 1T1R device. The opposite trend in the HCS and LCS of the 1T–1R device causes the dynamic range compression observed in Fig. 2a.

Fig. 3c further demonstrates this by presenting every 10 states from Fig. 2a, where the opposing trends in the higher and lower conductance states gradually converge around the 40$^{\text{th}}$ state. This indicates that at intermediate conductance levels, the competing influences of reduced carrier mobility in the HCS and the increased carrier concentration in the LCS balance out, resulting in minimal net change in conductance with the temperature.

### Online and offline training

Hence, the following steps were taken to assess the impact of dynamic range degradation resulting from elevated external temperatures. The P/D characteristics from Fig. 2a, specifically,

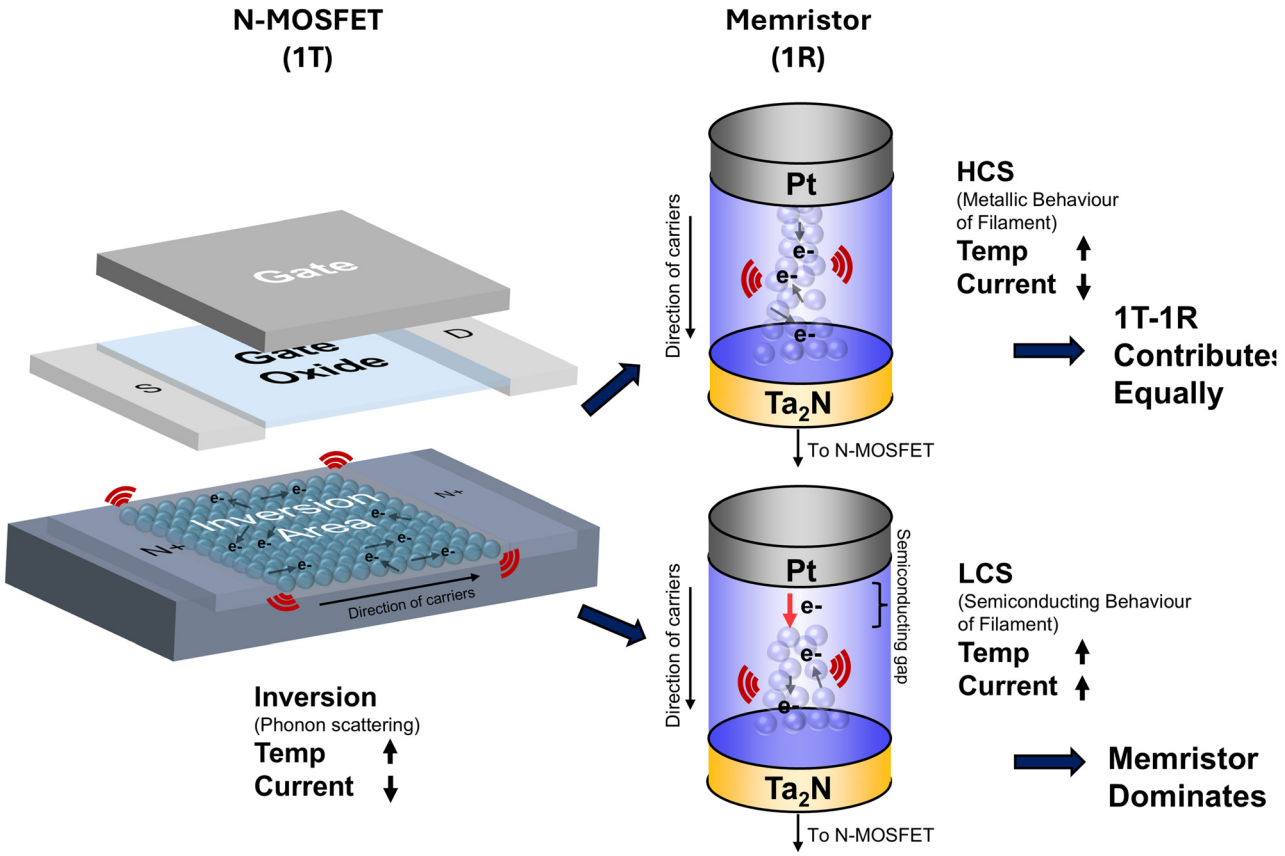


**Fig. 4** Schematic illustrating the temperature-dependent conduction behaviour of the 1T–1R device. As the temperature increases, the current through the NMOS transistor decreases due to enhanced carrier scattering from lattice vibrations. In the memristor, metallic-like filamentary conduction in the HCS is reduced by electron–phonon scattering, while the semiconducting LCS shows an increased current from thermally activated carrier hopping. The schematic highlights the coupled influence of the NMOS and memristor components on the overall dynamic conductance behaviour.

the five experimentally measured devices with each device subjected to the two full P/D cycles, were incorporated into a generic neural network simulation for training on the MNIST dataset. It was assumed that the performance variation, as expected in a larger array, would therefore be constrained within the variability envelope defined by these ten P/D cycles for each temperature. As depicted in Fig. 5a, a fully connected MLP network architecture with 784 input nodes (corresponding to 28 × 28 pixels), a hidden layer of 300 nodes, and 10 output nodes (one for each category) was constructed for this purpose. The sigmoid and SoftMax activation functions were digitally simulated. The classification accuracy dropped by approximately 7% as the operating temperature increased from room temperature (RT) to 150 °C, highlighting the direct influence of temperature-induced device non-idealities on online learning. In this work, only the accuracy degradation was analysed, because the goal of this network architecture was not to maximise the classification performance, but rather to examine how thermally induced changes in the memristor conductance affect the classification performance. The observed degradation is likely driven by the compression of the individual devices' dynamic range, which reduces the separation between conductance levels and consequently limits the precision of weight representation. This effect becomes increasingly significant in deeper and more complex neural network architectures, where a larger number of memristors are required. As each device experiences temperature-induced dynamic range degradation, the cumulative impact across the network exacerbates the overall network accuracy (*i.e.*, the errors from one layer are passed down to the next). For the sake of this argument, effects such as overfitting are assumed to be negligible, since relatively simple datasets like MNIST can be easily overfit using sufficiently large networks. Under such conditions, varying the network depth to compare performance could confound two competing influences. While increasing the number of layers can improve performance through enhanced hierarchical feature extraction, it also requires more memristor devices, which can lead to greater performance degradation, as discussed earlier. This conflict makes it difficult to isolate the individual effects of these two factors.

Next, to more precisely isolate the impact of temperature on conductance-to-state mapping, the MNIST and Fashion-MNIST (a more complex dataset) were trained offline under ideal con-

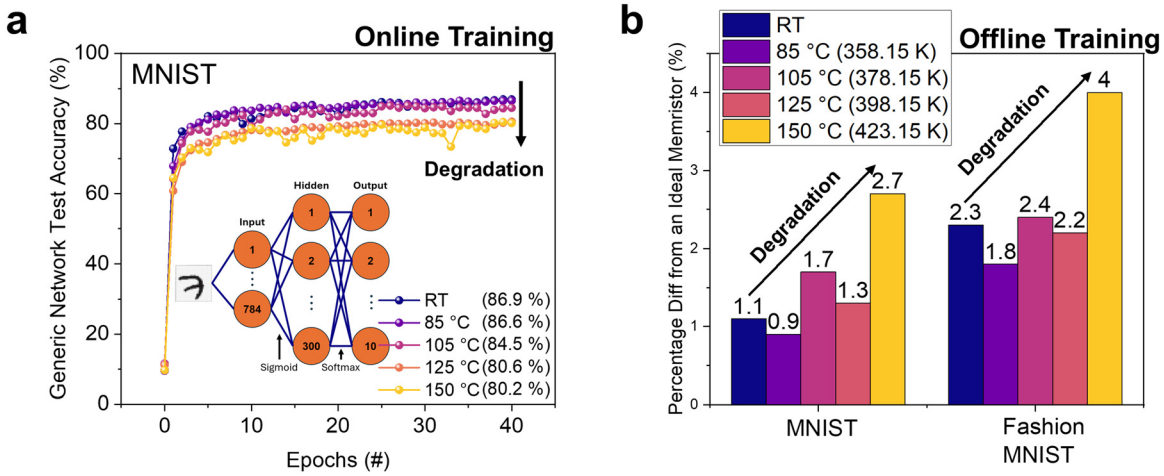


**Fig. 5** (a) Device trained online on the MNIST dataset, showing accuracy degradation as the external temperature increased. (b) Device trained offline on MNIST and Fashion-MNIST datasets also showed the same degradation.

ditions and subsequently mapped to the device. This approach also removes the influence of temperature-induced dynamic range shrinkage on the weight update process, allowing the temperature effect on inference performance to be examined independently. In Fig. 5a, the inference accuracy after online learning decreases nearly linearly with increasing temperature, whereas in Fig. 5b, the offline learning exhibits a nonlinear degradation. This difference arises from the distinct nature of the two training schemes. In online learning, weights are continuously updated through direct interaction with the device, so accuracy degradation closely follows the temperature-induced shrinkage of the device conductance window. In contrast, offline learning involves pre-training under idealised software conditions (infinite precision of states), with subsequent mapping of quantised weights to device conductances. Minor nonlinearities and quantisation artefacts during this mapping can produce small fluctuations in accuracy, sometimes yielding slightly higher performance at intermediate temperatures, such as at 85 °C or 125 °C. Importantly, these local variations do not contradict the device-level behaviour; the overall trend remains a degradation of accuracy with increasing temperature, consistent with the observed conductance shrinkage. This highlights that while the offline learning scheme may show small non-monotonicities, the underlying temperature-dependent reliability challenge persists, emphasising the importance of accounting for these effects in memristor-based neuromorphic systems. Furthermore, the degradation in classification accuracy observed in the offline learning case is notably smaller than that in online training (−1.6% *versus* −6.7%, respectively). This difference arises because, in offline learning, the synaptic weights are pre-trained, and computation primarily involves matrix–vector multiplication. As a result, the memristor's intrinsic switching characteristics, such as the nonlinearity and asymmetry of the potentiation/depression (P/D) curves, do not directly influence the inference process. In contrast, online learning requires continuous *in situ* weight updates, where these device-level nonlinearities play a significant role in determining network performance, leading to a more pronounced accuracy degradation with increasing temperature. To further understand the impact of temperature on network performance, we analysed the $\Delta G - G$ characteristics of our devices across temperatures from room temperature to 150 °C, as shown in Fig. S1 and S2. During learning, the average $\Delta G$ decreases monotonically from 9.2 μS at room temperature to 4.5 μS at 150 °C for potentiation and increases from −5.5 μS at room temperature to −0.8 μS at 150 °C for depression, while the standard deviation decreases more gradually, from 11.5 μS to 8.8 μS (for potentiation) and from 22.7 μS to 18.6 μS (for depression) over the same temperature range, as summarised in Fig. S3, resulting in a modest increase in the CV, driven primarily by the reduced mean update magnitude rather than an absolute increase in stochasticity. Importantly, across all temperatures, no polarity reversal, asymmetric drift, or abrupt broadening of the $\Delta G - G$ distributions is observed, indicating that the intrinsic update mechanism remains fundamentally unchanged. The slight increase in CV leads to local, zero-mean overlap between neighbouring conductance states, causing a minor reduction in update resolution, but does not introduce systematic global bias.

In contrast, the accessible conductance window contracts systematically with the temperature, as shown in Fig. 2, reducing the number of effectively usable states and suppressing weight contrast across the network. This deterministic, global compression of the conductance range is the dominant factor driving the observed degradation in both learning and inference accuracy. The proposed temperature-aware rescaling scheme restores the effective conductance span to its room-temperature equivalent, recovering global weight contrast and separability (w.r.t. Fig. S2 and S3). As a result, both learning and inference performance are restored to near-baseline levels, confirming that dynamic-range shrinkage, rather than altered update statistics or stochastic variability, is the primary mechanism governing temperature-induced accuracy loss.

Thermal crosstalk between neighbouring cells is not explicitly considered in the present simulation. However, the omission of such coupling can be regarded as a base-case scenario, since the inclusion of inter-cell thermal interactions would be expected to further accentuate conductance-range compression through local temperature elevation and spatially correlated drift, as reported in prior studies.[35,36]

Given that conductance-range compression remains a dominant limitation under elevated temperature conditions, it is important to examine how existing compensation strategies address this issue. Temperature compensation techniques for memristive systems can be broadly categorised into device-level, circuit-level, and system-level approaches. Device-level methods, such as material or stack engineering, aim to improve intrinsic thermal stability but require process modifications and provide no runtime adaptability (*i.e.*, effectively O(1) during operation). Circuit-level techniques, including PTAT/CTAT-based biasing, are typically designed under simplified assumptions of monotonic temperature dependence.[37] However, the devices investigated here exhibit asymmetric behaviour, where the HRS conductance increases while the LRS conductance decreases with the temperature, resulting in progressive compression of the conductance window. Programming-based methods such as write–verify compensate variability through iterative pulse–verify cycles, typically requiring O($N$) operations per weight update, where $N$ denotes the number of iterations needed to reach the target state. This leads to increased latency and energy consumption. Nevertheless, while one might assume that write–verify algorithms could ensure accurate conductance programming at elevated temperatures, such schemes only improve precision within the available conductance window and do not address the fundamental reduction of that window itself. When the entire conductance range contracts with the temperature, the achievable separation between programmed states is inherently reduced, and no algorithmic approach can program states beyond the physical limits imposed by the device. Therefore, the problem investigated in this work remains relevant regardless of the use of write–verify schemes, particularly for inference, where weights are fixed and temperature-dependent conductance drift directly translates into computational error without any opportunity for correction.

Hence, a temperature-aware conductance rescaling scheme is proposed in Fig. 6 to digitally compensate for thermally induced variations in device conductance. From Fig. 3c, it can be observed that the effects of the HCS current decrease and the LCS current increase approximately cancel each other around the 40th state, suggesting that a function can be designed to rescale these opposing effects and restore the effective conductance window. This approach leverages a calibrated temperature-dependent model that maps the measured conductance at elevated temperatures back to its equivalent room-temperature value, expressed as:

$$G_{\mathrm{RT}} = \frac{G(T) - \beta_0 - \beta_1(T - T_0)}{\alpha_0 + \alpha_1(T - T_0)} \tag{2}$$

where $\alpha_0$, $\alpha_1$, $\beta_0$, and $\beta_1$ are coefficients fitted *via* linear regression ($\alpha_0$ = 1.03, $\alpha_1$ = −0.0026, $\beta_0$ = −6.4, and $\beta_1$ = 0.5) with $T_0$ = 20 °C, yielding an RMSE of approximately 8.6 μS. This RMSE value refers to the accuracy with which the compensation restores the conductance distribution. Fig. 6a shows a comparison of the experimentally measured conductance (plotted as open circles) at increasingly elevated temperatures with the temperature-compensated results through the proposed digital rescaling scheme (plotted as filled circles). The compensation function was applied to each averaged data point, effectively restoring the original conductance window that had been compressed at elevated temperatures. Fig. 6b presents the full dataset from five devices, each subjected to two complete P/D cycles (one during heating and another during cooling) after post-processing. Additionally, the corrected data were also subjected to a conductance floor of 30 μS, preventing negative values while maintaining consistency across the temperature range. These processed datasets were subsequently used as input for the neural network simulations. Fig. 6c shows the corresponding *in situ* learning and inference results, once again obtained from simulation. At room temperature, as depicted in Fig. 5b, the baseline achieved an accuracy of 86.9% after 40 training epochs.

Following temperature compensation, the simulated inference accuracy remained stable across the temperature range of 85–150 °C, with a value of 85.8–86.7%, showing no observable degradation. Similarly, the inset in the same figure depicted the same for inference. The accuracy showed no noticeable difference across temperatures for both the MNIST and Fashion-MNIST datasets. This confirms that the proposed rescaling approach effectively mitigates temperature-induced conductance variability, thereby allowing network performance to recover.

The practical implementation of the proposed temperature-aware rescaling scheme introduces minimal overhead, as it primarily relies on lightweight digital post-processing rather than significant hardware modification. On-chip temperature sensors, which are already widely used in modern CMOS systems for thermal monitoring, can be leveraged and deployed at the array or tile level, resulting in negligible area and power overhead.[38–40] For the sake of estimation, the representative values from the literature (*e.g.*, the NeuRRAM chip) suggest that the additional area and power required for temperature sensing are on the order of <1% of the overall system.[41]

The compensation can be expressed as a temperature-dependent affine transformation of the measured conductance and implemented using low-complexity digital arithmetic or lookup-table-based approaches. As this requires only a constant number of operations per read, the associated latency is negligible relative to neural network inference timescales. Overall, the scheme provides an efficient and practical solution for mitigating temperature-induced conductance degradation in large-scale systems.

Nevertheless, although this work proposes a promising temperature-aware rescaling approach, its current implementation represents only a first step toward a comprehensive solution. Addressing the thermal sensitivity of memristive systems will require concerted efforts from the broader research community to develop more robust, scalable, and hardware-efficient compensation mechanisms.

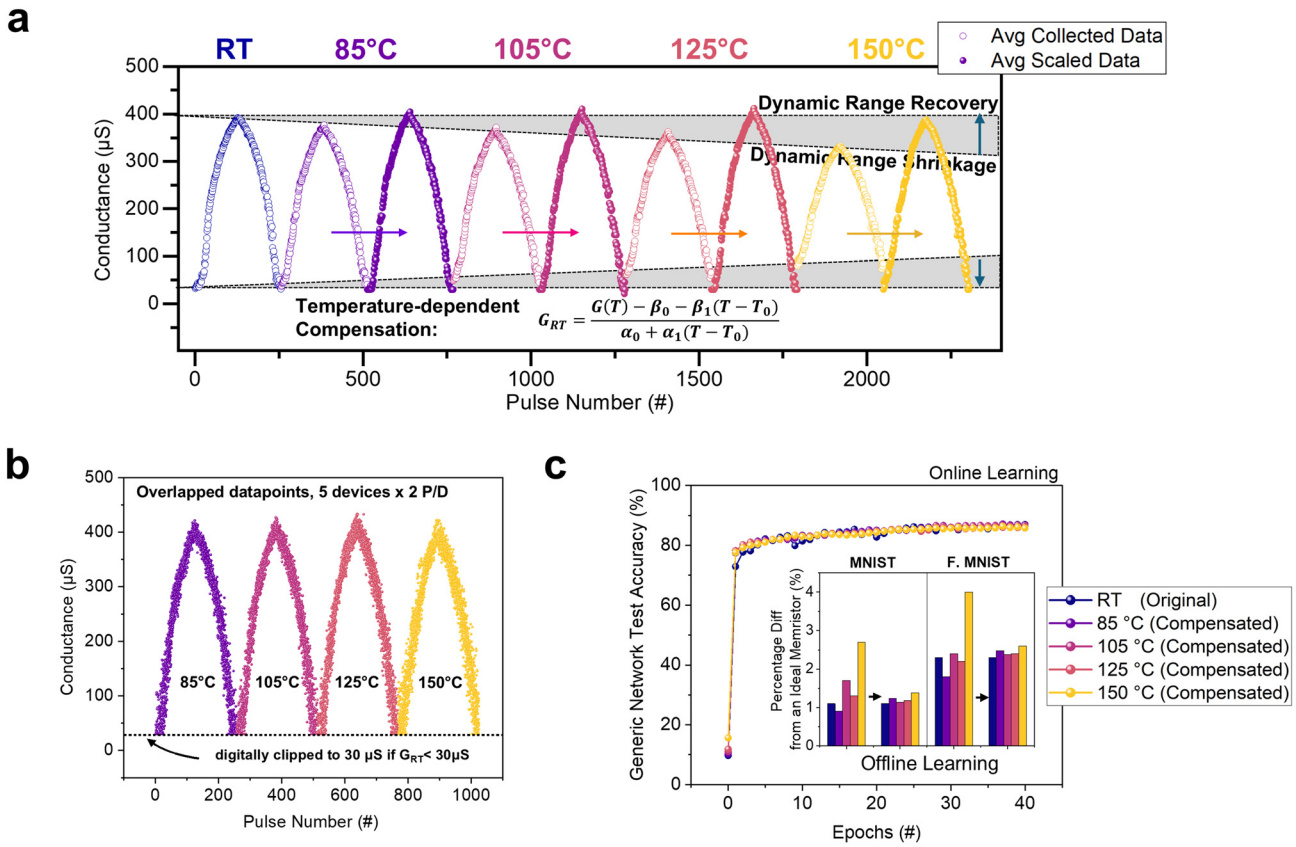


**Fig. 6** (a) Averaged conductance states collected at different temperatures. The open circles were replotted from Fig. 2a, showing the shrunken dynamic range at higher temperatures, while the filled circles represent the corresponding digitally compensated values obtained through the temperature-dependent rescaling function. (b) Aggregated conductance data from five devices, each undergoing two full P/D cycles, after digital post-processing and conductance floor clipping; these values were used as input for network simulations. (c) Neural network online and offline accuracy results demonstrating minimal degradation post-compensation across the elevated temperature compared to room temperature.

## Conclusions

This study highlights the significant influence of temperature on the synaptic weight stability of memristive devices, revealing opposing trends in conductance modulation due to the distinct conduction mechanisms in the HCS and LCS that ultimately reduce the dynamic range. These findings point to the importance of incorporating *in situ* temperature monitoring to manage and mitigate temperature-induced variability, thereby ensuring reliable neuromorphic system performance. It is also important to note that these findings are specific to the memristor structure studied here, which exhibits a distinct ohmic conduction mechanism. Devices based on different conduction mechanisms may exhibit less, or in some cases, even more severe degradation in dynamic range under thermal stress. This work proposed a temperature-aware rescaling scheme. While it effectively restores the conductance range and maintains network accuracy, several limitations remain. Firstly, its performance strongly depends on the precision and spatial resolution of on-chip temperature sensing, where any local hotspots or temporal fluctuations may result in imperfect compensation. Secondly, the model parameters were calibrated using a limited device set and may require re-fitting for other materials or stack configurations, which limits the generality of the approach. Thirdly, the compensation assumes a linear temperature dependence of conductance variation, which may not hold under wider thermal or stress conditions. Moreover, hardware implementation would require additional circuitry for real-time temperature feedback and digital correction, introducing potential power and area overhead. Finally, while the proposed approach corrects systematic temperature-induced drift, it does not mitigate random stochastic variability. However, this can be addressed by applying a write–verify scheme, which was beyond the scope of this work. Future research should focus on integrating similar innovative compensation mechanisms to further enhance the robustness and reliability of memristive synapses in practical neuromorphic applications.

## Methodology

As depicted in Fig. 1a, each memristor structure utilised in this work consists of a Pt top electrode and a $Ta_2O_5$ switching

layer deposited on $Ta_2N$ *via* connecting to the NMOS transistor fabricated using a 40 nm front-end-of-line process. A 3 nm $Ta_2O_5$ layer was deposited by magnetron sputtering using a 99.99% purity $Ta_2O_5$ ceramic target, with a 50 W RF power supply at 2 mTorr pressure and 20 sccm Ar gas flow. The 30 nm Pt electrode was deposited above the switching layer by a 50 W DC PVD step at the same pressure and gas flow. Transmission electron microscopy was carried out using a Thermo Fisher Talos F200X G2 system. The XPS depth profiling was conducted with a monochromatic Al Kα X-ray source using a Kratos Analytical Axis Supra Plus XPS system. All electrical current–voltage (*I*–*V*) measurements for the 1T, 1R, and 1T–1R devices were conducted using a Keithley 4200A-SCS semiconductor parameter analyser. All neural network simulations were performed using the CrossSim 2.0 framework.[42]

## Author contributions

E. K. K. and P. A. D. conceptualised this work and contributed to the formal analysis of the results. E. K. K. contributed to most of the work, including research methodology, data curation, formal investigation, and data visualisation. Y. S. Y. and W. S. L. supervised the research. W. S. L. acquired the funds for this research work. E. K. K. drafted the manuscript. All authors reviewed and revised the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Data availability

The data supporting the findings of this study are available within the article.

Supplementary information (SI) is available. See DOI: https://doi.org/10.1039/d5nr04850h.

## Acknowledgements

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).

## References

1 T.-Y. Wang, J.-L. Meng, L. Chen, H. Zhu, Q.-Q. Sun, S.-J. Ding, W.-Z. Bao and D. W. Zhang, *InfoMat*, 2021, **3**, 212–221.
2 Y. Shen, Z. Pan, M. Jin, J. Gao, Y. Sun, H. Tian and T.-L. Ren, *IEEE Trans. Electron Devices*, 2025, **72**, 266–270.
3 K.-C. Lin, H. Zuo, H.-Y. Wang, Y.-P. Huang, C.-H. Wu, Y.-C. Guo, S.-J. Jou, T.-H. Hou and T.-S. Chang, in *2024 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2024, pp. 1–6.
4 E. K. Koh, P. A. Dananjaya, L. Liu, C. X. X. Lee, G. J. Lim, Y. S. You and W. S. Lew, *ACS Nano*, 2024, **18**, 29602–29617.
5 K. Hou, S. Chen, R. A. John, Q. He, Z. Zhou, N. Mathews, W. S. Lew and W. L. Leong, *Adv. Sci.*, 2024, **11**, 2405902.
6 F. Nowshin and Y. Yi, in *2022 23rd International Symposium on Quality Electronic Design (ISQED)*, 2022, pp. 1–6.
7 T. Shi, L. Gao, Y. Tian, S. Tang, J. Liu, Y. Li, R. Zhou, S. Cui, H. Zhang, Y. Li, Z. Wu, X. Zhang, T. Li, X. Yan and Q. Liu, *Nat. Commun.*, 2025, **16**, 913.
8 L. Liu, P. A. Dananjaya, C. C. I. Ang, E. K. Koh, G. J. Lim, H. Y. Poh, M. Y. Chee, C. X. X. Lee and W. S. Lew, *Nanoscale*, 2023, **15**, 17076–17084.
9 J. B. Roldán, E. Miranda, D. Maldonado, A. N. Mikhaylov, N. V. Agudov, A. A. Dubkov, M. N. Koryazhkina, M. B. González, M. A. Villena, S. Poblador, M. Saludes-Tapia, R. Picos, F. Jiménez-Molinos, S. G. Stavrinides, E. Salvador, F. J. Alonso, F. Campabadal, B. Spagnolo, M. Lanza and L. O. Chua, *Adv. Intell. Syst.*, 2023, **5**, 2200338.
10 G. J. Lim and W. S. Lew, *Nat. Rev. Electr. Eng.*, 2024, **1**, 692–693.
11 R. Hulka Jr., B. Knoell, K. Kirby Jr., R. Haberl, H. Mehrooz, J. Larson, T. Mitchell and G. Fisher, Stress Test Qualification for Passive Components, AEC-Q200 Rev D, AEC Publication, 2010.
12 S. Aldana and H. Zhang, *ACS Omega*, 2023, **8**, 27543–27552.
13 W. Chen, L. Song, S. Wang, Z. Zhang, G. Wang, G. Hu and S. Gao, *Adv. Electron. Mater.*, 2023, **9**, 2200833.
14 W. Choi, C. Lee, S. Noh, J. Lee, H. Lee, S. Kim and H. Hwang, *IEEE Electron Device Lett.*, 2021, **42**, 763–766.
15 J. Hazra, M. Liehr, K. Beckmann, M. Abedin, S. Rafq and N. Cady, in *2021 IEEE International Memory Workshop (IMW)*, 2021, pp. 1–4.
16 W. Shim, J. Meng, X. Peng, J.-s. Seo and S. Yu, in *2021 IEEE International Reliability Physics Symposium (IRPS)*, 2021, pp. 1–4.
17 J. Zhang, X. Ma, Y. Xi, Y. Lu, K. Wang, H. Ren, J. Tang, L. Pan, L. Chen, D. Wu, B. Gao, H. Qian and H. Wu, in *2024 IEEE International Electron Devices Meeting (IEDM)*, 2024, pp. 1–4.
18 L. Cai, W. Chen, Y. Zhao, X. Liu, J. Kang, X. Zhang and P. Huang, *IEEE Trans. Electron Devices*, 2019, **66**, 3822–3827.
19 C. Nail, G. Molas, P. Blaise, G. Piccolboni, B. Sklenard, C. Cagli, M. Bernard, A. Roule, M. Azzaz, E. Vianello, C. Carabasse, R. Berthier, D. Cooper, C. Pelissier, T. Magis, G. Ghibaudo, C. Vallée, D. Bedeau, O. Mosendz, B. De Salvo and L. Perniola, in *2016 IEEE International Electron Devices Meeting (IEDM)*, 2016, p. 4.5.1–4.5.4.
20 W. Wan, R. Kubendran, C. Schaefer, S. B. Eryilmaz, W. Zhang, D. Wu, S. Deiss, P. Raina, H. Qian, B. Gao, S. Joshi, H. Wu, H.-S. P. Wong and G. Cauwenberghs, *Nature*, 2022, **608**, 504–512.
21 S. D'Agostino, F. Moro, T. Torchet, Y. Demirağ, L. Grenouillet, N. Castellani, G. Indiveri, E. Vianello and M. Payvand, *Nat. Commun.*, 2024, **15**, 3446.

 

22 E. Wu, Y. Wang, S. Huo, J. Xu, M. Sheng, H. Liu, L. Zhong, J. Gao, Y. Xie and C. Pan, *Adv. Funct. Mater.*, 2025, e18764.
23 Q. Wang, R. Luo, Y. Wang, W. Fang, L. Jiang, Y. Liu, R. Wang, L. Dai, J. Zhao, J. Bi, Z. Liu, L. Zhao, Z. Jiang, Z. Song, J. Schwarzkopf, T. Schroeder, S. Wu, Z.-G. Ye, W. Ren, S. Song and G. Niu, *Adv. Funct. Mater.*, 2023, **33**, 2213296.
24 M. Ismail, H. Abbas, A. Sokolov, C. Mahata, C. Choi and S. Kim, *Ceram. Int.*, 2021, **47**, 30764–30776.
25 K. X. Shi, H. Y. Xu, Z. Q. Wang, X. N. Zhao, W. Z. Liu, J. G. Ma and Y. C. Liu, *Appl. Phys. Lett.*, 2017, **111**, 223505.
26 R. Simpson, R. G. White, J. F. Watts and M. A. Baker, *Appl. Surf. Sci.*, 2017, **405**, 79–87.
27 S. Khan, M. J. M. Zapata, M. B. Pereira, R. V. Gonçalves, L. Strizik, J. Dupont, M. J. L. Santos and S. R. Teixeira, *Phys. Chem. Chem. Phys.*, 2015, **17**, 23952–23962.
28 D. Cristea, L. Cunha, C. Gabor, I. Ghiuta, C. Croitoru, A. Marin, L. Velicu, A. Besleaga and B. Vasile, *Nanomaterials*, 2019, **9**, 476.
29 C. Li, D. Belkin, Y. Li, P. Yan, M. Hu, N. Ge, H. Jiang, E. Montgomery, P. Lin, Z. Wang, W. Song, J. P. Strachan, M. Barnell, Q. Wu, R. S. Williams, J. J. Yang and Q. Xia, *Nat. Commun.*, 2018, **9**, 2385.
30 E. K. Koh, P. A. Dananjaya, H. Y. Poh, L. Liu, C. X. X. Lee, J. R. Thong, Y. S. You and W. S. Lew, *Nanoscale Horiz.*, 2024, **9**(5), 828–842.
31 G. Gaikwad, M. Sasmal and S. Lande, *Int. J. Sci. Eng.*, 2018, **3**, 15–19.
32 E. W. Lim and R. Ismail, *Electronics*, 2015, **4**, 586–613.
33 P. Huang, X. Y. Liu, W. H. Li, Y. X. Deng, B. Chen, Y. Lu, B. Gao, L. Zeng, K. L. Wei, G. Du, X. Zhang and J. F. Kang, *Int. Electron Devices Meet.*, 2012, 26.6.1–26.6.4.
34 C. M. M. Rosário, B. Thöner, A. Schönhals, S. Menzel, A. Meledin, N. P. Barradas, E. Alves, J. Mayer, M. Wuttig, R. Waser, N. A. Sobolev and D. J. Wouters, *Nanoscale*, 2019, **11**, 16978–16990.
35 P. Sun, N. Lu, L. Li, Y. Li, H. Wang, H. Lv, Q. Liu, S. Long, S. Liu and M. Liu, *Sci. Rep.*, 2015, **5**, 13504.
36 D. Schön and S. Menzel, in *2024 IEEE International Memory Workshop (IMW)*, 2024, pp. 1–4.
37 B.-D. Yang, Y.-K. Shin, J.-S. Lee, Y.-K. Lee and K.-C. Ryu, in *2009 IEEE Asian Solid-State Circuits Conference*, 2009, pp. 241–244.
38 A. L. Aita, M. A. P. Pertijs, K. A. A. Makinwa, J. H. Huijsing and G. C. M. Meijer, *IEEE Sens. J.*, 2013, **13**, 1840–1848.
39 T. Meng and C. Xu, *J. Semicond.*, 2009, **30**, 045002.
40 S. Park, C. Min and S. Cho, in *2009 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2009, pp. 1153–1156.
41 W. Wan, R. Kubendran, C. Schaefer, S. B. Eryilmaz, W. Zhang, D. Wu, S. Deiss, P. Raina, H. Qian, B. Gao, S. Joshi, H. Wu, H.-S. P. Wong and G. Cauwenberghs, *Nature*, 2022, **608**, 504–512.
42 T. P. Xiao, C. H. Bennett, B. Feinberg, M. J. Marinella and S. Agarwal, CrossSim 2.0: accuracy simulation of analog in-memory computing (version 2) https://github.com/sandia-labs/cross-sim.