Linking Device Dynamics to Neural Network Performance in Ionically Gated Synaptic Transistors

Nithil Harris Manimaran
Rochester Institute of Technology
Huayuan Han
Rochester Institute of Technology
Cory Merkel
Rochester Institute of Technology
Ke Xu
ke.xu@rit.edu

Rochester Institute of Technology

Article

Keywords: Ionically Gated Transistors, Electric Double Layer, Neuromorphic Computing, Artificial Neural Network, Weight Update Nonlinearity, Conductance States

Posted Date: April 21st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9248715/v1

Additional Declarations: No competing interests reported.

Linking Device Dynamics to Neural Network Performance in Ionically Gated Synaptic Transistors

*Nithil Harris Manimaran[1], Huayuan Han[2], Cory Merkel[3], and Ke Xu[1,2,4]**

[1]Microsystems Engineering, [2]School of Physics and Astronomy, [3]Department of Electrical and Computer Engineering, [4]School of Chemistry and Material Science, Rochester Institute of Technology, Rochester, New York 14623, United States

**Correspondence and requests for materials should be addressed to K.X. (email: ke.xu@rit.edu)*

Abstract

Neuromorphic computing based on artificial synapses requires devices capable of gradual, repeatable, and energy-efficient conductance modulation. Ionically gated transistors are promising candidates because their ion dynamics naturally produce synaptic behavior under low-voltage operation. However, how key device characteristics—conductance range, number of accessible conductance states N, and weight-update nonlinearity β—jointly influence neural network performance remains insufficiently understood. Here, we investigate MoS_2-based ionically gated synaptic transistors using a combined experimental and modeling framework that links device physics to hardware-aware artificial neural network (ANN) simulations across image-classification tasks of varying complexity. We show that under fixed-amplitude pulsing, increasing N introduces a fundamental trade-off: finer weight resolution is accompanied by stronger update nonlinearity. ANN simulations further reveal that, within the nonlinearity range studied here, classification accuracy is governed by a task-dependent optimal weight resolution rather than a simply maximized number of states, and is relatively insensitive to conductance range when N is fixed. To overcome the nonlinear weight updates, we employ a physics-informed transient model to develop a predictive pulse-engineering algorithm and experimentally demonstrate that it can linearize synaptic weight evolution in the same device. These linearized updates improve ANN accuracy by up to 7% at equivalent state resolution, establishing a quantitative link between device-level dynamics and neural network performance in ionically gated synaptic transistors.

Keywords

Ionically Gated Transistors, Electric Double Layer, Neuromorphic Computing, Artificial Neural Network, Weight Update Nonlinearity, Conductance States

Introduction

The rapid growth of artificial intelligence (AI) applications has intensified the need for computing hardware that can process large volumes of data with high energy-efficiency.[1] Conventional von Neumann architectures, in which memory and computation are physically separated, incur substantial energy and latency costs associated with frequent data transfer between these units.[2] Neuromorphic computing offers an alternative approach by co-locating memory and computation within networks of artificial neurons and synapses, thereby enabling highly parallel and potentially more energy-efficient information processing.[3,4] Realizing such systems in hardware requires electronic synaptic devices that can store and update neural network weights through gradual and repeatable conductance modulation.[5]

A wide range of device platforms have been explored as artificial synapses, including phase-change memory,[6,7] resistive memory,[8,9] ferroelectric field effect transistors (FeEFTs),[10,11] electrochemical RAM,[12,13] and electrostatically gated ionic devices such as electric double layer transistors (EDLTs).[14–16] Many of these technologies are primarily investigated for nonvolatile weight storage, whereas EDLTs operate through ionic motion within an electrolyte that naturally produces volatile and history-dependent conductance modulation under pulsed electrical inputs.[5] In an EDLT, mobile ions redistribute under an applied electric field and form electric double layers at the gate/electrolyte and electrolyte/channel interfaces, which electrostatically modulate the channel conductance. The large interfacial capacitance associated with EDL gating enables substantial carrier accumulation at relatively low operating voltages, while the coupled formation and relaxation of the ionic distribution produces dynamic synaptic behaviors such as potentiation and depression.[17–20] The resulting volatile conductance dynamics can emulate short-term synaptic plasticity, a key feature of biological synapses that enables temporal signal processing in neuromorphic systems.[21,22] EDL gating is particularly effective when combined with two-dimensional (2D) semiconductors, where modest gate voltages can induce large interfacial carrier densities (10^{13}–10^{14} cm^{-2}) and strong conductance modulation.[23–25]

For neuromorphic learning applications, several device-level characteristics are commonly considered important, including conductance range, number of accessible conductance states, and the linearity and symmetry of potentiation and depression. Increasing the number of conductance states is often viewed as beneficial because it improves weight resolution, while more linear and symmetric updates are generally expected to improve training accuracy.[4,26–29] However, in ionic devices these metrics are strongly coupled. The cumulative nature of ion migration and relaxation tends to produce nonlinear conductance evolution, in which early pulses induce larger weight changes than later pulses.[14,16,30–36] A larger conductance range

may delay this saturation and enable more distinguishable states, but it remains not well understood how these device metrics jointly influence neural network performance, whether increasing the number of states is always beneficial, and how differences in device characteristics such as conductance range affect learning behavior at the algorithm level. Addressing these questions is essential for establishing practical design guidelines for ionically gated synaptic hardware.

In this work, we investigate how three key EDLT synaptic characteristics—conductance range (on/off ratio), number of accessible conductance states N, and weight-update nonlinearity β—jointly influence ANN performance across tasks of different complexity. We experimentally study MoS_2-based EDLT synaptic transistors and employ a physics-informed model of EDLT dynamics to capture the coupled ionic and electronic processes governing weight modulation. Building on this framework, we develop a predictive pulse-engineering algorithm and experimentally demonstrate its ability to linearize synaptic weight updates. Using the experimentally extracted plasticity characteristics, we perform hardware-aware ANN simulations to evaluate how these device properties translate to classification accuracy. Together, this combined experimental and modeling approach establishes a device-to-system-level understanding of how conductance range, weight resolution, and update linearity should be balanced when designing ionically gated synaptic devices for neuromorphic hardware.

Results and Discussion

Fig. 1a illustrates the schematic of an ionically gated MoS_2 EDLT with a side-gate configuration, and an optical image of a representative device is shown in the inset of Fig. 1b. Device fabrication details are described in the Methods section. The transfer characteristics in Fig. 1b confirm effective electrostatic modulation of the MoS_2 channel, exhibiting predominately n-type conduction consistent with electron transport in MoS_2. The device shows an on/off ratio exceeding 10^4 and a threshold voltage of approximately -0.04 V. The large on/off ratio defines the available dynamic range for conductance modulation, while the near-zero threshold voltage enables substantial conductance change under relatively small gate voltages, beneficial for low power operation. This device represents the high on/off ratio device used in this study, while the transfer characteristics of a lower on/off ratio device are shown in Fig. S1.

To examine the transient response, single gate-voltage (V_G) pulses with a duration of ~1.8 s and varying amplitudes were applied, as shown in Fig. 1c. The resulting transient drain current (I_D) exhibits a rapid rise followed by a gradual decay, resembling the excitatory postsynaptic current (EPSC) observed in biological synapses. As the pulse amplitude increases from 0.1 V to 2.5 V, the peak ΔI_D increases from approximately 7 nA to 171 nA before relaxing back toward the baseline current. The rapid current increase arises from drift-driven ion migration toward the

electrolyte/channel interface during the gate pulse, while the subsequent decay reflects diffusion-driven ion relaxation after the pulse is removed.

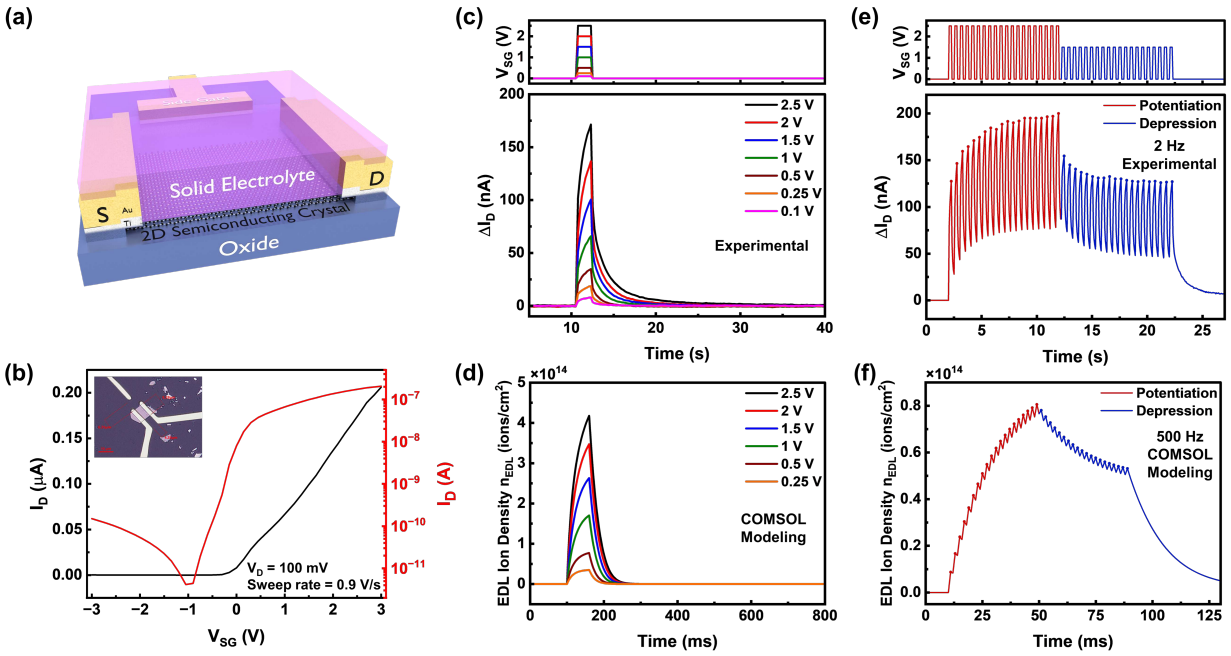

Figure 1: *(a) Schematic of a side-gated EDLT architecture. (b) Transfer characteristics showing predominantly n-type conduction with an on/off ratio exceeding 10^4, inset shows an optical image of a representative device. (c, e) Experimentally measured single and multi-pulse responses demonstrating EPSC-like behavior and multilevel conductance modulation under repeated identical gate-voltage pulses at 2 Hz. (d, f) COMSOL simulations showing EDL ion density evolution, and reproducing the nonlinear and saturating conductance updates observed experimentally at 500 Hz.*

To further understand the ionic dynamics governing this response, the experimental transient responses are compared with finite-element simulations using a COMSOL Multiphysics parallel-plate capacitor model. Details of this model were reported in a previous study.[20] The simulations reproduce both the transient relaxation behavior and the magnitude scaling with V_G (Fig. 1d), indicating that the experimental device response is governed by ionic redistribution within the electrolyte. As V_G increases from 0.25 V to 2.5 V, the simulated electric double layer ion density (n_{EDL}) increases from 0.32 x 10^{14} ions/cm^2 to 3.95 x 10^{14} ions/cm^2. While the experiment measures the transient drain current and the model tracks the EDL ion density, the two quantities are directly related because experimentally the accumulated ionic charge at the electrolyte/channel interface modulates the channel conductance. Consistent with this relationship, both the measured I_D and simulated n_{EDL} increase by approximately one order of magnitude as V_G increases from 0.25 V to 2.5 V, indicating consistent scaling between ionic

accumulation and channel conductance change.

The transient synaptic behavior of the device was further examined using sequences of fixed magnitude voltage pulses (Fig. 1e). Potentiation and depression were induced using 2.5 V and 1.5 V pulses, respectively, with a pulse period of 0.5 s and a duty cycle of 50%. During potentiation, the peak current gradually increases from approximately 127 nA to 200 nA, while depression pulses reduce the peak current back toward ~127 nA, demonstrating multilevel conductance modulation. The peak current following each pulse represents distinct channel conductance states analogous to synaptic weights. Consistent with prior reports, [14,16,30–36] the incremental conductance change decreases with successive pulses, leading to nonlinear weight updates and eventual saturation. While the experimentally accessible pulse frequency is limited by the instrumentation, COMSOL simulations performed at higher frequencies (500 Hz, Fig. 1f) reproduce the same nonlinear and saturating behavior, indicating that the saturation arises from intrinsic ionic dynamics rather than experimental limitations. The simulations further suggest that stronger potentiation can be achieved at higher operating frequencies, although such regimes are not accessible in the present experimental setup.

Having established that repeated identical pulses produce nonlinear conductance evolution, we next examine how this behavior influences the number of accessible synaptic states. In ionically gated transistors, the cumulative ionic response enables analog conductance modulation; however, the effective number of usable states depends on the device dynamic range (on/off ratio), the degree of weight update nonlinearity, and the asymmetry between potentiation and depression. Fig. 2a shows the experimentally extracted normalized synaptic weights for 10, 20, and 30 conductance states obtained from the same high on/off device using the pulsing scheme described in Fig. 1e. The corresponding transient current responses are shown in Fig. S2a–b, from which the peak current after each pulse were extracted and normalized to obtain the weight evolution curves. As the number of states increases, the weight evolution becomes increasingly nonlinear, characterized by a rapid initial increase followed by gradual saturation. Such nonlinear behavior may arise from the saturation of the transistor conductance and ionic accumulation dynamics during repeated pulsing. In this high on/off device, although the peak current during pulsing approaches the upper range of the transfer characteristics shown in Fig. 1b (~200 nA at V_G = 3V), the transfer curve itself does not exhibit clear saturation in this regime. Therefore, the observed nonlinear evolution during pulsing is more consistent with pulse-to-pulse ionic accumulation dynamics, arising from the competition between drift-driven ion migration during each pulse and diffusion-driven ion relaxation between pulses.[20] COMSOL simulations (Fig. 2b) reproduce the same qualitative behavior at a higher pulsing frequency of 500 Hz, indicating that the nonlinear conductance evolution originates primarily from ionic redistribution and persists at faster operation.

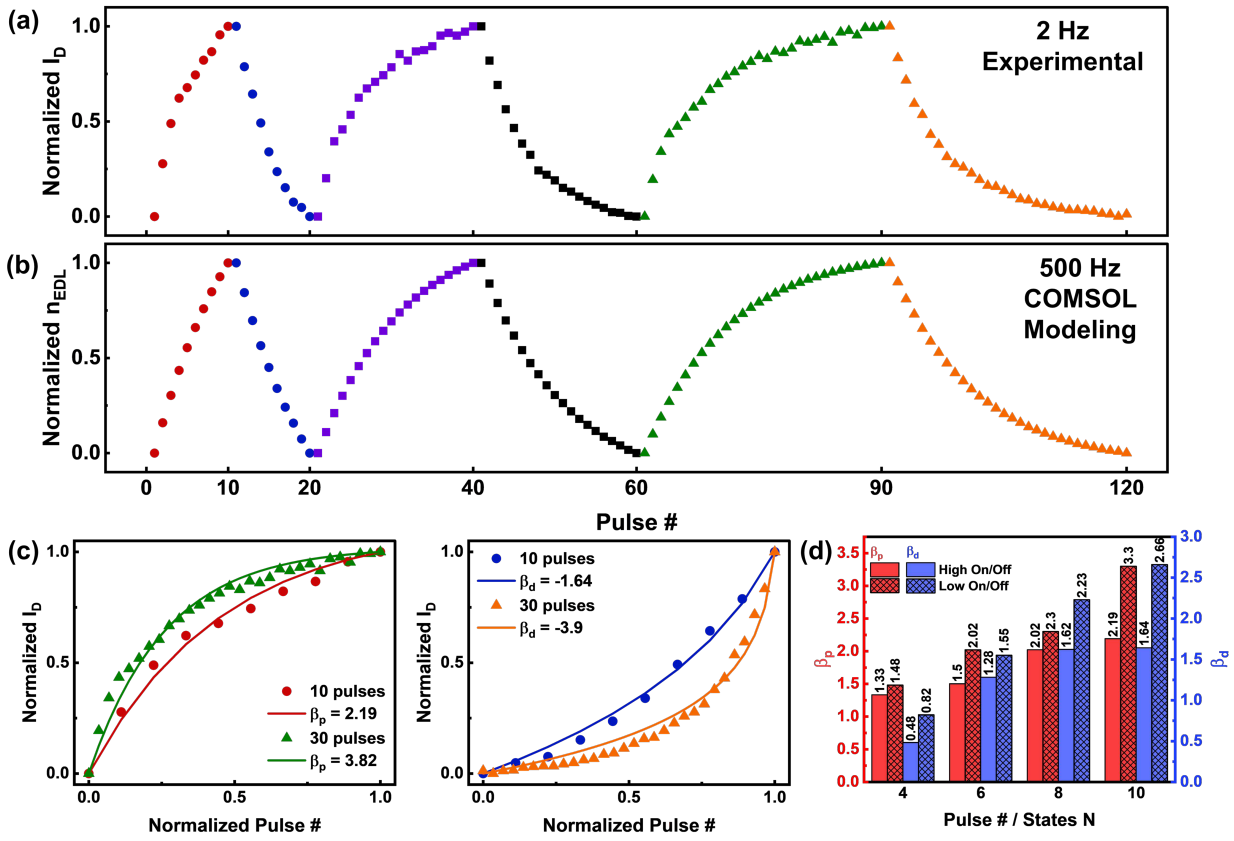

Figure 2: *(a) Experimentally extracted normalized synaptic weight updates for 10, 20, and 30 states from the high on/off device using the 2 Hz pulse scheme in Fig. 1e. (b) COMSOL simulations at 500 Hz reproducing the nonlinear weight evolution arising from ionic redistribution. (c) Extracted asymmetric nonlinearity (ANL) factor β for potentiation and depression as a function of state number N for the high on/off device. (d) Comparison of β values for high and low on/off devices, showing reduced nonlinearity in devices with larger tunable conductance range.*

To quantify the nonlinearity, the asymmetric nonlinearity (ANL) factor β was extracted using Eq. (1)[20] for both potentiation and depression (Fig. 2c).

$$Normalized\ I_D = \begin{cases} \dfrac{1 - exp\ (-\beta_p N)}{1 - exp\ (-\beta_p)} \\ \dfrac{ln\ [1 - N\{1 - exp(-\beta_d)\}]}{-\beta_d} \end{cases} \quad \ldots(1)$$

For this high on/off device, the magnitude of β increases as the number of states increases, with β_p (potentiation) increasing from 2.19 to 3.82 and β_d (depression) increasing from 1.64 to 3.9 when the number of states increases from 10 to 30. Under the fixed-amplitude pulse programming scheme used here, this trend indicates that achieving finer weight resolution through additional conductance states inherently introduces stronger nonlinear weight updates,

highlighting a trade-off between increasing the number of synaptic states and maintaining near-linear conductance modulation.

Fig. 2d compares the extracted β values for both the high and low on/off devices as a function of pulse/state number. For a given number of states N, the high on/off device consistently exhibits smaller β values for both potentiation and depression. For example, at 10 states the low on/off device shows approximately 1.5–1.6× larger β than the high on/off device, while at 30 states the difference increases to nearly 3.3×. Because the same electrolyte and same pulse scheme are used for both devices, this difference reflects primarily the larger conductance range of the high on/off device, which delays conductance saturation and enables more uniform incremental weight updates. In contrast, the limited conductance range of the low on/off device leads to earlier saturation and stronger nonlinear weight evolution. The conductance of the low on/off device saturates within approximately 10 pulses, beyond which additional pulses do not produce distinguishable conductance levels (Fig. S2c–d). These results indicate that two factors jointly govern the effective synaptic states in EDLT devices under fixed magnitude pulsing scheme: ion dynamics introduces an intrinsic trade-off between the number of states and update linearity, while the device conductance range determines how early conductance saturation occurs and therefore how nonlinear the weight updates become for a given number of states.

To examine how the device-level weight update characteristics identified in Fig. 2 influence network-level performance, the experimentally extracted potentiation and depression trajectories were incorporated into artificial neural network (ANN) simulations. Fig. 3a shows the fully connected, three-layer feedforward ANN architecture used in this study, consisting of 784 input neurons, 300 hidden neurons, and 10 output neurons corresponding to the classification classes. A conceptual crossbar-based hardware implementation is illustrated in Fig. 3b, where MoS_2 EDLTs serve as synaptic elements in a neuromorphic computing system. The experimentally measured nonlinear potentiation and depression characteristics were mapped directly to the synaptic weight update rules during ANN training (See Methods). Network performance was evaluated using three widely used image-classification benchmarks: MNIST, Kuzushiji-MNIST (KMNIST), and Fashion-MNIST (FMNIST), which represent progressively increasing task complexity. The validation accuracy was tracked after each training epoch and used to guide an early stopping function; training was terminated when the validation accuracy did not improve by more than 0.5% over 10 consecutive epochs, which balances convergence stability and computational cost. To account for stochastic variations during training, each simulation was repeated five times and the reported results represent the mean accuracy with the corresponding standard deviation. This evaluation framework is consistent with established approaches used in prior ANN studies[14,20,27,37–41] and allows a statistically meaningful comparison for network performance.

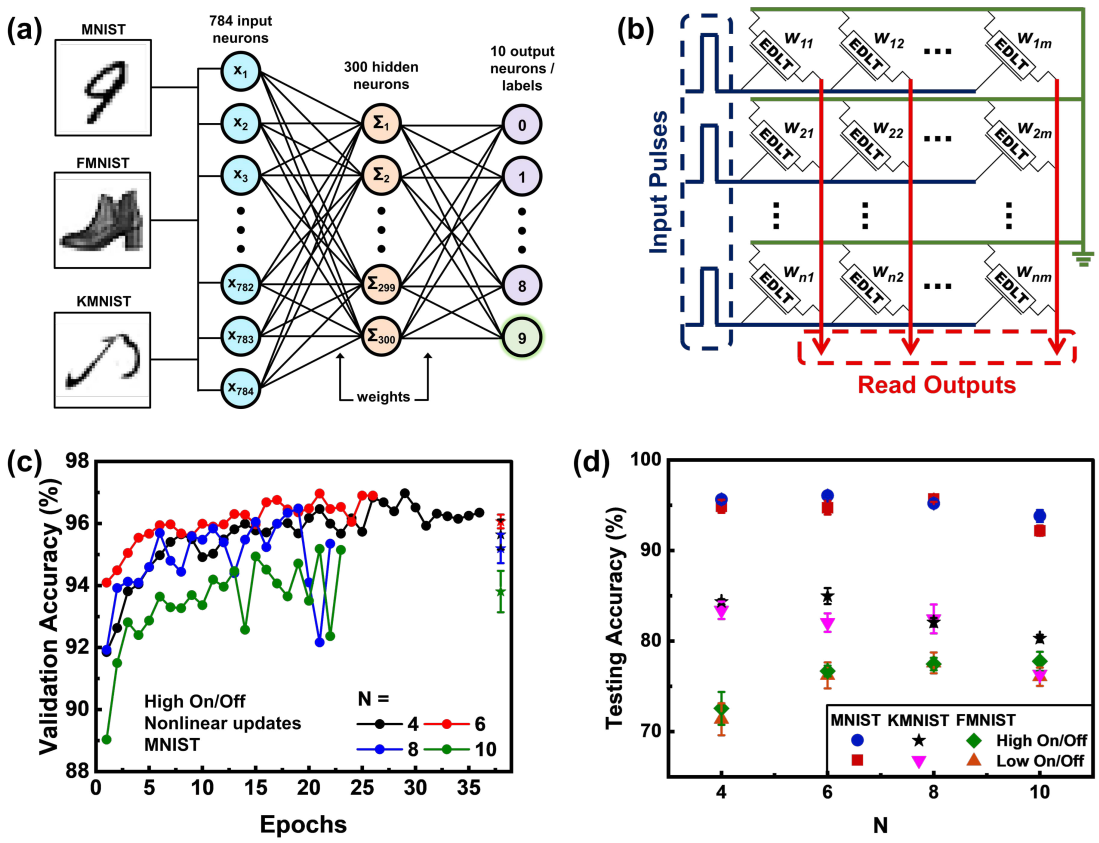

Figure 3: *(a) Fully connected, three-layer feedforward ANN architecture used in this study, and (b) corresponding crossbar-based hardware concept using MoS_2 EDLTs as synaptic elements. (c) Representative validation accuracy versus training epochs for different numbers of conductance states N. (d) Testing accuracy as a function of number of states (N) for high and low on/off devices across MNIST, KMNIST, and FMNIST datasets.*

Fig. 3d summarizes the testing accuracy as a function of the number of states *N* for both high and low on/off devices across the three datasets. As expected, the maximum achieved accuracy decreases with increasing task difficulty, reaching approximately 96% for MNIST, 86% for KMNIST, and 78% for FMNIST. Although KMNIST is often considered more challenging than FMNIST when evaluated using convolutional neural networks (CNNs),[42,43] the relative difficulty can differ for dense or other ANN architectures,[44] and in the present network KMNIST achieves slightly higher accuracy than FMNIST for $N = 4 – 10$. For a given *N*, however, the high and low on/off devices exhibit comparable classification accuracy, with the results generally falling within one standard deviation across all datasets. This is notable because Fig. 2d showed that the low on/off device exhibits 1.5–1.6× larger nonlinearity than the high on/off device within this *N* range. The absence of a distinguishable accuracy difference here suggests that, within this range of nonlinearity, network performance is relatively tolerant to differences in the device weight-update trajectories when *N* is fixed. A likely reason is that

synaptic weights are normalized during ANN training, which reduces the direct influence of absolute device conductance range on the effective weight values in ANN simulations. Nevertheless, in practical hardware implementations the device on/off ratio can influence the signal-to-noise margin of the synaptic weights, an effect that is not explicitly captured in the present ANN model. Consequently, in this regime the number of accessible conductance states play a larger role than the device on/off ratio itself. The influence of N and the role of weight update nonlinearity are examined in more detail in Fig. 4.

Fig. 4a shows how increasing the number of conductance states N affects ANN performance. Under the present pulsing conditions, the low on/off device is limited to approximately 10 distinguishable states, whereas the high on/off device can be extended up to $N = 30$. However, as established in Fig. 2, increasing N also increases the asymmetric nonlinearity parameter β, introducing a potential trade-off between improved weight resolution and degrades update linearity.

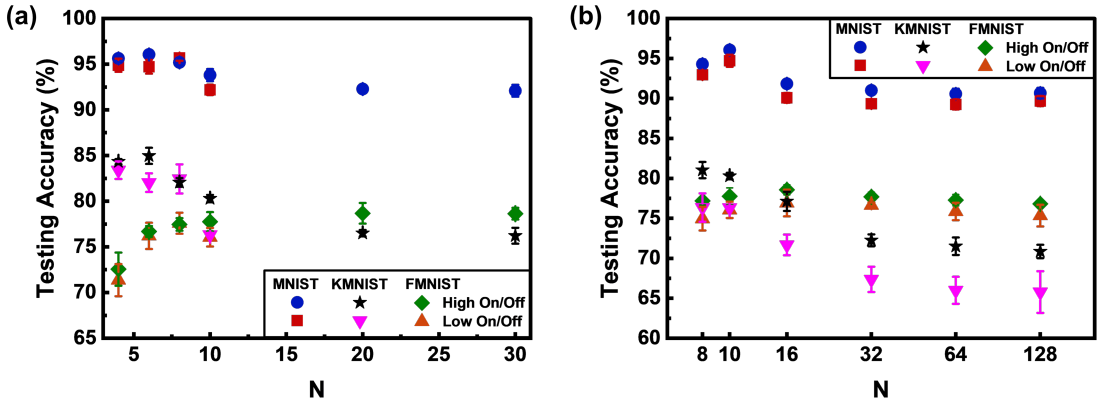

Figure 4: *Effect of conductance state number N on ANN classification accuracy for MNIST, KMNIST, and FMNIST datasets on (a) experimental device characteristics where increasing N is accompanied by increasing update nonlinearity β, and (b) a hypothetical case where β is held constant while N increases, approximating faster pulsing that accesses additional conductance states within the same modulation range.*

The testing accuracy does not increase monotonically with N. For the MNIST task, the accuracy peaks at approximately $N = 4 – 6$ (~96%) and gradually decreases to ~92% at $N = 30$. A similar behavior is observed for KMNIST, where the accuracy reaches a maximum near $N = 4 – 6$ (~85%) before decreasing more noticeably to ~76% at $N = 30$. In contrast, FMNIST exhibits a broader improvement region, increasing from ~72.5% at $N = 4$ to ~77.5% at $N = 10$, followed by a near-saturation behavior up to $N = 30$. Notably, KMNIST achieves slightly higher accuracy at small N, but becomes comparable to or slightly worse than FMNIST at larger N, suggesting

that the two tasks exhibit different sensitivity to weight resolution. These results indicate that each task exhibits a task-dependent optimal number of conductance states, beyond which additional resolution does not improve and can even degrade network performance.

One possible explanation for the observed behavior is the coupling between N and β identified in Fig. 2: increasing N improves weight resolution but simultaneously increases update nonlinearity, which can negatively affect learning dynamics. However, increasing N also increases the effective precision with which synaptic weights are represented, thereby increasing the representational capacity of the network.[45] While this can initially improve learning performance, once sufficient precision is reached for a given task, further increase provides limited benefit and can increase sensitivity to optimization noise and nonlinear update effects. Higher-capacity models are more susceptible to overfitting when additional precision is not required by the task complexity, as the network begins to fit dataset-specific variations rather than generalizable features. [46]

To decouple the effect of N and β, Fig. 4b considers a hypothetical scenario in which β is held constant while the number of states increases from $N = 8$ to $N = 128$. The experimental β values corresponding to the 10-state condition of both devices (β_p = 2.19, 3.3 & β_d = 1.64, 2.66) were used. This simulation also provides an estimate of the performance that could be achieved if faster pulsing instrumentation were used to sample the same conductance modulation range with finer temporal resolution, thereby accessing additional conductance states without increasing β. Under these conditions, the accuracy trends remain qualitatively similar to those observed in Fig. 4a. MNIST performance changes only marginally and saturates near 91–92%, KMNIST shows a gradual decrease from ~81% to ~76% before approaching saturation, and FMNIST improves slightly from ~74% to ~79% as N increases from 8 to 16 states, followed by saturation with a slight downward trend at larger N. Training accuracies, shown in Fig. S5a, exhibit similar trends where the performance decreases and then saturates. The persistence of the non-monotonic behavior when β is fixed indicates that increasing weight resolution alone does not necessarily improve the network's ability to fit the data. Instead, finer weight updates increase the effective model capacity and can enhance sensitivity to stochastic optimization noise, which may accelerate overfitting when the additional precision is not required by the task.[47] In the present implementation, this leads to an earlier onset of overfitting, causing the early stopping criterion to terminate training at fewer epochs and resulting in reduced training and test accuracy at larger N.

Comparing Fig. 4a and Fig. 4b therefore suggests that the dominant factor governing ANN accuracy is the task-dependent optimal weight resolution rather than the maximum achievable number of conductance states. Increasing update nonlinearity further amplifies the

performance degradation at large N, but within the β range investigated here its impact appears secondary to the effect of weight resolution and task complexity. Consequently, an optimal range of conductance states exists for a given task, beyond which increasing state resolution (e.g., through faster pulsing) provides limited benefit.

To further examine the influence of weight update linearity on network performance, Fig. 5 compares nonlinear and linear weight updates obtained from the same high on/off device. This comparison enables a direct evaluation of the role of β while keeping the number of accessible conductance states fixed. As discussed earlier and reported in previous studies, [14,16,30–36] fixed-amplitude pulse trains in ionically gated transistors typically produce nonlinear potentiation and depression. This behavior originates from the coupled ionic dynamics governing EDL formation and relaxation: drift-driven ion migration during each pulse increases the interfacial ion density, while diffusion-driven relaxation occurs between pulses. The competition between these processes leads to progressively smaller incremental conductance changes during repeated stimulation, resulting in saturation-like weight evolution.[20]

To address this limitation, the pulse magnitude can be gradually adjusted during programming to compensate for the intrinsic saturation behavior. Similar strategies are used in the memory industry; for example, incremental step pulse programming (ISPP) schemes employ progressively increasing pulse amplitudes to improve programming precision and reliability,[48–50] and analogous incremental pulsing approaches have been shown to improve learning accuracy in FeFET-based synaptic devices.[51]

Motivated by this concept, we developed a physics-informed predictive pulse-engineering framework that determines the voltage pulse sequence required to generate linear conductance updates. The model is based on the EDLT transient response model introduced by Sun *et al.*,[52] in which ionic drift-diffusion dynamics governing EDL formation and relaxation are coupled to the Enz-Krummenacher-Vittoz (EKV) transistor equation to relate interfacial ionic charge density to channel current. The model parameters were extracted from the transfer characteristics and single-pulse transient measurements shown in Fig. 1b–c. As shown in Fig. 5a, the resulting model reproduces the experimentally measured transient drain current with high fidelity (R^2 = 0.95–0.99), providing a quantitative description of the coupled ionic–electronic dynamics of the device.

Using this experimentally calibrated model, we developed a new algorithm (see Methods) that predicts the sequence of gate-voltage pulses required to produce a specified linear conductance trajectory. The inputs include the number of states N, pulse period T, duty cycle DC, the first pulse amplitude, and the target linearity. The algorithm then reverse-solves for the pulse

amplitudes needed to achieve nearly uniform incremental conductance updates. An example of model-predicted pulse train for 30 states at 1.3 Hz (T = 0.75 s, DC = 50%) is shown in Fig. 5b, where progressively varying pulse amplitudes compensate for the intrinsic ionic saturation dynamics.

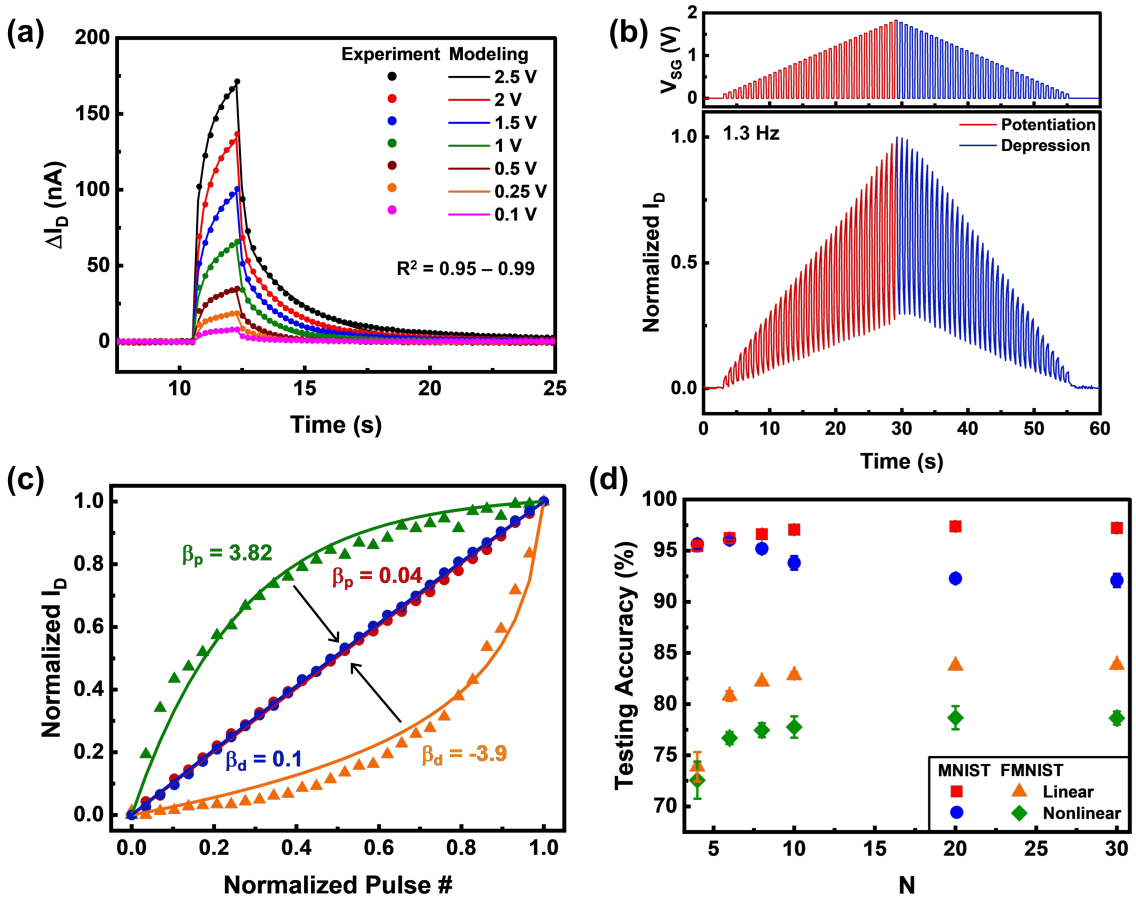

Figure 5: *(a) Validation of the physics-informed EDLT transient model by comparison with experimental current response. (b) Pulse sequence designed using predictive pulse-engineering and the corresponding experimentally measured drain current response demonstrating linear conductance updates. (c) Comparison of asymmetric nonlinearity (β) for nonlinear (fixed-amplitude) and linearized updates. (d) ANN classification accuracy for MNIST and FMNIST datasets showing improved performance with linear weight updates.*

Importantly, the experimentally measured conductance evolution under this predictive pulse-engineered sequence exhibits nearly linear weight updates, yielding β_p = 0.04 and β_d = 0.1. In contrast, the same device driven by fixed-amplitude pulses exhibits strongly nonlinear behavior (β_p = 3.82 and β_d = -3.9), as summarized in Fig. 5c. The extracted β values for all state numbers are listed in Table S1. These results provide the first experimental demonstration that predictive pulse-engineering can effectively linearize synaptic plasticity in ionically gated transistors.

The impact of this improved linearity on ANN performance is summarized in Fig. 5d. For MNIST, the testing accuracy obtained with linear and nonlinear updates is similar for small numbers of states ($N = 4 – 6$), consistent with the earlier observation that devices with fewer states exhibit relatively weak nonlinearity. However, as N increases, linear weight updates consistently yield higher classification accuracy than nonlinear updates for both MNIST and FMNIST, reflecting the improved stability and symmetry of the training process. For example, at $N = 10$, linear updates improve the accuracy of both datasets by approximately 4%. The corresponding KMNIST results (Fig. S5b) show an even larger improvement of ~7% at $N = 10$. For linear updates, the accuracy increases with N and approaches saturation between approximately 10–20 states, indicating that further increases in weight resolution provide diminishing returns for this network architecture and these task complexities. Although additional states could in principle be achieved through faster pulsing or larger pulse amplitudes, doing so would increase programming energy without a proportional improvement in network performance. Larger performance differences between nonlinear and linear weight updates are expected for more complex neural networks and more challenging datasets. For example, prior studies have reported accuracy differences exceeding 10–20% between nonlinear and ideal linear updates in memristor-based and floating-gate memory based networks and more demanding image-classification tasks such as CIFAR-10.[27,53–55]

Overall, these results establish a quantitative framework linking the physical behavior of ionically gated transistors to neural network learning performance. While maximizing the number of conductance states has been a common target for synaptic device optimization, our results show that synaptic performance is governed by coupled trade-offs between conductance range, weight resolution, and update nonlinearity. For relatively simple tasks such as MNIST, only a small number of states is sufficient to achieve near-optimal accuracy, whereas more complex tasks require moderately higher weight resolution. Importantly, the optimal number of states is task dependent and does not increase indefinitely; beyond a certain point, additional states provide diminishing benefit as nonlinear updates and optimization limitations begin to dominate.

Our results also show that when fixed-amplitude pulsing is used, devices with different on/off ratios—and therefore different degrees of nonlinearity—can produce similar network accuracy when the number of accessible states is fixed, indicating a degree of tolerance to device-level variability. However, predictive pulse-engineering provides a systematic pathway to further improve performance by suppressing nonlinear weight evolution. By combining experimental device characterization, physics-informed modeling of EDLT dynamics, and ANN simulations, this work provides the first experimental demonstration that predictive pulse-engineering can linearize synaptic plasticity in ionically gated transistors and translate this improvement into

measurable gains in neural network performance.

More broadly, these results highlight the importance of device-algorithm co-design for neuromorphic hardware. Rather than maximizing device precision alone, practical implementations must balance conductance resolution, update linearity, and programming energy while considering the requirements of the target learning task. The predictive modeling and pulse-engineering framework demonstrated here provides a pathway for systematically optimizing ionically gated synaptic devices for scalable and energy-efficient neuromorphic computing.

Methods

Electric Double Layer Transistor Fabrication

1 x 1 cm^2, 90 nm SiO_2/p-type Si substrates (Graphene Supermarket) were treated with O_2 plasma and MoS_2 flakes (2D Semiconductors) were mechanically exfoliated onto them using scotch tape. Few layer flakes of uniform thickness were selected by optical microscopy. Side-gate, source, and drain contacts were patterned using direct write laser lithography (405 nm, Heidelberg DWL 66^+) with LOR-5A and AZ1505 resists (Kayaku/Merck). CD-26 developer (Dow) was used to develop the patterns. Ti (10 nm)/Au (150 nm) contacts were deposited using e-beam evaporation (CHA), followed by an overnight lift-off in acetone. For the electrolyte preparation, 80 mg of PEO (M_w = 110,000 g/mol, Polymer Standard Service) and 9.6 mg of $LiClO_4$ (99.9% Sigma-Aldrich) were dissolved in 8.9 g acetonitrile to make a 1 wt% solution with 20:1 ether oxygen to Li^+ molar ratio, in an Ar glovebox. 35 – 40 μL of the solution was drop cast onto the substrates with the MoS_2 FETs and annealed at 80°C for 3 mins to evaporate and cooled naturally to room temperature, to form ~1 μm solid electrolyte film.

Electrical Measurements

All devices were vacuum annealed at 75°C for 30 mins in a Lakeshore cryogenic probe station and cooled to room temperature (298 K) overnight. All electrical measurements were performed using a Keysight B1500A semiconductor parameter analyzer at 298 K and under chamber vacuum of ~2 μtorr.

Artificial Neural Network Simulations

ANN simulations were performed using the open-source TensorFlow framework with the Keras interface in Python. A fully connected feedforward network was implemented and trained on the MNIST (handwritten digits), Kuzushiji-MNIST (Japanese characters), and Fashion-MNIST (clothing items) datasets, each consisting of 60,000 training images and 10,000 test images (28 x 28 pixels). The network architecture comprised three layers: an input

layer with 784 nodes corresponding to the flattened pixels, a hidden dense layer with 300 neurons, and an output layer with 10 neurons representing the classification categories. The hidden layer employed the rectified linear unit (ReLU) activation function and the output layer used a softmax activation for classification. Training was performed using stochastic gradient descent (SGD) with a batch size of 4. To emulate synaptic device behavior, the network weights were quantized to N discrete conductance levels between -1 and 1 using a custom quantized dense layer. During training, weight updates were modified using a custom callback that implements asymmetric nonlinear potentiation and depression dynamics parameterized by the experimentally extracted β_p and β_d. This hardware aware-training approach ensures that the weight evolution follows the measured device plasticity characteristics. 10% of the training dataset (6,000 images) was reserved for validation, and training was terminated using an early stopping function based on validation accuracy (patience = 10, min_delta = 0.005). For each parameter set, five independent training runs were performed and the reported accuracies correspond to the average testing accuracy.

Python Predictive Model

The physics-informed Python model was implemented using the NumPy and SciPy libraries, to determine the programming pulse sequences required to achieve controlled linear weight updates. The model parameters were extracted from experimentally measured single-pulse transient responses and transfer characteristics, which were fitting using a drift-diffusion-based description of ion dynamics combined with the EKV transistor model to capture the channel current response.[52] The solver accepts user-defined parameters including pulse frequency, initial gate voltage magnitude, desired linearity (rate of change of conductance), and number of pulses. Based on these inputs, the model first simulates the conductance response to the initial pulse and then recursively computes the subsequent voltage amplitudes required to maintain the specified conductance trajectory throughout the pulse sequence.

Data availability

All data generated or analyzed during this study are included in this published article and its Supplementary Information files. The source code is available from the corresponding author upon reasonable request.

Author Contributions

#N.H.M. and H.H. contributed equally to this work. All authors analyzed the results and wrote the paper.

Acknowledgements

The authors acknowledge Research Computing at the Rochester Institute of Technology for providing computational resources and support that have contributed to the research results reported in this publication.[56]

Funding

This work was partially supported by RIT College of Science.

References

1. *O'Leary, D. E. Artificial Intelligence and Big Data. IEEE Intell. Syst.* ***28****, 96–99 (2013).*
2. *Choi, S., Yang, J. & Wang, G. Emerging Memristive Artificial Synapses and Neurons for Energy-Efficient Neuromorphic Computing. Advanced Materials* ***32****, (2020).*
3. *Ornes, S. Can neuromorphic computing help reduce AI's high energy cost? Proceedings of the National Academy of Sciences* ***122****, (2025).*
4. *Mehonic, A. et al. Roadmap to neuromorphic computing with emerging technologies. APL Mater.* ***12****, (2024).*
5. *Xu, K. & Fullerton-Shirey, S. K. Ionically gated transistors based on two-dimensional materials for neuromorphic computing. 2d Mater.* ***12****, 023003 (2025).*
6. *Sarwat, S. G., Kersting, B., Moraitis, T., Jonnalagadda, V. P. & Sebastian, A. Phase-change memtransistive synapses for mixed-plasticity neural computations. Nat. Nanotechnol.* ***17****, 507–513 (2022).*
7. *Nandakumar, S. R. et al. Experimental Demonstration of Supervised Learning in Spiking Neural Networks with Phase-Change Memory Synapses. Sci. Rep.* ***10****, 8080 (2020).*
8. *Yang, S. et al. Relaxation Signal Analysis and Optimization of Analog Resistive Random Access Memory for Neuromorphic Computing. IEEE Trans. Electron Devices* ***71****, 560–566 (2024).*
9. *Park, S.-O., Jeong, H., Park, J., Bae, J. & Choi, S. Experimental demonstration of highly reliable dynamic memristor for artificial neuron and neuromorphic computing. Nat. Commun.* ***13****, 2888 (2022).*
10. *Yang, S. et al. High-Performance Neuromorphic Computing Based on Ferroelectric Synapses with Excellent Conductance Linearity and Symmetry. Adv. Funct. Mater.* ***32****, (2022).*
11. *Zhong, G. et al. Flexible electronic synapse enabled by ferroelectric field effect transistor for robust neuromorphic computing. Appl. Phys. Lett.* ***117****, (2020).*
12. *Lee, J., Nikam, R. D., Kwak, M. & Hwang, H. Improved Synaptic Characteristics of Oxide-Based Electrochemical Random Access Memory at Elevated Temperatures Using Integrated Micro-Heater. IEEE Trans. Electron Devices* ***69****, 2218–2221 (2022).*
13. *Hendy, H. et al. High-Throughput Time-domain Neuromorphic Computing using Memristors and ECRAM. in 2025 IEEE 34st Microelectronics Design & Test Symposium (MDTS) 1–5 (IEEE, 2025). doi:10.1109/MDTS64924.2025.11177076.*
14. *Zheng, C., Liao, Y., Wang, J., Zhou, Y. & Han, S.-T. Flexible Floating-Gate Electric-Double-Layer Organic Transistor for Neuromorphic Computing. ACS Appl. Mater. Interfaces* ***14****, 57102–57112 (2022).*

15. *Takayanagi, M. et al. Ultrafast-switching of an all-solid-state electric double layer transistor with a porous yttria-stabilized zirconia proton conductor and the application to neuromorphic computing. Mater. Today Adv.* ***18****, 100393 (2023).*
16. *Lee, D.-H., Park, H. & Cho, W.-J. Synaptic Transistors Based on PVA: Chitosan Biopolymer Blended Electric-Double-Layer with High Ionic Conductivity. Polymers (Basel).* ***15****, 896 (2023).*
17. *Yuan, H. et al. High-Density Carrier Accumulation in ZnO Field-Effect Transistors Gated by Electric Double Layers of Ionic Liquids. Adv. Funct. Mater.* ***19****, 1046–1053 (2009).*
18. *Awate, S. S. et al. Impact of Large Gate Voltages and Ultrathin Polymer Electrolytes on Carrier Density in Electric-Double-Layer-Gated Two-Dimensional Crystal Transistors. ACS Appl. Mater. Interfaces* ***15****, 15785–15796 (2023).*
19. *Xu, H., Fathipour, S., Kinder, E. W., Seabaugh, A. C. & Fullerton-Shirey, S. K. Reconfigurable Ion Gating of 2H-MoTe $_2$ Field-Effect Transistors Using Poly(ethylene oxide)-CsClO $_4$ Solid Polymer Electrolyte. ACS Nano* ***9****, 4900–4910 (2015).*
20. *Manimaran, N. H., Sutton, C. L. M., Streamer, J. W., Merkel, C. & Xu, K. Realizing linear synaptic plasticity in electric double layer-gated transistors for improved predictive accuracy and efficiency in neuromorphic computing. Journal of Physics: Materials* ***8****, 015008 (2025).*
21. *Sun, C. et al. Emerging electrolyte-gated transistors for neuromorphic perception. Sci. Technol. Adv. Mater.* ***24****, (2023).*
22. *Li, C. et al. Short-term synaptic plasticity in emerging devices for neuromorphic computing. iScience* ***26****, 106315 (2023).*
23. *Zhang, Y., Ye, J., Matsuhashi, Y. & Iwasa, Y. Ambipolar MoS $_2$ Thin Flake Transistors. Nano Lett.* ***12****, 1136–1140 (2012).*
24. *Li, H.-M. et al. Electric Double Layer Dynamics in Poly(ethylene oxide) LiClO $_4$ on Graphene Transistors. The Journal of Physical Chemistry C* ***121****, 16996–17004 (2017).*
25. *Efetov, D. K. & Kim, P. Controlling Electron-Phonon Interactions in Graphene at Ultrahigh Carrier Densities. Phys. Rev. Lett.* ***105****, 256805 (2010).*
26. *Li, P. et al. Ultra-highly linear Ga2O3-based cascade heterojunctions optoelectronic synapse with thousands of conductance states for neuromorphic visual system. Light Sci. Appl.* ***14****, 354 (2025).*
27. *Tang, J. et al. A Reliable All-2D Materials Artificial Synapse for High Energy-Efficient Neuromorphic Computing. Adv. Funct. Mater.* ***31****, (2021).*
28. *Cheng, X. et al. High Precision Conductance Modulation in CuCrP $_2$ S $_6$ Synaptic Devices for Enhanced Neuromorphic Computing. Adv. Funct. Mater.* ***35****, (2025).*
29. *Kim, T. et al. Spiking Neural Network (SNN) With Memristor Synapses Having Non-linear Weight Update. Front. Comput. Neurosci.* ***15****, (2021).*
30. *Yang, C. et al. All-Solid-State Synaptic Transistor with Ultralow Conductance for Neuromorphic Computing. Adv. Funct. Mater.* ***28****, (2018).*
31. *Hwang, T.-G., Park, H. & Cho, W.-J. Organic–Inorganic Hybrid Synaptic Transistors: Methyl-Silsesquioxanes-Based Electric Double Layer for Enhanced Synaptic Functionality and CMOS Compatibility. Biomimetics* ***9****, 157 (2024).*
32. *Jiang, J. et al. 2D electric-double-layer phototransistor for photoelectronic and spatiotemporal hybrid neuromorphic integration. Nanoscale* ***11****, 1360–1369 (2019).*
33. *Peng, C., Jiang, W., Li, Y., Li, X. & Zhang, J. Photoelectric IGZO Electric-Double-Layer Transparent*

Artificial Synapses for Emotional State Simulation. ACS Appl. Electron. Mater. ***1****, 2406–2414 (2019).*

34. *Kim, H.-S., Park, H. & Cho, W.-J. Biocompatible Casein Electrolyte-Based Electric-Double-Layer for Artificial Synaptic Transistors. Nanomaterials* ***12****, 2596 (2022).*
35. *Feng, X. et al. A novel CVD graphene-based synaptic transistors with ionic liquid gate. Nanotechnology* ***34****, 215201 (2023).*
36. *Wang, D. et al. Recent advanced applications of ion-gel in ionic-gated transistor. npj Flexible Electronics* ***5****, 13 (2021).*
37. *Xiao, T. P., Bennett, C. H., Feinberg, B., Agarwal, S. & Marinella, M. J. Analog architectures for neural network acceleration based on non-volatile memory. Appl. Phys. Rev.* ***7****, (2020).*
38. *Luo, Y. & Yu, S. Accelerating Deep Neural Network In-Situ Training With Non-Volatile and Volatile Memory Based Hybrid Precision Synapses. IEEE Transactions on Computers* ***69****, 1113–1127 (2020).*
39. *Zhang, H.-T. et al. Perovskite neural trees. Nat. Commun.* ***11****, 2245 (2020).*
40. *Seo, S. et al. The gate injection-based field-effect synapse transistor with linear conductance update for online training. Nat. Commun.* ***13****, 6431 (2022).*
41. *Moon, K. et al. RRAM-based synapse devices for neuromorphic systems. Faraday Discuss.* ***213****, 421–451 (2019).*
42. *Ma, X. et al. Mutual learning knowledge distillation framework for adaptive optimization of optical diffractive neural network. APL Photonics* ***10****, (2025).*
43. *Tunheim, S. A. et al. An All-Digital 8.6-nJ/Frame 65-nm Tsetlin Machine Image Classification Accelerator. IEEE Transactions on Circuits and Systems I: Regular Papers* ***73****, 1107–1120 (2026).*
44. *Li, Z. & Meng, L. Deep spiking neural networks for image classification. International Journal of Human Factors Modelling and Simulation* ***8****, 21 (2023).*
45. *Belkin, M., Hsu, D., Ma, S. & Mandal, S. Reconciling modern machine-learning practice and the classical bias–variance trade-off. Proceedings of the National Academy of Sciences* ***116****, 15849–15854 (2019).*
46. *Hastie, T., Tibshirani, R. & Friedman, J. The Elements of Statistical Learning. (Springer New York, New York, NY, 2009). doi:10.1007/978-0-387-84858-7.*
47. *Prechelt, L. Early Stopping - But When? in 55–69 (1998). doi:10.1007/3-540-49430-8_3.*
48. *Nam, K. et al. Origin of Incremental Step Pulse Programming (ISPP) slope degradation in charge trap nitride based multi-layer 3D NAND flash. Solid. State. Electron.* ***175****, 107930 (2021).*
49. *Du, Z. et al. Adaptive Pulse Programming Scheme for Improving the V_{th} Distribution and Program Performance in 3D NAND Flash Memory. IEEE Journal of the Electron Devices Society* ***9****, 102–107 (2021).*
50. *Kim, Y., Kim, J., Kong, J. J., K Vijaya Kumar, B. V & Li, X. Verify level control criteria for multi-level cell flash memories and their applications. EURASIP J. Adv. Signal Process.* ***2012****, 196 (2012).*
51. *Song, M.-K. et al. Recent Advances and Future Prospects for Memristive Materials, Devices, and Systems. ACS Nano* ***17****, 11994–12039 (2023).*
52. *Sun, D. et al. Dynamic Monolayer WSe_2 Electrolyte-Gated Transistor with Coexistent Double Relaxation Timescale for Enhanced Physical Reservoir Computing. Small https://doi.org/10.1002/smll.202504066 (2025) doi:10.1002/smll.202504066.*
53. *Huang, S., Sun, X., Peng, X., Jiang, H. & Yu, S. Overcoming Challenges for Achieving High in-situ Training Accuracy with Emerging Memories. in 2020 Design, Automation & Test in Europe Conference*

& Exhibition (DATE) 1025–1030 (IEEE, 2020). doi:10.23919/DATE48585.2020.9116215.

54. *Chang, C.-C. et al. Mitigating Asymmetric Nonlinear Weight Update Effects in Hardware Neural Network Based on Analog Resistive Synapse. IEEE J. Emerg. Sel. Top. Circuits Syst.* ***8****, 116–124 (2018).*
55. *Park, S.-O. et al. Linear conductance update improvement of CMOS-compatible second-order memristors for fast and energy-efficient training of a neural network using a memristor crossbar array. Nanoscale Horiz.* ***8****, 1366–1376 (2023).*
56. *Rochester Institute of Technology. Research Computing Services. https://www.rit.edu/researchcomputing/ (2026).*

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SIManimaranLinkingEDLTDynamicstoANN.pdf