**University of Bath**

**Alternative formats**
If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

# High-performance neuromorphic computing based on ferroelectric synapses with excellent conductance linearity and symmetry

*Shu-Ting Yang, Xing-Yu Li, Tong-Liang Yu, Jie Wang, Hong Fang, Fang Nie, Bin He, Le Zhao\*, Wei-Ming Lü\*, Shi-Shen Yan, Alain Nogaret, Gang Liu, and Li-Mei Zheng\**

S.-T. Yang, L. Zhao
School of Electrical Engineering and Automation
Qilu University of Technology (Shandong Academy of Sciences)
Jinan 250353, P. R. China.
E-mail: dianxinzl@qlu.edu.cn

S.-T. Yang, X.-Y. Li, T.-L. Yu, F. Nie, S.-S. Yan, L.-M. Zheng
School of Physics
State Key Laboratory of Crystal Materials
Shandong University
Jinan 250100, P. R. China
E-mail: zhenglm@sdu.edu.cn

J. Wang, H. Fang, B. He, W.-M. Lü
Spintronics Institute
University of Jinan
Jinan 250022, P. R. China.
E-mail: weiminglv@hit.edu.cn

J. Wang, H. Fang, W.-M. Lü
Condensed Matter Science and Technology Institute
School of Instrumentation Science and Engineering
Harbin Institute of Technology
Harbin 150080, P.R. China.
E-mail: weiminglv@hit.edu.cn

A. Nogaret
Department of Physics

University of Bath

Bath, BA2 7AY, United Kingdom.


G. Liu

Center for High Pressure Science and Technology Advanced Research,

Shanghai 201203, P.R. China.

Artificial synapses can boost neuromorphic computing to overcome the inherent limitations of von Neumann architecture. As a promising memristor candidate, ferroelectric tunnel junctions (FTJ) enable us to successfully emulate spike-timing dependent synapses. However, the nonlinear and asymmetric synaptic weight update under repeated presynaptic stimulation hampers neuromorphic computing by favoring the runaway of synaptic weights during learning. Here we demonstrate an FTJ whose conductivity varies linearly and symmetrically by judiciously combining ferroelectric domain switching and oxygen vacancy migration. The artificial neural network based on this FTJ-synapse achieves classification accuracy of 96.7% during supervised learning, which is the closest to the maximum theoretical value of 98% achieved to date. This artificial synapse also demonstrates stable unsupervised learning in a noisy environment for its well-balanced spike timing-dependent plasticity response. Our novel concept of controlling ionic migration in ferroelectric materials paves the way towards highly reliable and reproducible supervised and unsupervised learning strategies.

## 1. Introduction

Neuromorphic computing is a candidate technology for replacing conventional von Neuman computing, achieving parallel and energy-efficient computation as well as autonomous learning.[1-6] Designing innovative devices that emulate core functions of synapses is key to implementing neuromorphic computing in hardware. Long-term synaptic plasticity, including long-term potentiation (LTP), long-term depression (LTD), and spike-timing-dependent plasticity (STDP), has been recognized as the basis for learning and memory.[7-9] Nonvolatile memristors, whose conductance can be tuned continuously based on the history of applied voltages or currents and remains stable over hours or days, have the ability to emulate long-term synaptic plasticity.[10, 11] Therefore memristors are considered as suitable electronic

2

materials for artificial synapses. In recent years, artificial synapses have been built via ion migration-,[12-15] phase change-,[16, 17] magnetic-[18, 19] and ferroelectric-based memristors.[20-22] However, a significant drawback of these synaptic devices is the lack of linear response to repeated stimuli and the asymmetric update of conductance caused by electrical impulses of equal amplitude but opposite polarity (Figure S1, Supporting Information).[13, 23] The nonlinear and asymmetric conductance change is known to introduce difficulties for updating weights during learning, reducing network capacity and selectivity in classifying input patterns.[24] In particular, nonlinear and asymmetric conductance changes can result in unbalanced LTP and LTD in the STDP rule. The precise balance between LTP and LTD is critical for STDP-based unsupervised learning, which ensures the steady bidirectional evolution of synaptic strength during learning and leads to a stable bimodal distribution of synaptic weights. Without this balance, synaptic weights may easily run away to extreme values, which compromises network adaptation during the learning phase and reduces the robustness of learning against noise.[25]

Many strategies are proposed to improve the linearity and symmetry of conductance changes, such as optimized device/circuit with 1-transisor/2-resistor (1T2R) structure,[26] complicated pulses programming scheme.[27] However, these approaches either face challenges in terms of limited scalability or high-power consumption. As the nonlinearity and asymmetry issues primarily arise from the intrinsic resistance switching mechanism within materials, a sensible approach is to engineer these microscopic mechanisms to support stable unsupervised learning.

Ferroelectric tunnel junctions (FTJ) offer several advantages for emulating synapses given their giant ON/OFF ratio,[28] fast write/read speed,[29] low power consumption,[30] and good cycling endurance.[31, 32] FTJs exhibits memristor-like response through switching ferroelectric domains with a moderate electric field, resulting in discrete resistance levels which are suitable for encoding the synaptic weights. However, the initial domain nucleation gives an initial fast conductance change followed by a slower conductance variation as the domain expands. This mechanism is responsible for the nonlinearity of synaptic adaptation.[31] In addition, residual asymmetry in the electrodes tends to induce asymmetric switching behavior as a function of positive and negative voltages.[33] It is known that ion migration is also sensitive to electric field, one may expect that ion migration can correct the asymmetry in the interfacial potential drop and slow down the domain nucleation rate in FTJs. Hence judicious control of ion migration would effectively improve the linearity and asymmetry of synaptic conductance changes.

Here we demonstrate an FTJ-type artificial synapse, $Au/Cr/BaTiO_3/Nb:SrTiO_3$ (Au/Cr/BTO/NSTO), in which highly linear and symmetric weight changes were achieved by combining ferroelectric domain switch and oxygen vacancy migration mechanisms. To validate their performance in neuromorphic computing, we simulated two neural network hardware to perform supervised and unsupervised learning, respectively. The supervised pattern recognition exhibited high classification accuracies of 96.7%, the highest level to date, while the unsupervised pattern recognition rate showed high stability and excellent robustness against noise.

## 2. Results

### 2.1. Structural and ferroelectric characterizations

The schematic FTJ structure is shown in **Figure 1**a. A 1.6 nm thick BTO film was deposited on a 0.7 wt% NSTO substrate through pulse laser deposition (PLD). The Au and Cr films with thickness of 50 nm and 5 nm respectively were sputtered on the top of the BTO film as top electrodes. The X-ray diffraction (XRD) patterns of a 100 nm-thick BTO film prepared similarly reveal the good epitaxial growth (Figure 1b). The atomically flat surface with roughness of 113 pm and step-terrace topography can be observed for the 1.6 nm BTO film (Figure 1c). Piezoresponse force microscopy (PFM) phase image shows a ~180° contrast (Figure 1d), and the PFM hysteresis loops exhibit phase and amplitude saturation (Figure 1e), demonstrating the ferroelectric nature of the 1.6 nm BTO film with initial downward polarization pointing to NSTO substrate (Figure S2, Supporting Information). The polarization-voltage (*P-V*) loops of the BTO film demonstrate a high level of leak current (Figure S3, Supporting Information). After annealing under high oxygen pressure, the leak current is significantly inhibited, indicating a large number of oxygen vacancies $V_O^{\cdot\cdot}$ free to migrate under applied electric field in the as-grown BTO film.

### 2.2. Nonvolatile resistance switching

**Figure 2**a shows the current-voltage (*I-V*) characteristics of the FTJs. During the positive voltage scan $(0 \rightarrow V_{max}^+ \rightarrow 0$ V), the FTJ changes from a high resistance state (HRS) to a low resistance state (LRS), whereas it changes from LRS to HRS during the negative voltage scan $(0 \rightarrow V_{max}^- \rightarrow 0$ V), demonstrating bipolar switching. The resistance-voltage (*R-V*) hysteresis loops were conducted with various pulse widths. The HRS increases gradually with pulse

width (Figure 2b). The LRS remains at around 1 kΩ in the range of 1-100 μs (LRS1), then decreases to 0.4 kΩ (LRS2) as the pulse width increases to 1 ms. Two important features can be observed in the *R-V* loops: first, the device exhibits a wide resistance switch window ~2 V (the green shadow area), about 4 times of that in other FTJs (~0.5 V).[34] For the devices with a narrow voltage window, the resistance changes abruptly at the threshold bias $V_{th}$, easily
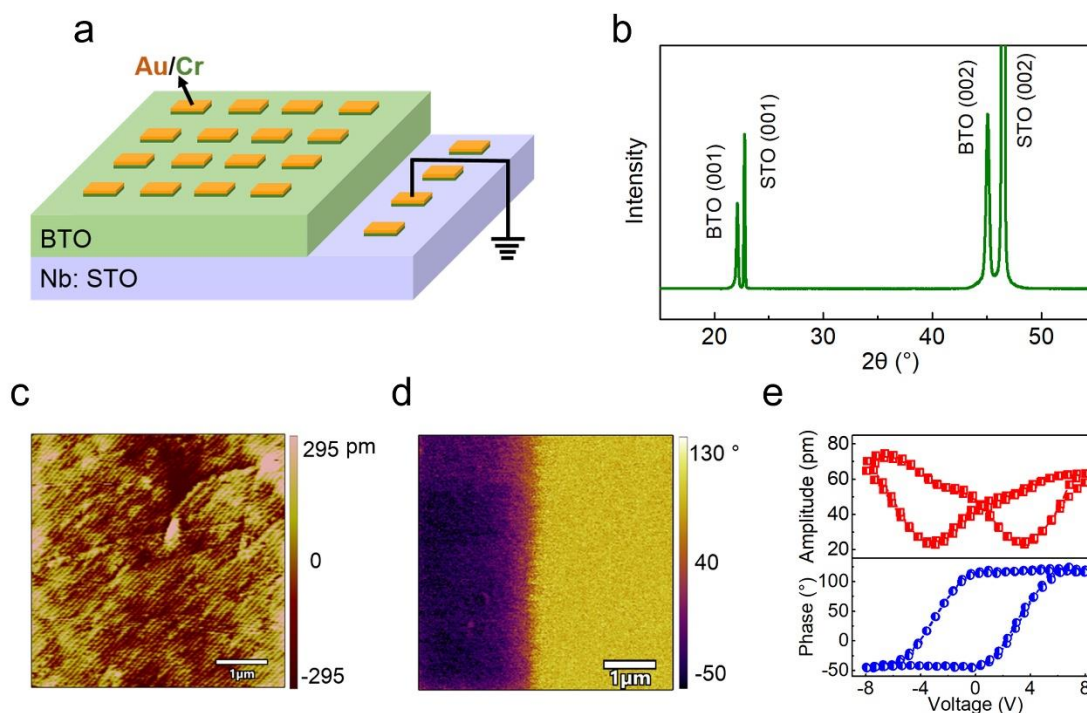


**Figure 1.** Device structure and basic film properties. (a) Schematic diagram of the FTJs; (b) XRD of BTO films; (c) Surface topography, (d) Piezoresponse force microscopy (PFM) phase image, and (e) PFM phase and amplitude loops of the 1.6 nm thick BTO film.
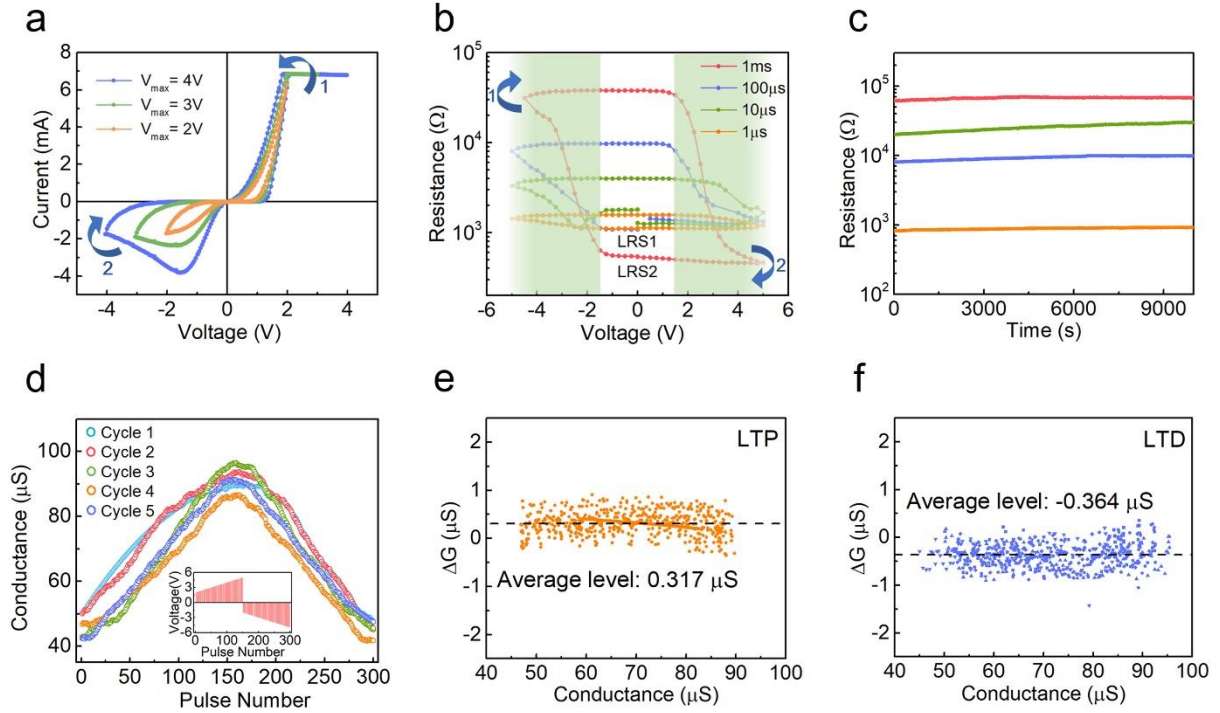
**Figure 2.** Electrical characterization of the FTJ devices. (a) *I-V* curves, (b) *R-V* loops, and (c) retention properties of the device. (d) Linear conductance change with sequential positive and negative pulses measured over 5 cycles. The amplitudes of pulses were increased from |±2| V to |±5| V with a step voltage of 20 mV. The pulse width was fixed at 1ms. (e) and (f) the statistical *ΔG* versus conductance plots for potentiation and depression processes.

forming a discontinuous and nonlinear resistance variation.[35] In contrast, for our device, the resistance exhibits a smoother and almost linear change over a large voltage window. Another feature is the highly symmetrical character of the HRS→LRS and LRS→HRS processes, indicating the same long-term potentiation and depression is achieved when the polarity of electrical impulses changes. All resistance states show good stability over time and remain unchanged for more than $10^4$ s (Figure 2c).

## 2.3. Highly linear and symmetric LTP/LTD and supervised learning

By alternatively applying positive and negative pulses sequence (Inset of Figure 2d), LTP and LTD with high-density nonvolatile states were obtained within a conductance range of 40-90 μS  (Figure 2d). The conductance change  $\Delta G$  induced by a single pulse was extracted and plotted as a function of device conductance $G$ (Figure 2e and f). The mean  $\Delta G$  was approximately 0.32 and -0.36 μS  for the potentiation and depression processes, respectively,

demonstrating a precise modulation below 0.6% has been achieved. Here, the nonlinearity (NL) of weight update was defined as:[36]

$$NL = \frac{\max |G_P(n) - G_D(151-n)|}{G_P(150) - G_P(1)} \; for \; n = 1 \; to \; 150, \tag{1}$$

where $G_P(n)$ and $G_D(n)$ are the conductance values after the $n^{th}$ potentiation pulse and $n^{th}$ depression pulse, respectively. A small NL value indicates high linearity. NL=0 corresponds to a completely linear weight update. Our device shows an extremely low NL value of 0.13-0.17. To best of our knowledge, it is the most linear weight update achieved to date among all two-terminal synapses (Table S1, Supporting Information). In addition, the similar $\Delta G$ values in potentiation and depression processes also guarantee a highly symmetric and equilibrium LTP and LTD process. Especially, the realistic device properties including the cycle-to-cycle and device-to-device variations were studied (Figure S4a-d, Supporting Information). The cycle-to-cycle variation in conductance of our FTJ devices is 5.4%, and the device-to device variation is estimated to be 8.3%. The small variations show that our FTJ device has good uniformity and reliability.

To demonstrate the advantage of the linear and symmetric weight update in neural computing, we performed a supervised pattern classification task in a simulated artificial neural network (ANN) hardware with these FTJ-based synapses (see Methods). A three-layer network (**Figure 3**a) was built to execute the backpropagation algorithm on the basis of the conductance update in Figure 2d. The 28×28 pixel image version of handwritten digits from the "Modified National Institute of Standards and Technology (MNIST)" data set was employed to for the pattern classification task. As shown Figure 3b, the classification accuracy exceeds 90% on the second epoch, and achieves an accuracy as high as 96.7% after
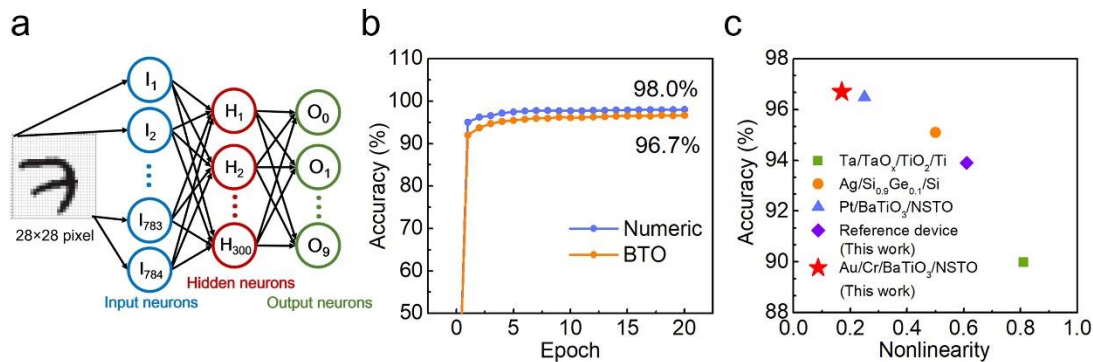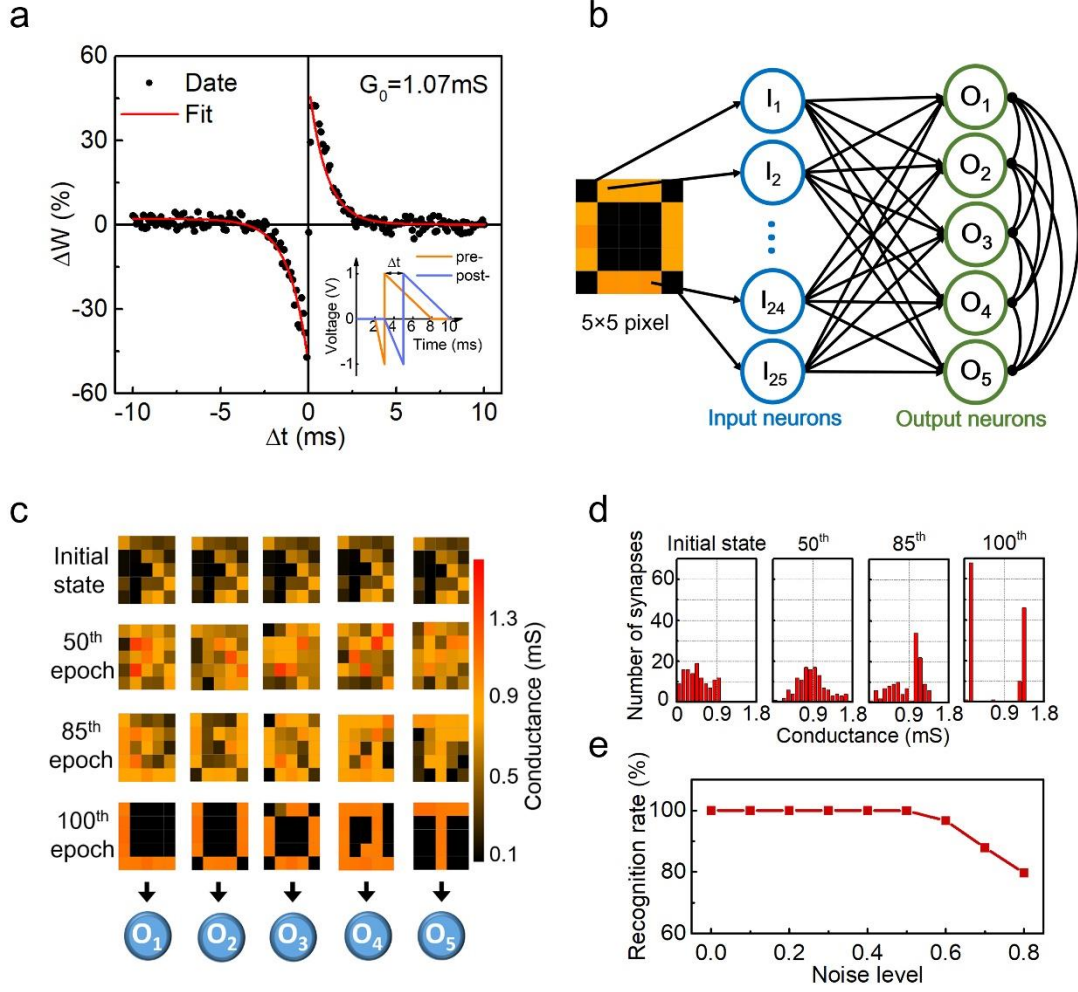


**Figure 3.** Supervised learning with FTJs-based synapses. (a) The designed neural network for implementing the backpropagation algorithm. (b) The training accuracy of FTJs-based synapses and ideal devices. (c) The classification accuracy vs. nonlinearity of some typical two-terminal artificial synapses.[13, 22, 36]

20 training epochs. The accuracy is the highest one to date in the ANNs, very close to the ideal floating point-based neural network performance of 98%. In particular, the accuracy remains as high as 92.1% even after considering the influence of the cycle-to-cycle and device-to-device variations on the neural network simulation (Figure S4e, Supporting Information). In contrast, the reference FTJ device with a NL coefficient of 0.61-0.65, shows a reduced accuracy of 93.9% without considering the cycle-to-cycle and device-to-device variations (Figure S5, Supporting Information). Figure 3c lists the classification accuracy *vs.* nonlinearity of some typical two-terminal artificial synapses. The accuracy depends strongly on the NL value: the lower the nonlinearity, the higher the classification accuracy. The linear weight update makes the classification accuracy in our device the best of all two-terminal devices. Although the weight update linearity of many devices show a dependence of the pulse scheme applied,[37, 38] our device demonstrates low NL values under both stepped and identical pulse schemes, confirming that the high linearity mainly arises from the intrinsic resistance switching mechanism rather than the pulses programming (Figure S6 and Table S2, Supporting Information).

**2.4. STDP and unsupervised learning**

With the highly symmetric weight update, an STDP learning with balanced LTP and LTD is anticipated. STDP is a typical form of long-term plasticity, through which the synaptic strengths evolve depending on the relative timing between presynaptic and postsynaptic spikes. To simulate STDP, two spikes similar to the biological nerve signals (Insert of **Figure 4**a and Figure S7, Supporting Information) with time delay $\Delta t = t_{post} - t_{pre}$ were fed into the presynaptic neuron (the Au/Cr electrode) and postsynaptic neuron (the NSTO electrode), respectively. $\Delta t$ varied as the weight of the FTJ synapse was recorded. As shown in Figure 4a, the change of synaptic weight $\Delta W$ depends strongly on $\Delta t$: LTP occurs if $\Delta t > 0$ and LTD occurs if $\Delta t < 0$. The $\Delta W(\Delta t)$ curve of our device is phenomenologically consistent with the antisymmetric Hebbian learning STDP form. The STDP curves were further quantitatively analyze by fitting with the formula:

**Figure 4.** Unsupervised learning with FTJ-based synapses. (a) The STDP learning curve with balanced LTP and LTD. Inset: The spike schemes of the pre- and post-synaptic. (b) The proposed fully connected spiking neural network for implementing the unsupervised learning through STDP. It has two layers: input layer including 25 input neurons, and output layer including 5 output neurons. The output neurons are lateral inhibited connected. (c) Synaptic weight changes during a learning session. Each $5 \times 5$ matrix represents the group of 25 synapses contributing to the output neurons 1 to 5. As the number of epochs (from top to bottom) increases, the SNN specializes each post-neuron to recognize a different character. (d) Distribution of the synaptic weights during the training session. (e) The variation of the recognition rate with the noise level.

$$\Delta W(\Delta t) = \begin{cases} A_+ \exp\left(-\frac{\Delta t}{\tau_+}\right) + \Delta W_{0+} & \Delta t > 0 \\ -A_- \exp\left(\frac{\Delta t}{\tau_-}\right) + \Delta W_{0-} & \Delta t < 0 \end{cases}, \qquad (2)$$

Where $A_+$ and $A_-$ are parameters regulating the modulation magnitude of synaptic potentiation and depression; $\tau_+$ and $\tau_-$ regulate the temporal range over which LTP and

LTD occur; $\Delta W_{0+}$ and $\Delta W_{0-}$ are constants. The best fit of experimental data gives $A_{\pm} = 50.01/52.18$, $\tau_{\pm} = 1.00/1.11$, and $\Delta W_{0\pm} = 0.20/2.14$. The $\frac{A_+}{A_-} = 0.96$ and $\frac{\tau_+}{\tau_-} = 0.90$ shows rough equality between LTP and LTD, indicating a balanced STDP curve.[39, 40] It is noteworthy that the balanced STDP is independent of initial states (Figure S8, Supporting Information). Thus our device is superior to conventional FTJ synapses for the latter exhibit initial weight-dependent STDP curves,[22] and thus lead to extremely enhanced LTP over LTD for small initial weight and extremely enhanced LTD over LTP for large initial weight.

To demonstrate the advantage of the balanced STDP, an unsupervised pattern recognition task was performed in the simulated spiking neural network hardware (see Methods).[34, 41] Figure 4b shows the SNN including 125 FTJ-based synapses that can perform the STDP-based unsupervised learning. The construction of the SNN follows winner-takes-all strategy. During training, 5 capital characters (Q, L, U, O, and T) given as 5×5 pixel images were presented to the SNN. Each input neuron codes for one pixel of the image. Every output neuron integrates the input signals flowing through the FTJ synapses and is specialized to recognize one specific image at the end of learning. The evolution of synaptic weight and the weight distribution of the 125 synapses as a function of training epochs are demonstrated in Figure 4c and d, respectively. Initially, synaptic weights are set to be randomly distributed and the patterns are completely unrecognizable. During learning, the weights of the 125 synapses autonomously evolve according to the STDP curve in Figure 4a. The learning ends up with a stable bimodal distribution of synaptic weights after 100 epochs (Figure 4d). Correspondingly, the characters are clearly recognized at the synaptic level and the 5 output neurons are specialized to each of the 5 images (Figure 4c). In addition, the simulation also reveals that the recognition rate is robust against noise. As shown in Figure 4e, for images with noise levels below 0.5 times of the input amplitude, the recognition rate remains ~100%. The recognition rate still maintains 80% when noise level increases to 0.8 times of the input amplitude. In contrast, an unbalanced STDP curve fails the unsupervised learning task even with a noise free image (Figure S9, Supporting Information). Our device demonstrates that the balanced STDP contributes greatly to the robustness of unsupervised learning.

## 2.5. Mechanism of resistance switching

The highly selective pattern classifiers in our work are obtained purely by improving the conductance variations of their memristive synapses. The results suggest a novel strategy for improving the capabilities of pattern recognition of supervised and unsupervised neuromorphic computers. To understand the mechanism of the linear and symmetric

conductance changes of the FTJ-based memristors, we further explore their resistance switching kinetics. The common mechanism for FTJs is based on the ferroelectric polarization switching, which could control the tunnel resistance states by changing tunneling potential barrier height and width.[42] This ferroelectric domain switching mechanism is confirmed to exist in our FTJ device (Note 10, Supporting Information). In addition, $V_O^{\cdot\cdot}$ migration could also contribute to resistance switching in the oxide either by forming oxygen vacancy filaments, or by modulating the Schottky barrier. Considering that the typical switching time of ferroelectric polarization is $10^{-4}$-$10^{-7}$ s,[43] the enhanced HRS and decreased LRS by 1 ms pulse width (Figure 2b) confirms the contribution of $V_O^{\cdot\cdot}$ migration to resistance switching. For filamentary-type conduction, the resistance of LRS is independent of electrode area.[28] When the top electrode area of our device is increased 10-fold, the resistance is reduced by an order of magnitude (Figure S11, Supporting Information), thus a filamentary conduction mechanism may be excluded. The $V_O^{\cdot\cdot}$ migrate and accumulate at the interfaces, modulating the Schottky barrier height, hence the resistance.

To identify the tunneling mechanism underpinning the device resistance (Figure 2a and b), we analyzed the *I-V* curves of the device for both LRS and HRS (Note 12, Supporting Information). It is found that the direct tunneling (DT) mechanism dominates the LRS in the low-bias regime, and the Fowler-Nordheim tunneling (FNT) mechanism dominates the high-bias regime (**Figure 5**a). The DT barrier height at Cr/BTO and BTO/NSTO interfaces are 0.577 and 0.164 eV, respectively. For the HRS state, the thermionic injection (TE) mode is responsible for the transport (Figure 5b), where the Schottky barrier height is estimated to be 0.59 eV (Figure S12, Supporting Information). The schematic band diagram of the junction at LRS and HRS may then be deduced, as shown in Figure 5c. At LRS, the polarization points downward to the NSTO, meanwhile the $V_O^{\cdot\cdot}$ migrate from the Cr/BTO interface to the interior of the BTO film under the positive bias. In this case, DT conduction dominates the transport process. At HRS, $V_O^{\cdot\cdot}$ accumulate at the Cr/NSTO interface, and the ferroelectric polarization switches upward pointing to the top electrode. As a result, a depletion layer is formed in the NSTO substrate. Thus, a wide barrier can be obtained and TE is the dominant mechanism. For an FTJ hosting a polarization only switch mechanism, the device should exhibit a narrow voltage window for the resistance switch because of the uniform polarization switch kinetics.[34] On the other hand, with the pure oxygen vacancy migration mechanism, an asymmetric *R-V* loop would be obtained because $V_O^{\cdot\cdot}$ diffuses more easily than it accumulates.[44] The above two mechanisms, ferroelectric polarization switch and oxygen

11
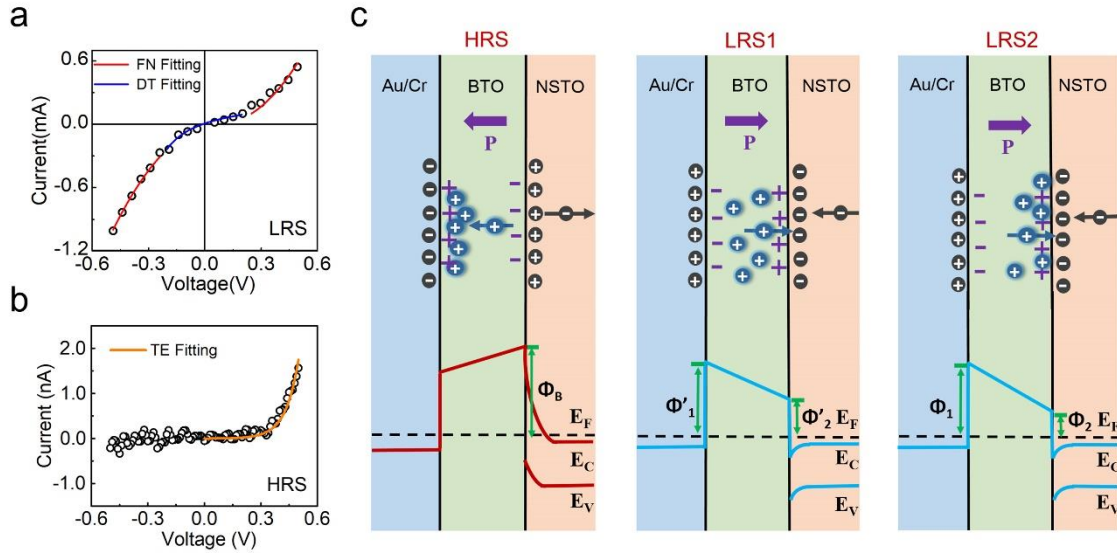
**Figure 5.** Transport mechanism for the FTJs. (a) and (b), measured and fitted *I-V* curves of LRS and HRS for the device. (c) Band diagram for HRS, LRS1, and LRS2. HRS is achieved by applying a negative bias on the top electrode, where the polarization points upward and $V_O^{..}$ accumulate at the Cr/BTO interface. LRS1 is obtained when a positive bias is applied, the polarization point downward to the NSTO substrate and $V_O^{..}$ drift away from the Cr/BTO interface. With a sufficiently long pulse width, the $V_O^{..}$ would accumulate at BTO/NSTO interface which further reduces barrier height and LRS2 is obtained.

vacancy migration under applied bias, coexist in the devices, and contribute collaboratively to the resistance switching. The two mechanisms interact with each other through Coulomb interactions, *e.g.*, domain nucleation rate can be inhibited by $V_O^{..}$ migration, and the $V_O^{..}$ diffusion is suppressed by ferroelectric domains. The combined action of polarization and diffusion widens the voltage window over which the resistance switches and gives a linear change in conductance. To further confirm the contribution of $V_O^{..}$ migration on the conductance variation, we fabricated a reference FTJ device, in which the BTO film is deposited under a high oxygen pressure of 0.5 Torr, and the dominant resistance switching mechanism is ferroelectric domain reversal. Without the contribution of $V_O^{..}$ migration, the reference device exhibits a much higher nonlinear character with NL of 0.61-0.65 (Figure S5a, Supporting Information).

Increasing the pulse width drives more $V_O^{..}$ to accumulate at the Cr/BTO interface, enhancing the Schottky barrier, and consequently increasing HRS. Considering that the mobility of $V_O^{..}$ in BTO is about $10^{-11} \sim 10^{-12}$ cm$^2$/Vs,[45] it takes several milliseconds for the $V_O^{..}$ to migrate from the Cr/BTO interface to the BTO/NSTO interface. A pulse width of 1 ms is therefore

sufficient for the dynamics of $V_O^{\cdot\cdot}$ diffusion to reach the steady state and $V_O^{\cdot\cdot}$ to fully accumulate at BTO/NSTO interface under the positive bias, further reducing barrier height and enhancing the conduction of the LRS (Figure 2b).

The selection of electrodes is another critical factor to achieve highly symmetric conductance change in the FTJ. The two electrodes in our device, Cr and NSTO exhibit small differences in work function (4.5 eV for Cr and ~4.08 eV for NSTO), this causes band bending between the top and bottom interface is easily compensated by $V_O^{\cdot\cdot}$, arising a symmetric conductance variation. In contrast, for the FTJ with Pt and NSTO electrodes, symmetric resistance switching is difficult to achieve even at extremely high $V_O^{\cdot\cdot}$ concentration because the top electrode (5.6 eV for Pt) shows a much higher work function than the bottom one.[43]

## 3. Conclusion

In summary, we presented a high-performance FTJ-type artificial synapse with highly symmetric, linear, and accurate weight update by combining ferroelectric domain switch and oxygen vacancy migration mechanisms. Based on these synaptic devices, a simulated ANN hardware that can perform supervised learning exhibited classification accuracy as high as 96.7% for the MNIST character set, which is the highest to date. Moreover, a balanced STDP curve is obtained. Consequently, unsupervised learning with high robustness against noise is achieved. The elegance of this approach is to avoid redundant circuits that correct nonlinearity in conventional synapses. Our findings demonstrate the potential of ferroelectric materials incorporating vacancy migration for high performance neuromorphic computation.

## 4. Experimental Section/Methods

*Device Preparation*: Epitaxial BTO thin films were fabricated on (001)-oriented $Nb:SrTiO_3$ (Nb:0.7wt%) substrates by pulsed laser deposition using a KrF excimer laser at a frequency of 2 Hz (λ=248 nm). During deposition, the substrate temperature was held at 780 ᵒC, and the oxygen partial pressure was 0.1 Torr. The artificial synapse with an electrode area of 30 ×30 $\mu m^2$ was defined by Laser direct-writting and evaporated Cr (5 nm thick) followed by a capping layer of Au (50 nm thick).

*Characterization*: X-ray diffraction (XRD) was performed on a Rigaku SmartLab X-ray diffractometer with Cu Kα radiation. Surface topographies of the heterostructure were ensured by AFM at contact mode (MFP-3D, Asylum Research). Ferroelectric properties were measured by means of piezoresponse force microscopy (PFM) in the DART (dual a.c.

resonance tracking) mode. The room temperature *P-V* hysteresis loops were measured by using a Radiant Precision Premium Ⅱ ferroelectric tester (Radiant Technology, USA).

*Electrical Measurements*: Electrical characterizations were measured by a Keithley 2410 SourceMeter with homemade programs and an ArC ONE system. The sample was placed in ESP4 probe station with 1μm tungsten steel gold-plated probe.

*Supervised learning*: The artificial neural network (ANN) for performing supervised learning was simulated with a Python package called CrossSim. [46] The package can be used to simulate the neuromorphic computing based on memristor crossbar. It is widely used in research related to memristor-based neuromorphic computing. [6, 47, 48] As shown in Figure 3a, the simulated ANN consisted of three layers, including the input layer, hidden layer and output layer. The number of the input points was dependent on the number of image pixels, and the number of output points was dependent on the number of images to be identified. For MNIST images ($28 \times 28$ pixels), the network size of 784 (input layer)$\times 300 \times 10$ (output layer) was used. A 6000-image training set was used and then a 1000-image testing set was employed to assess the performance of the supervised learning.

*Unsupervised learning*: The spiking neural network (SNN) for performing unsupervised learning was simulated in MATLAB environment by modifying the 'Kohonen network'. The details are displayed in Note 13, Supporting Information.

**Supporting Information**

Supporting Information is available from the Wiley Online Library or from the author.

References

[1]     J. J. Yang, D. B. Strukov, D. R. Stewart, *Nat. Nanotech.* **2013**, 8, 13.

[2]     M. Prezioso, F. Merrikh-Bayat, B. D. Hoskins, G. C. Adam, K. K. Likharev, D. B. Strukov, *Nature* **2015**, 521, 61.

[3]     Z. Wang, S. Joshi, S. E. Savel'ev, H. Jiang, R. Midya, P. Lin, M. Hu, N. Ge, J. P. Strachan, Z. Li, Q. Wu, M. Barnell, G. L. Li, H. L. Xin, R. S. Williams, Q. Xia, J. J. Yang, *Nat. Mater.* **2017**, 16, 101.

[4]     P. Yao, H. Wu, B. Gao, S. B. Eryilmaz, X. Huang, W. Zhang, Q. Zhang, N. Deng, L. Shi, H. P. Wong, H. Qian, *Nat. Commun.* **2017**, 8, 15199.

[5]     Q. Xia, J. J. Yang, *Nat. Mater.* **2019**, 18, 309.

[6]     Y. van de Burgt, E. Lubberman, E. J. Fuller, S. T. Keene, G. C. Faria, S. Agarwal, M. J. Marinella, A. Alec Talin, A. Salleo, *Nat. Mater.* **2017**, 16, 414.

[7]     J. T. Yang, C. Ge, J. Y. Du, H. Y. Huang, M. He, C. Wang, H. B. Lu, G. Z. Yang, K. J. Jin, *Adv. Mater.* **2018**, 30, 1801548.

[8]     C. Ge, C. Liu, Q. Zhou, Q. Zhang, J. Du, J. Li, C. Wang, L. Gu, G. Yang, and K. Jin, *Adv. Mater.* **2019**, 1900379.

[9]     R. Guo, W. Lin, X. Yan, T. Venkatesan, J. Chen, *Appl. Phys. Rev.* **2020**, 7, 011304.

[10]    L. Chua, *Appl. Phys. A* **2011**, 102, 765.

[11]    T. Zhang, K. Yang, X. Xu, Y. Cai, Y. Yang, R. Huang, *Phys. Status Solidi RRL* **2019**, 13, 1900029.

[12]    Y. Shi, L. Nguyen, S. Oh, X. Liu, F. Koushan, J. R. Jameson, D. Kuzum, *Nat. Commun.* **2018**, 9, 5312.

[13]    S. Choi, S. H. Tan, Z. Li, Y. Kim, C. Choi, P. Y. Chen, H. Yeon, S. Yu, J. Kim, *Nat. Mater.* **2018**, 17, 335.

[14]    S. Ham, S. Choi, H. Cho, S. I. Na, G. Wang, *Adv. Funct. Mater.* **2019**, 29, 1806646.

[15]    H. Hu, Y. Li, Y. Yang, W. Lv, H. Yu, W. Lu, Y. Dong, Z. Wen, *Appl. Phys. Lett.* **2021**, 119, 023502.

[16]    D. Kuzum, R. G. Jeyasingh, B. Lee, H. S. Wong, *Nano. Lett.* **2012**, 12, 2179.

[17]    H. Y. Lo, C. Y. Yang, G. M. Huang, C. Y. Huang, J. Y. Chen, C. W. Huang, Y. H. Chu, W. W. Wu, *Nano Energy* **2020**, 72, 104683.

[18]    S. Fukami, T. Anekawa, C. Zhang, H. Ohno, *Nat. Nanotech.* **2016**, 11, 621.

[19]    S. Fukami, C. Zhang, S. DuttaGupta, A. Kurenkov, H. Ohno, *Nat. Mater.* **2016**, 15, 535.

[20]    A. Chanthbouala, V. Garcia, R. O. Cherifi, K. Bouzehouane, S. Fusil, X. Moya, S. Xavier, H. Yamada, C. Deranlot, N. D. Mathur, M. Bibes, A. Barthelemy, J. Grollier, *Nat. Mater.* **2012**, 11, 860.

[21]    Z. Wen, D. Wu, A. Li, *Appl. Phys. Lett.* **2014**, 105, 052910.

[22]    J. Li, C. Ge, J. Du, C. Wang, G. Yang, K. Jin, *Adv. Mater.* **2019**, 32, 1905764.

[23]    S. Choi, J. Yang, G. Wang, *Adv. Mater.* **2020**, 32, 2004659.

[24]    J. Woo, K. Moon, J. Song, S. Lee, M. Kwak, J. Park,H. Hwang, *IEEE Electron Device Lett.* **2016**, 37, 994.

[25]    S. Brivio, D. Conti, M. V. Nair, J. Frascaroli, E. Covi, C. Ricciardi, G. Indiveri, S. Spiga, *Nanotechnology* **2019**, 30, 015102.

[26]    K. Moon, M. Kwak, J. Park, D. Lee,H. Hwang, *IEEE Electron Device Lett.* **2017**, 38, 1023.

[27]    P. Chen, B. Lin, I. Wang, T. Hou, J. Ye, S. Vrudhula, J. Seo, Y. Cao, S. Yu, *IEEE Electron Device Lett.* **2015***, 36, 194.

[28]    R. Guo, Y. Zhou, L. Wu, Z. Wang, Z. Lim, X. Yan, W. Lin, H. Wang, H. Y. Yoong, S. Chen, Ariando, T. Venkatesan, J. Wang, G. M. Chow, A. Gruverman, X. Miao, Y. Zhu, J. Chen, *ACS Appl. Mater. Interfaces* **2018**, 10, 12862.

[29]    C. Ma, Z. Luo, W. Huang, L. Zhao, Q. Chen, Y. Lin, X. Liu, Z. Chen, C. Liu, H. Sun, X. Jin, Y. Yin, X. Li, *Nat. Commun.* **2020**, 11, 1439.

[30]    A. Chanthbouala, A. Crassous, V. Garcia, K. Bouzehouane, S. Fusil, X. Moya, J. Allibe, B. Dlubak, J. Grollier, S. Xavier, C. Deranlot, A. Moshar, R. Proksch, N. D. Mathur, M. Bibes, A. Barthelemy, *Nat. Nanotech.* **2011**, 7, 101.

[31]    H. Y. Yoong, H. Wu, J. Zhao, H. Wang, R. Guo, J. Xiao, B. Zhang, P. Yang, S. J. Pennycook, N. Deng, X. Yan,J. Chen, *Adv. Funct. Mater.* **2018**, 28, 1806037.

[32]    Y. Yang, M. Wu, X. Li, H. Hu, Z. Jiang, Z. Li, X. Hao, C. Zheng, X. Lou, S. J. Pennycook, Z. Wen, *ACS Appl. Mater. Interfaces* **2020**, 12, 32935.

[33]    H. Ryu, H. Wu, F. Rao, W. Zhu, *Sci. Rep.* **2019**, 9, 20383.

[34]    S. Boyn, J. Grollier, G. Lecerf, B. Xu, N. Locatelli, S. Fusil, S. Girod, C. Carretero, K. Garcia, S. Xavier, J. Tomas, L. Bellaiche, M. Bibes, A. Barthelemy, S. Saighi, V. Garcia, **2017**, 8, 14736.

[35]    L. Chen, T. Y. Wang, Y. W. Dai, M. Y. Cha, H. Zhu, Q. Q. Sun, S. J. Ding, P. Zhou, L. Chua, D. W. Zhang, *Nanoscale* **2018**, 10, 15826.

[36]    I. T. Wang, C. C. Chang, L. W. Chiu, T. Chou, T. H. Hou, *Nanotechnology* **2016**, 27, 365204.

[37]    J. Woo, S. Yu, IEEE *Nanotechnol. Mag*. **2018**, 12, 36.

[38]    M. Jerry, P. Y. Chen, J. Zhang, P. Sharma, K. Ni, S. Yu, S. Datta, in *IEEE International Electron Devices Meeting* **2017**, pp. 6.2.1-6.2.4.

[39]    S. Song, K. D. Miller, L. F. Abbott, *Nat. Neurosci.* **2000**, 3, 919.

[40]    M. Yoshioka, *Phys. Rev. E* **2002**, 65, 011903.

[41]    E. Covi, S. Brivio, A. Serb, T. Prodromakis, M. Fanciulli, S. Spiga, *Front. Neurosci.* **2016**, 10, 482.

[42]    Z. Wen, D. Wu, *Adv. Mater.* **2020**, 32, 1904123.

[43]    J. Li, N. Li, C. Ge, H. Huang, Y. Sun, P. Gao, M. He, C. Wang, G. Yang, K. Jin, *iScience* **2019**, 16, 368.

[44]    Q. H. Qin, L. Akaslompolo, N. Tuomisto, L. Yao, S. Majumdar, J. Vijayakumar, A. Casiraghi, S. Inkinen, B. Chen, A. Zugarramurdi, M. Puska, S. van Dijken, *Adv. Mater.* **2016**, 28, 6852.

[45]    F. El Kamel, P. Gonon, L. Ortega, F. Jomni, B. Yangui, *J. Appl. Phys.* **2006**, 99, 094107.

[46]    E. Park, M. Kim, T. S. Kim, I. S. Kim, J. Park, J. Kim, Y. Jeong, S. Lee, I. Kim, J. K. Park, G. T. Kim, J. Chang, K. Kang, J. Y. Kwak, *Nanoscale* **2020**, 12, 24503.

[47]    E. J. Fuller, F. E. Gabaly, F. Leonard, S. Agarwal, S. J. Plimpton, R. B. Jacobs-Gedrim, C. D. James, M. J. Marinella, A. A. Talin, *Adv. Mater.* **2017**, 29, 1604310.

[48]    B. W. Yao, J. Li, X. D. Chen, M. X. Yu, Z. C. Zhang, Y. Li, T. B. Lu, J. Zhang, *Adv. Funct. Mater.* **2021**, 31, 2100069.

By combining ferroelectric domain switching and oxygen vacancy migration, a ferroelectric tunnel junction artificial synapse with intrinsic nonlinearity as low as 0.13-0.17 and symmetric weight updating was developed, which greatly improved the classification accuracy of neural network hardware in supervised learning to 96.7% and enhanced robustness to noise during unsupervised learning.

Shu-Ting Yang, Xing-Yu Li, Tong-Liang Yu, Jie Wang, Hong Fang, Bin He, Le Zhao*, Wei-Ming Lü*, Shi-Shen Yan, Alain Nogaret, and Li-Mei Zheng*

**High-performance neuromorphic computing based on ferroelectric synapses with excellent conductance linearity and symmetry**