An Adaptive Solid-State Synapse with Bi-Directional Relaxation for Multimodal Recognition and Spatio-Temporal Learning

Fang Nie, Hong Fang, Jie Wang, Le Zhao,* Chen Jia, Shuanger Ma, Feiyang Wu, Wenbo Zhao, Shuting Yang, Shizhan Wei, Shuang Li, Chen Ge, Alain Nogaret,* Shishen Yan,* and Limei Zheng*

The brain's unique processing power, such as perception, understanding, and interaction with the multimodal world, is achieved through diverse synaptic functionalities, which include varied temporal responses and adaptation. Although specific functions in brain-like computing have been successfully realized, emulating multimodal recognition and spatio-temporal learning remain significant challenges due to the difficulties in achieving multimodal signal processing and adaptive long-term plasticity in a single electronic synapse. Here, a purely electrically-modulated ferroelectric tunnel junction (FTJ) memristive synapse which realizes multimodal recognition and spatio-temporal pattern identification, through the integration of oxygen vacancies migration and ferroelectric polarization switching mechanisms, providing bi-directional relaxation and adaptive long-term plasticity simultaneously in the isolated device. The bi-directional relaxation enables multimodal recognition in the purely electrically-modulated FTJ device by encoding distinct sensory signals with different electrical polarities. The multimodal perception task is implemented with a multimodal computing system combining visual and speech pattern recognition. Moreover, the adaptive long-term plasticity allows spatio-temporal pattern recognition, which is demonstrated by identifying object orientation and direction of motion with a neural network incorporating the arrayed synapses. This work provides a feasible approach for designing bio-realistic electronic synapses and achieving highly intelligent neuromorphic computing.

1. Introduction

The human brain's intellectual prowess in learning, memorization, and multimodal information processing, derives from complex neural networks comprising billions of neurons interconnected via trillions of synapses. Connections between neurons are reinforced or depressed by synaptic plasticity which occurs on different timescales with powerful adaptability in different brain functions.[1] Long-term synaptic plasticity is traditionally considered as a primary mechanism of learning and memory,[2,3] while short-term synaptic plasticity contributes to the processing of temporal information and the generation of adaptive behavior.[4,5] It is worthy to mention that the coexistence of synaptic plasticity on multiple time scales gives the biological nervous system powerful adaptability, which plays a key role in the remodeling of human sensory cortices. For instance, incorporating adaptation into long-term plasticity is necessary for spatio-temporal learning to occur during stimulation of the human visual cortex.[4,6–8] Therefore, to

F. Nie, C. Jia, S. Ma, F. Wu, W. Zhao, S. Yang, S. Wei, S. Yan, L. Zheng
School of Physics
Shandong University
Jinan 250100, P. R. China
E-mail: shishenyan@sdu.edu.cn; zhenglm@sdu.edu.cn
H. Fang, J. Wang
School of Physics
Harbin Institute of Technology
Harbin 150080, P. R. China
L. Zhao, S. Li
School of Information and Automation Engineering
Qilu University of Technology (Shandong Academy of Sciences)
Jinan 250353, P. R. China
E-mail: dianxinzl@qlu.edu.cn

C. Ge
Beijing National Laboratory for Condensed Matter Physic
Institute of Physics
Chinese Academy of Sciences
Beijing 100190, P. R. China
A. Nogaret
Department of Physics
University of Bath
Bath BA2 7AY, UK
E-mail: pysarn@bath.ac.uk
S. Yan
Spintronics Institute
University of Jinan
Jinan 250022, P. R. China

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
MATERIALS**

www.advmat.de

faithfully replicate the human brain capabilities in neuromorphic devices, implementing adaptive long-term plasticity—specifically, long-term plasticity that adapts to presynaptic stimulation—in a single synaptic emulator is necessary but remains a challenge.

In order to achieve the above aims, synaptic emulators, such as memristors and memtransistors,[9,10] have been considered as the most promising candidates for artificial synapses. Depending on the resistance switching behavior in response to electrical pulses, a synaptic emulator can be modified either non-volatilely or volatilely.[11,12] Non-volatile devices successfully emulate long-term plasticity, and support functions such as pattern recognition in artificial neural networks.[13,14] Volatile emulators are used to simulate short-term plasticity, showing great potential for processing temporal information to meet the needs of reservoir computing[12,15,16] or sensorimotor devices. Since most non-volatile devices and volatile devices are independent with each other, recently some emulators are reported to realize the coexistence of short-term and long-term plasticity. However, these emulators still fail to simulate the adaptability of the long-term synaptic behavior, because their volatile and nonvolatile resistance switching are induced by separate mechanisms without mutual interaction. Therefore, effectively integrating volatile and non-volatile mechanisms in the same solid-state device to achieve adaptive long-term plasticity through their interplay, remains a significant challenge.

Another major challenge is that most synaptic emulators, whose conductance is modulated solely by electric pulses, can only integrate sensory information from a single receptor. This fails to represent the human brain's ability to simultaneously perceive multimodal information, such as visual and auditory signals. Although devices with multi-sensory functions have been proposed, such as optoelectronic devices,[17–20] they are cumbersome, difficult to integrate, and involving complicated processing technologies, making their realization a daunting prospect. Therefore, it would be of great significance to realize multimodal signal processing in a purely electric-modulated device with bi-directional relaxed synaptic plasticity (spontaneous relaxation occurs after both potentiation and depression behavior), where dual voltage polarities can be employed to encode distinct sensory signals.

In this study, we demonstrate a novel ferroelectric tunnel junction (FTJ)-based memristive synapse which combines the oxygen vacancy (represented by $V_O^{\cdot\cdot}$) migration dynamics and polarization switching dynamics. These two mechanisms cooperate to achieve both volatile and non-volatile behaviors in the device. Therefore, the memristive synapse not only forms bi-directional relaxed synaptic plasticity, endowing the electronic synapse with multimodal recognition capability, but also integrates the adaptation characteristics into long-term synaptic behavior, enabling the electronic synapse to demonstrate spatio-temporal learning ability. To demonstrate the uniqueness of the memristive synapses in dual pattern recognition and spatio-temporal pattern identification, we used them to simultaneously recognize image and speech patterns, and to implement the shape orientation and motion detection functions of the human visual cortex.

2. Results and Discussion

2.1. Multi-Timescale Plasticity and Bi-Directional Relaxation Induced by Oxygen Vacancies Migration and Ferroelectric Polarization Switching

In this study, a classical FTJ structure, Au/Cr/BaTiO$_3$ (1.6 nm)/Nb:SrTiO$_3$ (Au/Cr/BTO/NSTO), is employed to implement synaptic behavior. **Figure 1a** schematically illustrates the similarity between the FTJ-based electronic synapse and a biological synapse in terms of microscopic dynamics. In the biological synapse, the rise and spontaneous decay of postsynaptic Ca^{2+} concentration underpins the internal dynamics, and hence provides the synapse's temporal response and the synaptic weight adaptation (top panel of Figure 1a). In the FTJ synaptic emulator, the BTO ferroelectric layer is deposited at a low oxygen pressure of 0.02 Torr to introduce abundant oxygen vacancies, $V_O^{\cdot\cdot}$s (see the Methods section). The $V_O^{\cdot\cdot}$ drift along with the ferroelectric polarization switching under external voltage stimulus gives rise to memristive behavior. The interaction between $V_O^{\cdot\cdot}$ migration/diffusion dynamics and ferroelectric polarization (represented by P) switching/back-switching dynamics are supposed to provide the integrated volatile and nonvolatile conductance changes,[21,22] mimicking the transient and long-term weight changes of biological synapses. The ferroelectric nature of the BTO film is confirmed by piezo-response force microscopy (Figure S1, Supporting Information). The presence of $V_O^{\cdot\cdot}$s is verified by X-ray photoelectron spectroscopy along with the enlarged lattice parameters of the BTO film (Note S1, Supporting Information).[23–26] Figure 1b shows the current–voltage (I–V) characteristics of the memristors. Following repeated positive/negative bias cycles, the magnitude of the current gradually increases/decreases, indicating the tunability of device conductance. Taking the conductance as the synaptic weight, a series of positive or negative pulses cause consecutive potentiation or depression of the memristive synapse (Figure 1c). The cycle-to-cycle and device-to-device variations in conductance were estimated to be below 1.21% and 1.93% (Note S2, Supporting Information), respectively, confirming our device's excellent endurance and good uniformity. Both potentiation and depression states partially relax over time, simultaneously presenting long-term plasticity and short-term plasticity (Figure 1d) in an isolated device. The spontaneous decay behavior indicates that volatile dynamics have been successfully introduced into the device. Of particular interest is the bi-directional conductance relaxation in this synaptic emulator (Figure 1d), which is rarely observed in ferroelectric synapses. Furthermore, our device retains the typical advantages of ferroelectric synapses, including excellent controllability, low energy consumption, and minimal cycle-to-cycle and device-to-device variations (Note S2, Supporting Information). These features collectively enhance the device's potential as an ideal neuromorphic component suitable for programming applications. The bi-directional relaxation feature indicates that the device can be independently programmed by both positive and negative pulses, making multimodal signal coding and multimodal perception feasible. Figure 1e further confirms the bi-directional relaxation feature of multiple conductance states.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com
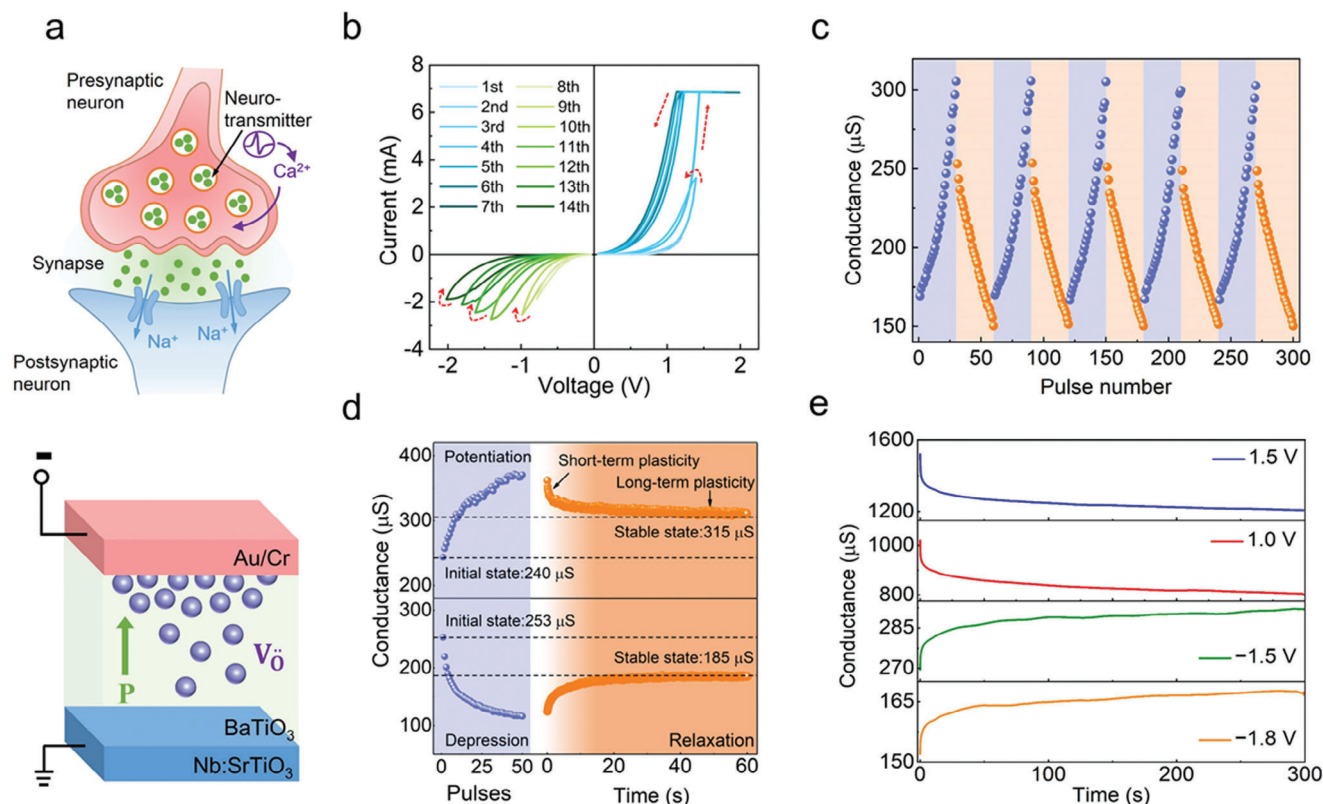
**ADVANCED
MATERIALS**

www.advmat.de

Figure 1. Synaptic efficacy of the Au/Cr/BTO/NSTO memristor. a) Similarity between biological synapses and Au/Cr/BTO/NSTO device in terms of microscopic dynamics. The kinetics of $V_{\ddot{O}}$ in the memristor should be analogous to Ca^{2+} in the biological synapses. b) $I-V$ curves of the device measured across seven sweeps of positive and negative voltage cycles, where the red arrows indicate the cyclic direction of electrical scanning. c) Potentiation and depression of the memristive synapse; the programing pulse scheme is shown in Figure S4b (Supporting Information). d) Spontaneous decay of potentiation and depression states. During the relaxation process, the pre-synaptic voltage was removed leaving only small reading voltage (0.1 V) to read the conductance (Figure S5, Supporting Information). e) Conductance relaxation of 4 different states: 1520 μS, 1020 μS, 269 μS, and 152 μS, which are induced by pulse trains of a fixed frequency of 3.33 Hz, a pulse number of 10, a pulse width of 100 μs, and different amplitudes: 1.5, 1.0, −1.5, and −1.8 V, respectively.

By fitting *I-V* curves of this FTJ device, it is verified that at high conductance state (HCS), the direct tunneling mechanism dominates in the low-bias regime (below 0.3 V) and the Fowler-Nordheim tunneling mechanism dominates in the high-bias regime. While at the low conductance state (LCS), the thermionic emissionmechanism is responsible for the transport (Note S3, Supporting Information).[27] The barrier height Φ_B of the memristor can also be obtained by fitting the *I-V* curves[28–30] (Figure S7, Supporting Information). As shown in **Figure 2**a, Φ_B decreases/increases gradually under the positive/negative bias. The equivalent barrier changes induced by $V_{\ddot{O}}$ drift (Φ_V) and *P* switching (Φ_P) during the transition from low to high conductance are extracted (Note S4, Supporting Information) and summarized in Figure 2b, and those for the high-to-low transition are shown in Figure S9 (Supporting Information). It can be seen that both $V_{\ddot{O}}$ drift and *P* switching play important roles in the conductance modulation. In addition, the conductance relaxation is well fitted by a double exponential decay function, giving rise to two activation energies of 0.23 and 0.7 eV (Note S5, Supporting Information). The former activation energy (0.23 eV) is ascribed to ferroelectric *P* reversal because its activation energy coincides with the barrier energy of *P* back-

switching.[31] The latter activation energy is ascribed to $V_{\ddot{O}}$ back-diffusion as this activation energy (0.7 eV) matches that of $V_{\ddot{O}}$ back-diffusion in the Au/Cr/STO/NSTO device (Figure S10c, Supporting Information).[32] Therefore, it is the combined effect of $V_{\ddot{O}}$ back-diffusion and *P* back-switching that leads to conductance relaxation.

Based on the above experimental results, the schematic description of internal dynamics is illustrated in Figure 2c, which is responsible for the conductance modulation and bi-directional relaxation behavior. When a positive voltage pulse is applied to the Au/Cr electrode, *P* switches gradually to point to the BTO/NSTO interface. Meanwhile, the positively charged $V_{\ddot{O}}s$ also migrate and accumulate on the interface, forming an electron accumulation layer at the NSTO side, lowering the barrier height at the BTO/NSTO interface (left panel of Figure 2c). Thus, the positive pulse results in a gradual increase in conductance. On the other hand, when negative pulses are applied, both *P* switching and $V_{\ddot{O}}$ migration contribute to the increase in barrier height, resulting in a decrease in device conductance (right panel of Figure 2c).[33–35] In turn, the accumulation of $V_{\ddot{O}}s$ at electrode/BTO interfaces forms a $V_{\ddot{O}}$ concentration gradient, leading to the $V_{\ddot{O}}$ back-diffusion. In addition, the repulsive force
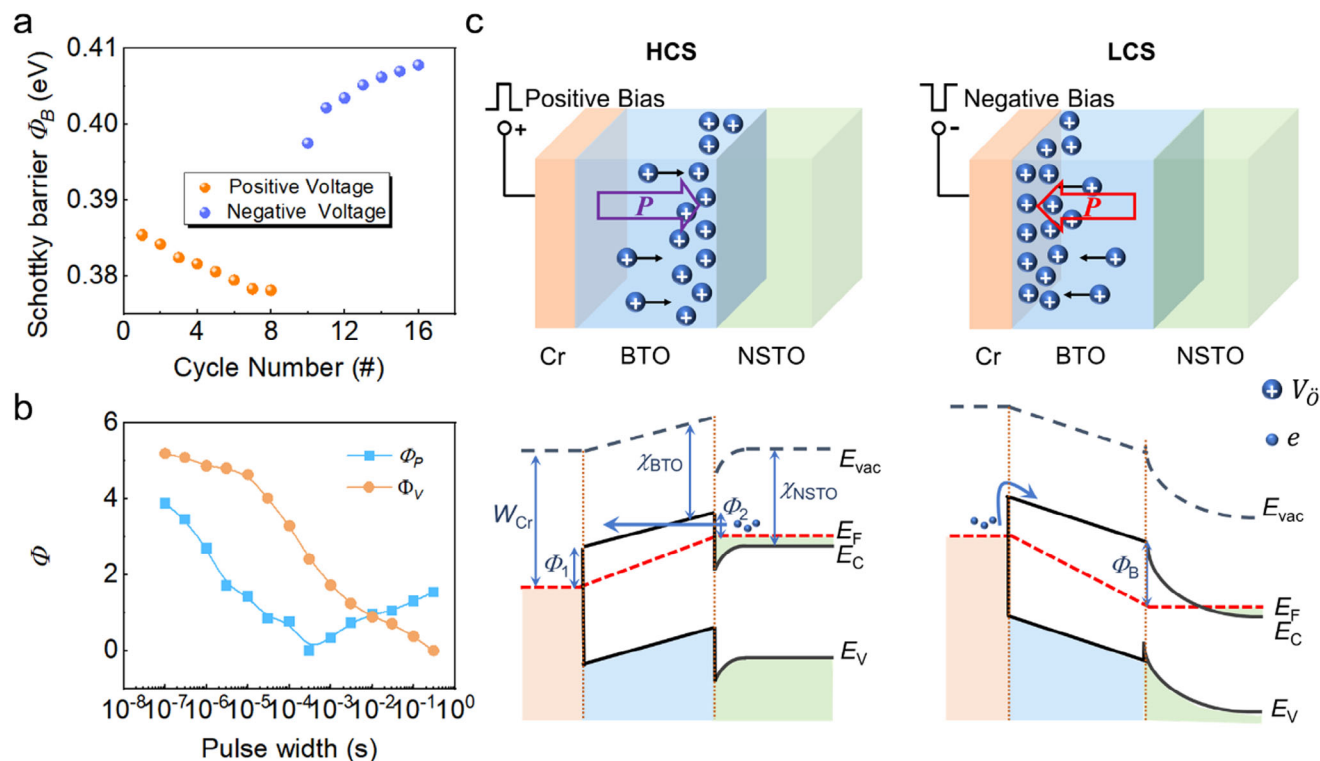
Figure 2. Internal dynamics of the memristor. a) Evolution of barrier height Φ_B with continuous positive and negative bias sweeping. b) Evolution of equivalent barrier changes, Φ_V induced by $V_{\ddot{O}}$ migration and Φ_P induced by P switching, as a function of pulse width for the low-to-high conductance transition. c) Schematic demonstration of $V_{\ddot{O}}$ and P dynamics in the HCS induced by positive pulses (left panel) and the LCS induced by positive pulses (right panel), and the corresponding band diagrams. At HCS, direct tunneling mechanism is dominant, thus Φ_1 and Φ_2 are adopted to demonstrate the two barrier heights at the interface. At LCS, thermionic emission mechanism is mainly responsible for the transport, thus Φ_B is used to demonstrate the barrier height (Note S3, Supporting Information).

between P bound charges and the accumulated $V_{\ddot{O}}s$ at the interfaces causes the back-switching of polarization. P back-switching and $V_{\ddot{O}}$ back-diffusion can happen after the removal of both positive and negative pulses, thus producing bi-directional relaxation behaviors (Figure 1d,e).

It should be noted that the appropriate amount of $V_{\ddot{O}}s$ in the BTO layer is crucial to generate the combined volatile and non-volatile resistance switching in the FTJ memristor. For the FTJ with negligible $V_{\ddot{O}}$ content, non-volatile characteristic is obtained, and therefore can only mimic the long-term learning activities (Figure S11, Supporting Information). Increasing the $V_{\ddot{O}}$ concentration, achieves both the direct internal temporal dynamics (the $V_{\ddot{O}}$ back-diffusion) and indirect dynamics (P back-switching due to the interactions between $V_{\ddot{O}}$ and P), needed to successfully integrate short-term and long-term plasticity in a single FTJ synapse (Figure 1d), providing adaptability of long-term plasticity for spatio-temporal learning. More importantly, the temporal dynamics induced by $V_{\ddot{O}}$ gives bi-directional relaxation behavior, enabling the FTJ synapse capable to encode multimodal signals. It should be noted that this strategy is applicable to a wide range of materials, and that bi-directional relaxation behavior was also achieved in a Pb(Zr,Ti)O$_3$-based FTJ device by introducing appropriate amount of $V_{\ddot{O}}s$ into the ferroelectric layer (Figure S14, Supporting Information).

2.2. Short-Term Synaptic Dynamics with Bi-Directional Relaxation for Multimodal Perception

Short-term dynamics play an important role in encoding temporal information in auditory or visual signals.[36] Paired-pulse facilitation (PPF) and paired-pulse depression (PPD) are two important forms of short-term plasticity, which describe the enhancement and depression of excitatory postsynaptic currents (EPSCs) evoked by the second pulse in comparison to the first, respectively. Simulations of PPF and PPD in our memristor are shown in **Figure 3**a,c, respectively. The PPF and PPD index can be quantified using the following Equation (1):[37]

$$\text{PPF or PPD} = \frac{P2 - P1}{P1} \times 100\% \tag{1}$$

where $P1$ and $P2$ are the EPSC peaks evoked by the first and second presynaptic pulse, respectively. A typical exponential decay of PPF or PPD index as a function of pulse interval Δt was found (Figure 3b,d), which can be fitted by a double exponential Equation (2):[38]

$$\gamma = c_1 \exp\left(-\frac{\Delta t}{\tau_1}\right) + c_2 \exp\left(-\frac{\Delta t}{\tau_2}\right) \tag{2}$$

ADVANCED
SCIENCE NEWS

www.advancedsciencenews.com

ADVANCED
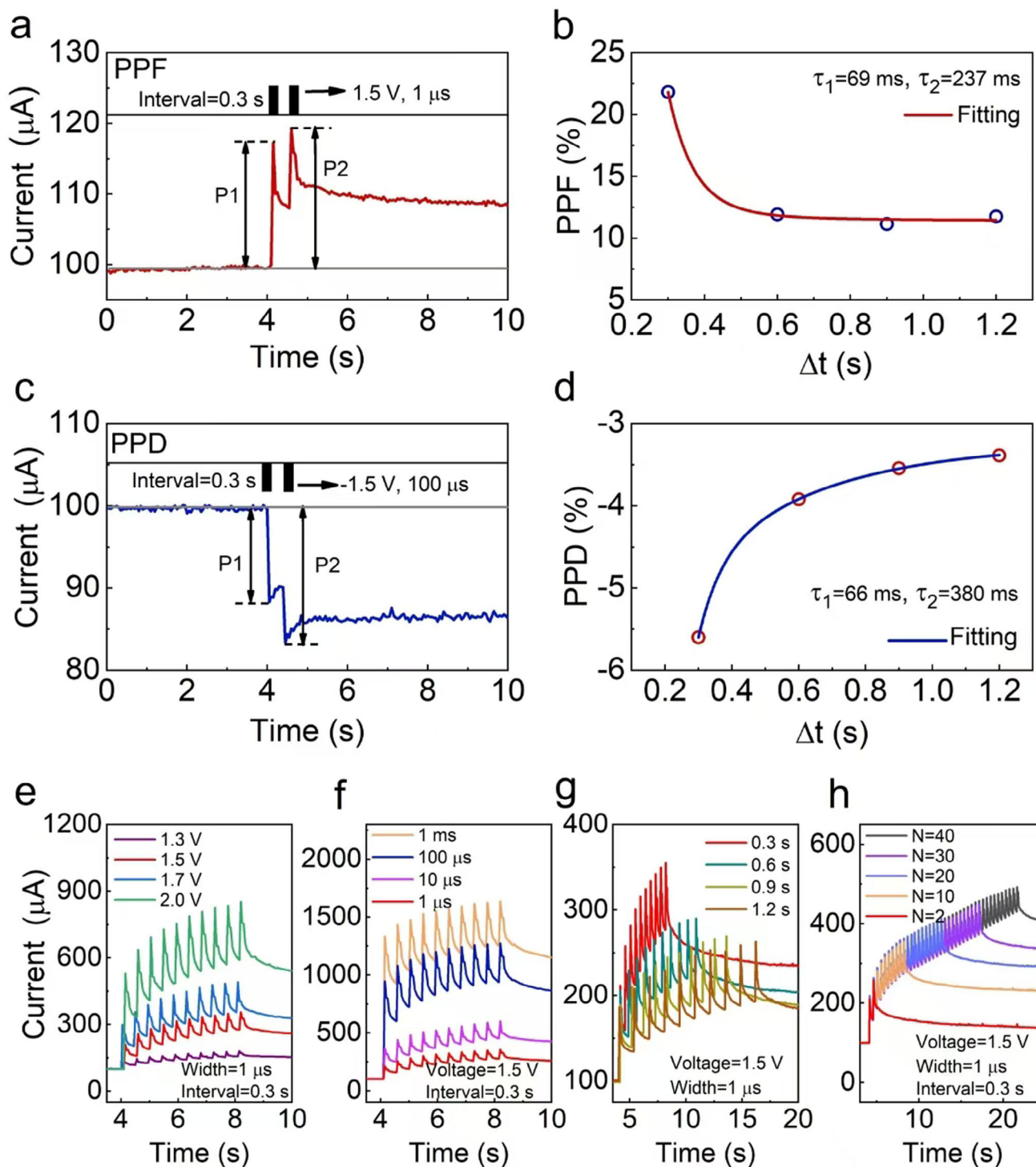**MATERIALS**

www.advmat.de

Figure 3. Short-term dynamics. a) and c) PPF and PPD realized by applying a pair of positive and negative pulses, respectively. b) and d) Varying PPF and PPD index versus pulse internal Δt. e–h) Response of EPSCs to stimuli pulses of varying amplitudes, pulse widths, pulse intervals, and pulse numbers.

where c_1 and c_2 are the initial magnitudes, and τ_1 and τ_2 are the relaxation times. Fitting the experimental results gives $\tau_1 = 69$ ms, $\tau_2 = 240$ ms for PPF, and $\tau_1 = 66$ ms and $\tau_2 = 380$ ms for PPD, respectively. The similarity of the two relaxation times between PPD and PPF further confirms that the relaxation in both

potentiation and depression states arises from the same mechanism, i.e., the combined $V_O^{\cdot\cdot}$ back-diffusion and P back-switching. To further verify our device's temporal response, the EPSC induced by a presynaptic pulse of different amplitude, width, time interval, and number is monitored, the results are shown in

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
MATERIALS**

www.advmat.de

Figure 3e–h. Increasing the amplitude, width, frequency, and number of input pulse trains can all enhance the EPSC peaks, as well as increase the final conductance states, indicating that both long-term and short-term synaptic behavior show excellent adaptability to the stimulus. It is noteworthy that in our device, the long-term and short-term plasticity are deeply intertwined, in contrast to other synaptic devices in which short-term plasticity and long-term plasticity are achieved separately by adjusting pulse schemes.[39,40]

Through its bi-directional relaxed short-term dynamics, the memristive synapse is expected to encode multimodal signals. To demonstrate this functionality, we conducted digit recognition tasks based on both image and speech patterns. Three numbers, "0", "5" and "8" (Figure S15, Supporting Information), were chosen for their susceptibility to confusion when image corruption occurs, so that additional spoken information may provide resolution. Both the image and speech recognition functions were implemented on a reservoir computing (RC) network. The RC system is an efficient neural network well-suited for processing temporal information.[41] It typically consists of an input layer, a reservoir, and an output layer (**Figure** 4a). The reservoir can nonlinearly map a time-dependent input $\vec{u}(t)$ into a feature space represented by the reservoir internal state $\vec{x}(t)$, and here the reservoir is composed of our FTJ memristors. The output layer is a neural network that can be trained to recognize the internal state $\vec{x}(t)$. As only the output network needs to be trained during learning while the reservoir does not, the training cost can be significantly reduced. The reservoir is generally constructed from electronic synapses with unidirectional relaxation, thus it can only be programmed by unipolar (negative or positive) input pulses, limiting it to single-mode signal coding (left panel of Figure 4b). One advantage of bi-directional relaxation, is that our FTJ memristor may be programmed independently by both negative and positive electrical pulses. When signals of different modes are encoded separately by pulses of different polarities, the reservoir exhibits multimodal perception and multimodal signal recognition capability (right panel of Figure 4b). In order to achieve these functions, we designed a FTJ memristor array consisting of two functional parts: one part for encoding image information with positive electrical pulses, and the other for encoding voice information with negative electrical pulses (Figure 4c).

For image "5" recognition, the image is divided into a 5 × 4-pixel spatial pattern (left panel of Figure 4d). The red pixel corresponds to a positive write pulse (1 V, 1 ms duration, and 20 ms interval), while the white one corresponds to no pulse. In this way, the four pixels in each row are sequentially encoded in one voltage waveform, and the entire pattern is encoded as a set of five such waveforms which constitute the image inputs of the RC system (right panel of Figure 4d). Five Au/Cr/BTO/NSTO memristors were utilized in the reservoir layer to process the input information, with each memristor processing a pulse stream. The reservoir state is then represented by the collective EPSCs of the five Au/Cr/BTO/NSTO memristors. The reservoir states of the three digital images "0", "5" and "8" are shown in Figure S17 (Supporting Information). They exhibit significant differences, demonstrating the ability of the reservoir to clearly distinguish different pulse streams. The sensitivity of the reservoir states to stimulation and their reproducibility were further validated (Note S8, Supporting Information). The recognition rate de-

creases greatly for the for corrupted images (Figure S22, Supporting Information). Therefore, voice recognition is necessary to strength the discrimination between different digits.

For recognizing the voice pattern "5", the audio waveforms of digits are taken from the TI-46 database.[42] Feature extraction of the audio signals was performed with procedures shown in Figure 4e–g. First, the original audio waveform (Figure 4e) was converted to a cochlea-gram with 64 frequency channels per frame using Lyon's passive ear model (Figure 4f).[43] Then 5 spectral lines in the middle of the cochlea-gram (circled by the red dashed box in Figure 4f) that best represent the waveform features are retained in a 5 × 64 matrix. This matrix is then down sampled by multiplying it with a 64 × 4 mask matrix composed of randomly assigned binary values (1 and −1).[42] This produces a matrix with 5 × 4 negative elements that best encapsulate the features of the original recording. The 20 elements of the matrix (5 × 4 = 20) are represented as a time series and appropriately scaled. This produces the voltage sequence, shown in Figure 4g, which is then input into an Au/Cr/BTO/NSTO memristor. The current sequence output by the memristor (Figure 4h) serves as a reservoir state encoding the spoken digit. We experimentally confirmed that our device has small cycle-to-cycle and device-to-device variation when generating these speech reservoir states (Note S10, Supporting Information). Finally, the reservoir states that correspond to image and voice information of the digits are fed into the output layer of the RC system for digit recognition. This multimodal approach to digit recognition is particularly crucial when the image or/and the audio is partially lost. For corrupted pixel images with only the first two rows left (as shown in Figure S21, Supporting Information), the classification accuracy is only 64% (Figure S22b, Supporting Information). And when the audio loses 80% of its information, the classification accuracy drops to 62.7% (Figure S23a, Supporting Information). By leveraging multimodal information through the combination of the damaged spoken signal and the corrupted image signal, we significantly improved the recognition accuracy to 95.4% (Figure S23b, Supporting Information).

2.3. Adaptive Long-Term Plasticity for Spatio-Temporal Recognition

History adaptation describes the cumulative changes in synaptic weight that have occurred from past sequences of stimuli. The frequency, amplitude, polarity, and number of historically experienced spikes may affect the subsequent synaptic plasticity.[39] Combining history adaptation characteristics into long-term synaptic behavior to obtain the adaptive long-term plasticity is the prerequisite for spatio-temporal related learning.[44] One evidence for the adaptive long-term plasticity in our device is shown in Figure 3e–h, where the final conductance states depend strongly on the amplitude, width, frequency, and number of prior pulses. Additionally, metaplasticity, which describes the phenomenon that a pre-stimulus alters the subsequent long-term plasticity,[45,46] is also successfully implemented in the memristive synapse (Note S11, Supporting Information). The results demonstrate that both long-term potentiation and long-term depression can be modulated by the polarity of the pre-stimulus, showing the adaptation of long-term
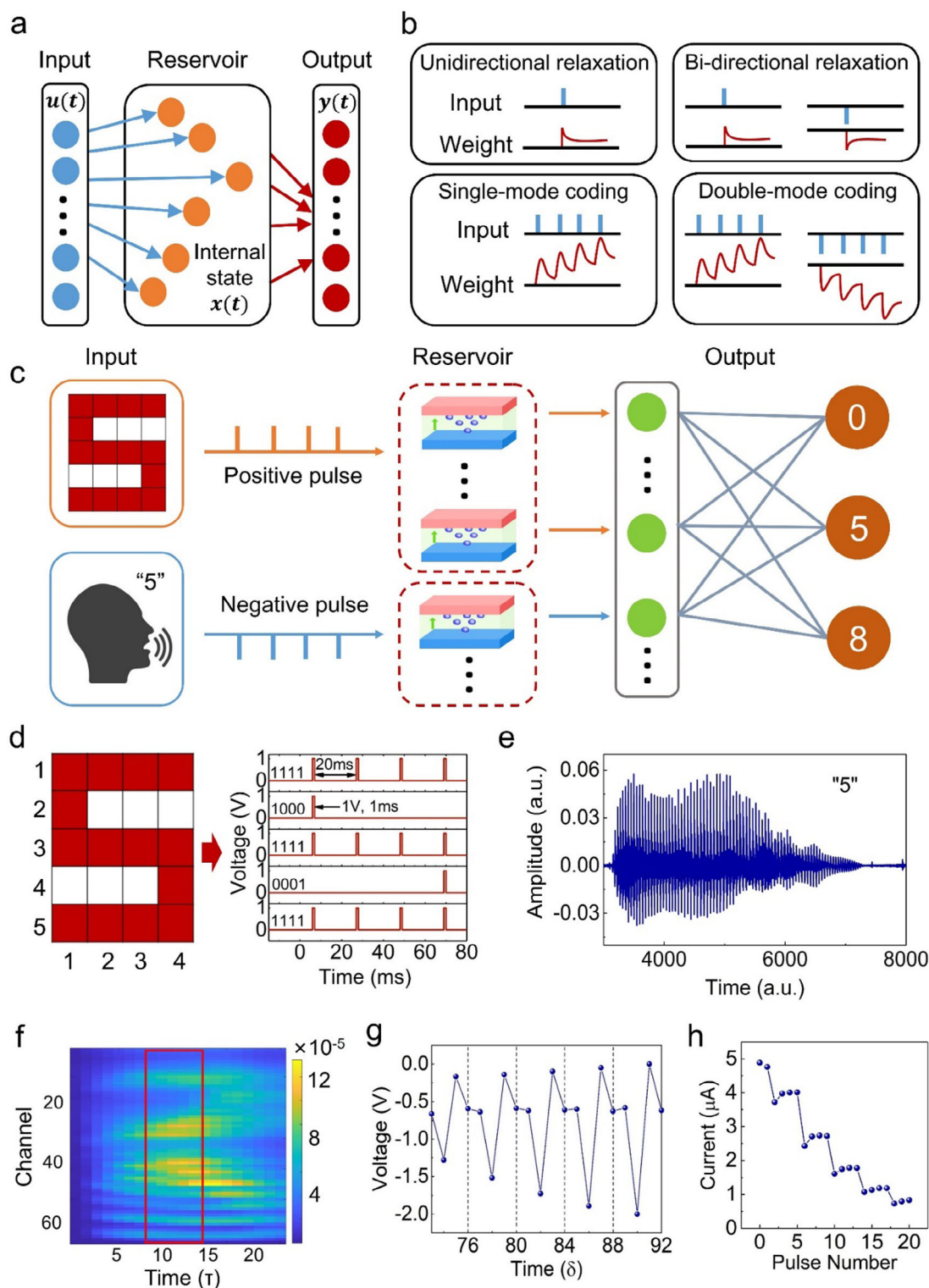
Figure 4. Multimodal recognition based on image and speech patterns. a) Schematic of an RC system. b) Schematics of unidirectional and bi-directional relaxation synapses, which are used for single-mode and double-mode signal coding, respectively. c) The schematic of the multimodal digit recognition system, where the positive pulse streams are used to encode image information and the negative pulse streams are employed to encode voice information. d) The pulse scheme and the reservoir states of image "5". e) A typical audio waveform of voice "5". f) The cochlea-gram converted from the original audio waveform using Lyon's passive ear model. It has 64 frequency channels in each frame. 5 frames in the middle of the cochlea-gram are retained to produce the input negative voltage sequence. g) The input negative voltage and h) current response of the dynamic memristor to the input voltage.
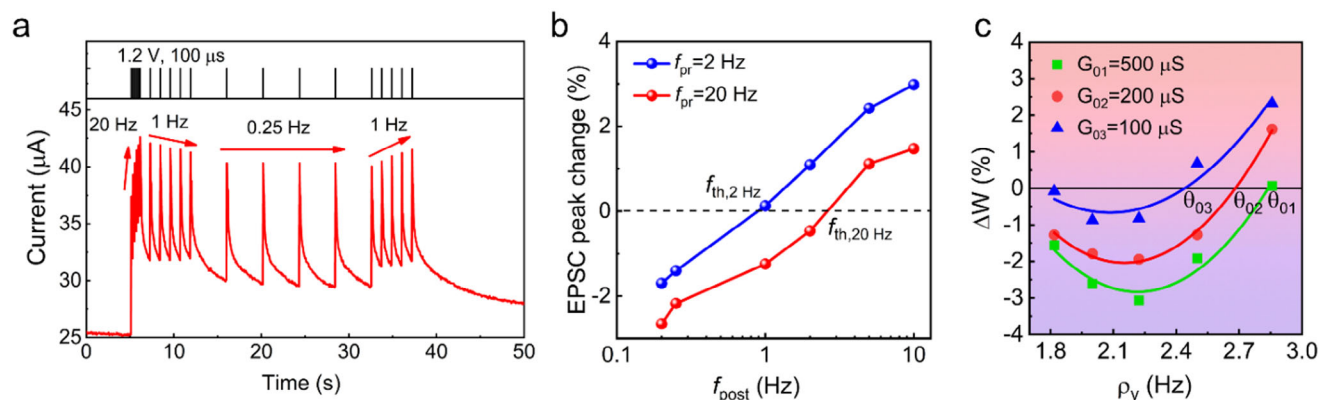
**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
MATERIALS**

www.advmat.de

Figure 5. Adaptive characteristics and BCM learning rule. a) EPSC in response to consecutive pulse trains with different frequencies (20Hz → 1Hz → 0.25Hz → 1Hz). b) Influence of a prior pulse train with a different frequency f_{pr} on the weight changes: a larger f_{pr} causes a higher f_{th}. Both a) and b) demonstrate frequency-dependent history adaptation. c) The BCM learning rule extracted from the triplet-STDP landscape with $\Delta t_1 = -150$ ms (the yellow line in Figure S30, Supporting Information) for different initial states: $G_0 = 100$, 200, and 500 μS.

plasticity to pulse polarity, which is also an important type of adaptive long-term plasticity.

Here, we take the weight changes influenced by previous input pulse frequencies as an example to further demonstrate the adaptation behavior of our memristive synapse. Voltage pulse trains with various frequencies are employed to modulate the device conductance (top panel in **Figure 5**a), and the EPSC is monitored (bottom panel in Figure 5a). The initial pulse train (20 Hz) increases EPSC, demonstrating synaptic potentiation. The EPSC is then reduced by applying a slower pulse train (1 Hz), indicating synaptic depression. Reducing the pulse train frequency further to 0.25 Hz does not cause obvious changes in the EPSC. Restoring the frequency to 1 Hz, however, causes synaptic potentiation. Although the pulse frequency in the second and fourth epochs is the same (1 Hz), the synaptic weight is modified in opposite directions. The same synapse can either be strengthened or depressed, depending on the frequency of experienced stimulations, demonstrating frequency-influenced history-adaptive characteristics. The history adaptation was also demonstrated by setting the device into two different states using prior pulse trains (1 V, 100 μs, 10 pulses) with frequencies f_{pr} of 2 and 20 Hz. The posterior pulse trains (1 V, 100 μs, five pulses) with frequencies f_{post} varying from 0.2 to 10 Hz are then employed to monitor the changes in EPSC peaks. Figure 5b shows that the low and high posterior frequency stimulations lead to synaptic depression and potentiation, respectively. The posterior frequency threshold f_{th}, where the EPSC peak change is 0, depends on the prior stimulation frequency f_{pr}: a higher f_{pr} leads to a large f_{th}.

To demonstrate the spatio-temporal learning ability enabled by our electronic synapse, we have implemented the functionalities of orientation selectivity and moving directions recognition possessed by the visual cortex in a neuromorphic visual system (**Figure 6**a). The orientation selectivity function is realized by the Bienenstock–Cooper–Munro (BCM) learning rule, which is a fundamental principle of spatiotemporal learning in the brain's cortex. Faithfully replicating BCM rule is a prerequisite for neural networks to effectively simulate the spatiotemporal learning mechanisms of the human brain. Specifically, the BCM rule describes the long-term synaptic weight changes mod-

ulated by spiking frequency and history-adaptive characteristics (Note S12, Supporting Information). Here the BCM learning rule is faithfully produced by triplet-STDP (spike-timing-dependent plasticity),[47] see details in Notes S13 and S14 (Supporting Information). The BCM curves with initial states $G_0 = 100$, 200, and 500 μS are shown in Figure 5c. The calculated BCM theoretical curves for the three different initial states (see details in Note S15, Supporting Information) are also shown in Figure 5c and fit well with the experimental data. As can be seen, ΔW transforms from depression to potentiation as the spiking frequency exceeds the threshold value θ. The frequency threshold θ varies with initial conductance states G_0, demonstrating the threshold sliding effect. Additionally, ΔW initially decreases and then increases with spiking frequency in the depression region, which is consistent with the enhanced depression effect (Figure S28, Supporting Information). Thus, the BCM learning rule is faithfully embodied in our FTJ memristor.

Then the BCM learning rule was used to perform orientation and direction recognition task.[44] As shown in Figure 6a, there are eight different patterns consisting of four bars with different orientations, each moving in one of two directions. We use neural networks with an array of FTJ memristors (Figure 6b) at the core to sequentially identify the orientations represented by these patterns and the moving directions of these orientations. The network is composed of input neurons, a synaptic array constructed with our FTJ memristors, and output neurons (Figure 6b), corresponding to the eyes, receptive fields, and cortical neurons, respectively. The 81 input neurons are mapped to one of the four output neurons via an $81 \times 4 = 324$ synapses. Each orientation pattern is organized into a $9 \times 9 = 81$ grid (left panel of Figure 6b), which is associated with 81 input neurons. The 81 pixels of orientation patterns are encoded with Poisson-distributed spike frequency signals, among which the 9 pixels corresponding to the orientation bars (yellow blocks in Figure 6b) have a high average frequency of $<\rho_x> = 2$ Hz, whereas the other 72 pixels (blue blocks in Figure 6b) show a low average frequency of $<\rho_x> = 0.7$ Hz. Prior to learning, all 324 synaptic weights were initialized to stochastic states with a low G_0. During learning, the input neurons were randomly presented
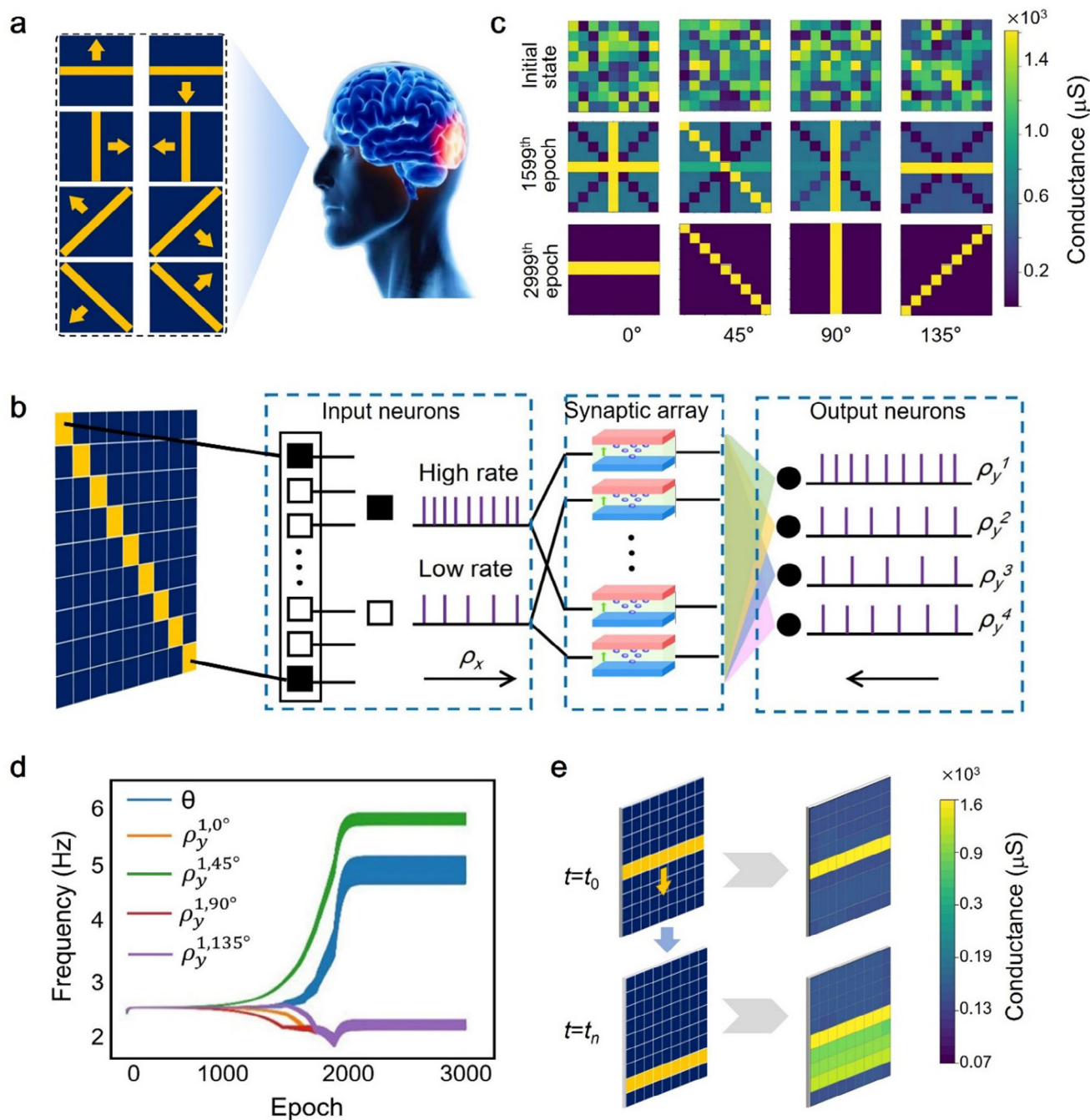
ADVANCED
SCIENCE NEWS

www.advancedsciencenews.com

ADVANCED
MATERIALS

www.advmat.de

Figure 6. Orientation and moving direction recognition. a) Four different orientation bars (0°, 90°, 45°, and 135°), each with two moving directions, were recognized by the artificial receptive fields constructed from our FTJ array. b) Schematic diagram of the feedforward neural network consisting of FTJ devices as a synaptic array. c) Evolution of the synaptic weights connected to the output neuron for different orientation bars. At the end of learning, the conductance of the devices corresponding to the pixels of the bar's orientation in the synaptic array evolved to the high conductance state. d) Dependence of the frequency ρ_y^1 and threshold frequency θ of the synapses connected to the output Neuron 1 on the number of training epochs. e) The implementation of moving direction identification task, during which the moving orientation bar are mapped to conductance matrix of the synaptic array.

with four orientation patterns with equal probability. The output neurons respond to the inputs following a function with a firing frequency of $\rho_y^n = \sum_{m=1}^{81} \rho_x^m \times W_m^n$ ($m = 1,2,...,81$, $n = 1,2,3,4$.[48,49] To select the specific neuron from the four orientations, the ρ_y^n

of the four output neurons were compared after each pattern submission. Only the neuron with maximal ρ_y^n could generate the post-synaptic spikes being fired following the winner-take-all rule, thereby modifying the corresponding 81 synaptic weights

ADVANCED
SCIENCE NEWS

www.advancedsciencenews.com

ADVANCED
MATERIALS

www.advmat.de

together with presynaptic spikes. The synaptic weights W and the thresholds θ then update following the BCM learning curves in Figure 5c, finally forming the binary weight distributions corresponding to the four orientation patterns. As a result, the output neurons are able recognize one of the orientation patterns.

The output Neuron 1 is taken as an example to demonstrate the evolution of synaptic weights W, firing frequency ρ_y^1, and threshold frequency θ with the learning epoch. At the early learning stage, the weights, threshold frequency, and firing frequencies relevant to the four orientations are indistinguishable (Figure 6c; Figure S32a, Supporting Information), without selectivity capacity. After 2000 learning epochs, a higher frequency $\rho_y^{1,45°}>\theta$ corresponding to the 45° orientation occurs, indicating that Neuron 1 became trained to recognize the 45° pattern. Meanwhile, a relatively low postsynaptic firing frequency ($\rho_y^{1,0°}$, $\rho_y^{1,90°}$, $\rho_y^{1,135°}<\theta$) was observed for the other three orientations (Figure 6d; Figure S32a, Supporting Information). Correspondingly, the synaptic weights of the 45° orientation bar increased upon training on further epochs, while the values of the other weights decayed (Figure 6c and left panels of Figure S32a, Supporting Information). Although the weights corresponding to the other three orientations were suppressed in the output of Neuron 1, they were selected by the other three output neurons, generating strong potentiation with a higher firing frequency ρ_y^n (N = 2,3,4) (Note S16, Supporting Information).

Once the orientation recognition task was completed, the moving direction identification of the orientation bars was carried out. Assuming that the orientation selectivity task is finished at time t_0, and the bar starts to move from this moment. Taking the up/down movement of the 0° orientation bar as an example, at time t_0, the conductance of the devices corresponding to this orientation bar's pixels in the synaptic array has evolved to the high weight state (Figure 6c). Then the pixels corresponding to the orientation bar are stimulated by positive pulses to illustrate its trajectory. As the orientation bar moves up or down, all the memristors encountered are subjected to positive pulses sequentially, while the memristors that were previously stimulated gradually decay to lower conductivity levels due to the relaxation effect (Figure 6e). This process allows us to represent the position of orientation bars at various times as distinct matrices of conductance values. By analyzing these matrices, the moving directions of the orientation bars can be identified.

It is noteworthy that the orientation and moving direction recognition task based on our device is significantly different from motion perception tasks achieved by optoelectronic devices.[50] Our device takes advantages of its "computing" capability by replicating the BCM learning rule of the brain. In contrast, optoelectronic devices rely on their "sensory" characteristics, while requiring additional neural network modules to complete the recognition of motion.

3. Conclusion

In summary, to faithfully replicate the multimodal perception and spatio-temporal learning capabilities of the human brain, we have proposed an FTJ-based memristive synapse in which multi-timescale plasticity is achieved by introducing a volatile $V_{\ddot{O}}$ migration mechanism into the ferroelectric layer to mod-

ulate the device conductance together with the nonvolatile polarization switching mechanism. This strategy not only generates bi-directional relaxed short-term temporal dynamics, which ensures multiple signal processing abilities, but also integrates adaptation characteristics into long-term synaptic behavior, enabling the electronic synapse to demonstrate spatio-temporal learning. A typical example of digital recognition of bimodal information comprising visual and acoustical signals was successfully realized by encoding images and speech into positive and negative electric pulses respectively. Then, an orientation and direction recognition task for spatio-temporal functions was also successfully implemented in neural networks composed of the FTJ memristor array. Our study closely replicates the computing functions of the human brain with a purely electrically modulated synaptic emulator, demonstrating the great potential of neuromorphic devices that integrate synergetic physical mechanisms achieving bimodal modulation on multiple timescales.

4. Experimental Section

Device Fabrication: Ultrathin BaTiO$_3$ (BTO) films with a thickness of 1.6 nm (4 u.c.) were epitaxially grown by pulsed laser deposition (PLD) on (001)-oriented 0.7 wt.% Nb-doped SrTiO$_3$ substrates using a KrF (λ = 248 nm) excimer laser with an energy density of 3.0 J cm^{-2} and a repetition rate of 2 Hz. BTO films were deposited at 780 °C under continuous flow of oxygen at a low pressure of 0.02 Torr followed by cool down to room temperature within 30 min. The artificial synapse with an electrode area of 30 × 30 µm^2 was defined by direct laser writing and evaporated Cr (5 nm thickness)/Au (50 nm thickness).

Structure Characterization: Ferroelectric properties were measured by means of piezo-response force microscopy (PFM; MFP-3D-Origion$^+$ Asylum Research) in the DART (dual a.c. resonance tracking) mode with triangular waveforms. The chemical states and the concentration of oxygen vacancies in the BTO films were evaluated using the ESCALAB XI+ X-ray photoelectron spectrometer (XPS, Thermo Fisher Scientific) with monochromatized Al K$_\alpha$X-ray radiation (1486.6 eV). The phase structures and crystallographic orientations of BaTiO$_3$ films were studied by X-ray diffraction using Cu K$_\alpha$ radiation (XRD, Smartlab3kw, Rigaku).

Electrical Measurements: The Au/Cr/BTO/NSTO FTJs were placed in an ESP4 probe station with a 1 µm tungsten steel gold-plated probe, and the probe station was connected to the electrical measurement equipment. Electrical characterizations were undertaken using a Keithley 2410 Source Meter with homemade programs and an ArC ONE memristor characterization platform.

Neural Network Simulations: The output layer of the RC system is an artificial neural network (ANN) simulated in Matlab 2022a. It is a single-layer feedforward neural network that uses Delta rule along with the tanh activation function for weight updates to perform classification. The ANN used to achieve the image consists of 5 input neurons, 3 output neurons, and 5 × 3 = 15 synapses. The ANN used to achieve speech recognition consists of 20 input neurons, 3 output neurons, and 20 × 3 = 60 synapses. The ANN used to achieve multimodal digital information consists of 6 input neurons, 3 output neurons, and 6 × 3 = 18 synapses. The following method was used to prepare the digital image, digital speech, and multimodal digital datasets: First, a set of 5 measured reservoir states were used to represent a single digital image. 30 such sets were obtained from different devices to formed the ideal digital image dataset containing 30 samples. Then, a set of 20 measured reservoir states were used to represent a single digital voice. Similarly, 30 sets were obtained from different devices formed the ideal digital speech dataset containing 30 samples. Next, a corrupted image sample was prepared by using the first 2 reservoir states out of total 5 states of an image. For the corrupted digital speech sample, 4 reservoir states were randomly selected from the 20 of an sample. Finally, the corrupted image reservoir states were combined

with those of the corrupted digital speech to create the multimodal digital dataset, which consists of 5400 samples, with 900 used for training the neural network and 4500 used for testing. The neural network for performing rate-based orientation selectivity is a two-layer feedforward neural network simulated with the spiking neural network simulator Brain2. It is composed of 81 input neurons, 4 postsynaptic neurons, and 324 synapses. The neural network can realize orientation recognition with the BCM learning rule.

Ethical Statement: The experiments did not involve animals or human biological samples.

Supporting Information

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

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

F.N., H.F., and J.W. contributed equally to this work. L.Z. and L.M.Z. conceived the idea and project and conducted the experimental design. Then F.N., H.F., and J.W. prepared the sample; F.N. H.F., J.W., and S.T.Y. performed the electrical and PFM measurements; S.E.M., F.Y.W., and W.B.Z. simulated the output neural network of the RC system in Matlab; C.J., S.Z.W., and S.L. simulated the rate-based orientation selection in *Brain2*; C.G., S.S.Y., and A.N. carried out result analysis and provided important suggestions during the preparation of the manuscript; L.Z., L.M.Z, S.S.Y., and A.N. wrote the manuscript.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

[1] J. C. Magee, C. Grienberger, *Annu. Rev. Neurosci.* **2020**, *43*, 95.
[2] T. Takeuchi, A. J. Duszkiewicz, R. G. M. Morris, *Philos. Trans. R. Soc. B-Biol. Sci.* **2014**, *369*, 20130288.
[3] S. J. Martin, R. G. M. Morris, *Hippocampus* **2002**, *12*, 609.
[4] A. Citri, R. C. Malenka, *Neuropsychopharmacology* **2008**, *33*, 18.
[5] E. S. Fortune, G. J. Rose, *Trends Neurosci.* **2001**, *24*, 381.
[6] A. Kohn, *J. Neurophysiol.* **2007**, *97*, 3155.
[7] C. A. Patterson, S. C. Wissig, A. Kohn, *J. Neurosci.* **2013**, *33*, 532.
[8] Z. Wang, T. Zeng, Y. Ren, Y. Lin, H. Xu, X. Zhao, Y. Liu, D. Ielmini, *Nat. Commun.* **2020**, *11*, 1510.
[9] X. Yan, J. H. Qian, V. K. Sangwan, M. C. Hersam, *Adv. Mater.* **2022**, *34*, 2108025.
[10] C. Li, D.-L. Yang, L.-F. Sun, *Acta Phys. Sin.* **2022**, *71*, 218504.
[11] G. Zhou, Z. Wang, B. Sun, F. Zhou, L. Sun, H. Zhao, X. Hu, X. Peng, J. Yan, H. Wang, W. Wang, J. Li, B. Yan, D. Kuang, Y. Wang, L. Wang, S. Duan, *Adv. Electron. Mater.* **2022**, *8*, 2101127.
[12] R. Wang, J.-Q. Yang, J.-Y. Mao, Z.-P. Wang, S. Wu, M. Zhou, T. Chen, Y. Zhou, S.-T. Han, *Adv. Intell. Syst.* **2020**, *2*, 2000055.
[13] H.-M. Huang, Z. Wang, T. Wang, Y. Xiao, X. Guo, *Adv. Intell. Syst.* **2020**, *2*, 2000149.
[14] S. Boyn, J. Grollier, G. Lecerf, B. Xu, N. Locatelli, S. Fusil, S. Girod, C. Carrétéro, K. Garcia, S. Xavier, J. Tomas, L. Bellaiche, M. Bibes, A. Barthélémy, S. Saïghi, V. Garcia, *Nat. Commun.* **2017**, *8*, 14736.
[15] D. Kim, B. Jeon, Y. Lee, D. Kim, Y. Cho, S. Kim, *Appl. Phys. Lett.* **2022**, *121*, 10501.
[16] G. Zhang, J. Qin, Y. Zhang, G. Gong, Z. Xiong, X. Ma, Z. Lv, Y. Zhou, S. Han, *Adv. Funct. Mater.* **2023**, *33*, 2302929.
[17] J. Yu, Y. Wang, S. Qin, G. Gao, C. Xu, Z. Lin Wang, Q. Sun, *Mater. Today* **2022**, *60*, 158.
[18] Z. Guo, J. Zhang, X. Liu, L. Wang, L. Xiong, J. Huang, *Adv. Funct. Mater.* **2023**, *33*, 2305508.
[19] D. Hao, Z. Yang, J. Huang, F. Shan, *Adv. Funct. Mater.* **2023**, *33*, 2211467.
[20] K. Liu, T. Zhang, B. Dang, L. Bao, L. Xu, C. Cheng, Z. Yang, R. Huang, Y. Yang, *Nat. Electron.* **2022**, *5*, 761.
[21] 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. Barthélémy, J. Grollier, *Nat. Mater.* **2012**, *11*, 860.
[22] D. J. Kim, H. Lu, S. Ryu, C.-W. Bark, C.-B. Eom, E. Y. Tsymbal, A. Gruverman, *Nano Lett.* **2012**, *12*, 5697.
[23] K. Ye, K. Li, Y. Lu, Z. Guo, N. Ni, H. Liu, Y. Huang, H. Ji, P. Wang, *TRAC Trends Anal. Chem.* **2019**, *116*, 102.
[24] M. C. Biesinger, L. W. M. Lau, A. R. Gerson, R. St. C. Smart, *Appl. Surf. Sci.* **2010**, *257*, 887.
[25] Y. Zhu, L. Zhang, B. Zhao, H. Chen, X. Liu, R. Zhao, X. Wang, J. Liu, Y. Chen, M. Liu, *Adv. Funct. Mater.* **2019**, *29*, 1901783.
[26] X. Liu, L. Zhang, Y. Zheng, Z. Guo, Y. Zhu, H. Chen, F. Li, P. Liu, B. Yu, X. Wang, J. Liu, Y. Chen, M. Liu, *Adv. Sci.* **2019**, *6*, 1801898.
[27] J. Li, N. Li, C. Ge, H. Huang, Y. Sun, P. Gao, M. He, C. Wang, G. Yang, K. Jin, *iScience* **2019**, *16*, 368.
[28] S. Yang, X. Li, T. Yu, J. Wang, H. Fang, F. Nie, B. He, L. Zhao, W. Lü, S. Yan, A. Nogaret, G. Liu, L. Zheng, *Adv. Funct. Mater.* **2022**, *32*, 2202366.
[29] Y. Li, J. Chu, W. Duan, G. Cai, X. Fan, X. Wang, G. Wang, Y. Pei, *ACS Appl. Mater. Interfaces* **2018**, *10*, 24598.
[30] A. Gramm, T. Zahner, U. Spreitzer, R. Rössler, J. D. Pedarnig, D. Bäuerle, H. Lengfellner, *Europhys. Lett. EPL* **2000**, *49*, 501.
[31] L. Zheng, L. Yang, Y. Li, X. Lu, D. Huo, W. Lü, R. Zhang, B. Yang, W. Cao, *Phys. Rev. Appl.* **2018**, *9*, 64028.
[32] Z. Q. Wang, H. Y. Xu, X. H. Li, H. Yu, Y. C. Liu, X. J. Zhu, *Adv. Funct. Mater.* **2012**, *22*, 2759.
[33] Z. Wen, C. Li, D. Wu, A. Li, N. Ming, *Nat. Mater.* **2013**, *12*, 617.

[34] Z. Xi, J. Ruan, C. Li, C. Zheng, Z. Wen, J. Dai, A. Li, D. Wu, *Nat. Commun.* **2017**, *8*, 15217.

[35] N. Siannas, C. Zacharaki, P. Tsipas, D. J. Kim, W. Hamouda, C. Istrate, L. Pintilie, M. Schmidbauer, C. Dubourdieu, A. Dimoulas, *Adv. Funct. Mater.* **2024**, *34*, 2311767.

[36] Z. Chen, W. Li, Z. Fan, S. Dong, Y. Chen, M. Qin, M. Zeng, X. Lu, G. Zhou, X. Gao, J.-M. Liu, *Nat. Commun.* **2023**, *14*, 3585.

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

[38] J. Gong, H. Yu, X. Zhou, H. Wei, M. Ma, H. Han, S. Zhang, Y. Ni, Y. Li, W. Xu, *Adv. Funct. Mater.* **2020**, *30*, 2005413.

[39] J. Li, C. Ge, J. Du, C. Wang, G. Yang, K. Jin, *Adv. Mater.* **2020**, *32*, 1905764.

[40] P. S. Subin, A. S. Asha, K. J. Saji, M. K. Jayaraj, *J. Mater. Sci. Mater. Electron.* **2021**, *32*, 13051.

[41] C. Du, F. Cai, M. A. Zidan, W. Ma, S. H. Lee, W. D. Lu, *Nat. Commun.* **2017**, *8*, 2204.

[42] Y. Zhong, J. Tang, X. Li, B. Gao, H. Qian, H. Wu, *Nat. Commun.* **2021**, *12*, 408.

[43] R. Lyon, in *ICASSP 82 IEEE Int. Conf. Acoust. Speech Signal Process*, Institute of Electrical and Electronics Engineers, Paris, France, **1982**, pp. 1282–1285.

[44] Y. Zhai, P. Xie, Z. Feng, C. Du, S. Han, Y. Zhou, *Adv. Funct. Mater.* **2022**, *32*, 2108440.

[45] Z. Tan, R. Yang, K. Terabe, X. Yin, X. Zhang, X. Guo, *Adv. Mater.* **2016**, *28*, 377.

[46] Q. Wu, H. Wang, Q. Luo, W. Banerjee, J. Cao, X. Zhang, F. Wu, Q. Liu, L. Li, M. Liu, *Nanoscale* **2018**, *10*, 5875.

[47] R. A. John, A. Milozzi, S. Tsarev, R. Brönnimann, S. C. Boehme, E. Wu, I. Shorubalko, M. V. Kovalenko, D. Ielmini, *Sci. Adv.* **2022**, *8*, eade0072.

[48] M. Lawlor, S. W. Zucker, in *Conf. Neural Inf. Process. Syst.*, MIT Press, Cambridge, MA, USA **2014**, pp. 2564–2572.

[49] L. C. Udeigwe, P. W. Munro, G. B. Ermentrout, *J. Math. Neurosci.* **2017**, *7*, 2.

[50] J. Chen, Z. Zhou, B. J. Kim, Y. Zhou, Z. Wang, T. Wan, J. Yan, J. Kang, J.-H. Ahn, Y. Chai, *Nat. Nanotechnol.* **2023**, *18*, 882.

2412006 (12 of 12)