npj | 2D Materials and Applications www.nature.com/npj2dmaterials

ARTICLE OPEN

Graphene–ferroelectric transistors as complementary synapses for supervised learning in spiking neural network

Yangyang Chen[1], Yue Zhou[1], Fuwei Zhuge[2], Bobo Tian[3], Mengge Yan[3], Yi Li[1], Yuhui He[1] and Xiang Shui Miao[1]

The hardware design of supervised learning (SL) in spiking neural network (SNN) prefers 3-terminal memristive synapses, where the third terminal is used to impose supervise signals. In this work we address this demand by fabricating graphene transistor gated through organic ferroelectrics of polyvinylidene fluoride. Through gate tuning not only is the nonvolatile and continuous change of graphene channel conductance demonstrated, but also the transition between electron-dominated and hole-dominated transport. By exploiting the adjustable bipolar characteristic, the graphene–ferroelectric transistor can be electrically reconfigured as potentiative or depressive synapse and in this way complementary synapses are realized. The complementary synapse and neuron circuit is then constructed to execute remote supervise method (ReSuMe) of SNN, and quick convergence to successful learning is found through network-level simulation when applying to a SL task of classifying 3×3-pixel images. The presented design of graphene–ferroelectric transistor-based complementary synapses and quantitative simulation may indicate a potential approach to hardware implementation of SL in SNN.

npj 2D Materials and Applications (2019)3:31 ; https://doi.org/10.1038/s41699-019-0114-6

INTRODUCTION

By mimicking the plasticity of brain, neuromorphic computing is capable of self-learning, while with revolutionary speed and energy efficiency, and is thus regarded as a promising candidate to next generation computing.[1] At hardware level, it calls for materials and devices that emulate the nonvolatile modulation of synaptic strengths.[2–5] Currently, two-terminal memristors made from various kinds of materials such as resistive random access memory,[6] phase-change memory,[7] and conducting bridge random access memory,[8] etc. are intensely studied as electronic synapses for artificial neural networks.[9–11] In a typical supervised learning (SL) task, the two-terminal memristors implement algorithms with iterative read-and-write operations: during the forward step the outputs are obtained through multiplying voltages from input neurons by the conductance of memristive synapses (read), while during the update step the conductance of memristors is delicately tuned in order to minimize the error between real outputs and the desired ones (write). In this way, the outputs are calibrated to the targets through the SL process. By further adopting structures such as 1-transistor-1-memristor[12] and 1-selector-1-memristor,[13,14] the network-level computing is substantially facilitated by removing the sneak paths in the synaptic arrays, and excellent performances on face recognition[12] and handwritten digit classification[6] have been reported.

On the other side, a spike-based computing paradigm namely spiking neural network (SNN) has emerged as the third generation neural network.[15] Since it is more similar to the operation of biological brains, SNN shares the advantages of real brains, such as ultralow power consumption and larger processing capacity. From the viewpoint of hardware, 3-terminal nonvolatile transistors that accommodate direct feedback modulation to synapse weight are desired.[16,17] As seen in Fig. 1a, spike signals transmitting from drain to source of the transistor mimics that from presynaptic neuron to postsynaptic one in biological systems. Once neurons in the output layer fire, the error message is generated by comparing the timings of actual outputs and desired ones. The required amounts of synapse strength modulation, represented by the channel conductance changes, are then calculated through those auxiliary modules in SNN circuits and the corresponding conductance tuning is implemented by the feedback gate voltage. In this regard, memristive transistors such as 3-terminal ferroelectric memristor[17] and organic ferroelectric synapses[18] have been proposed, whereas further development on the device materials, functions, and the related implementation of advanced algorithms are called for.

In order to address the above demands, we fabricate field effect transistors using graphene as the channels and P(VDF-TrFE) ferroelectric polymer as the gate dielectric (graphene–ferroelectric field effect transistor, abbreviated as GrFeFET) as seen in Fig. 1b. The graphene channel serves as the synapse connecting the pre- and postsynaptic neurons, while the gate terminal accepts supervise signals and modulates the channel conductance. Here the gate tuning to channel conductance is nonvolatile due to the remnant of ferroelectric polarization in the polyvinylidene fluoride (PVDF) layer (100 nm thick), which then emulates the plastic changes of synapse strengths.[19] Figure 1c illustrates a 3-dimensional optical view of our device with false color of the source, drain, top electrodes, and graphene channel (5 μm long and 10 μm wide). Such GrFeFETs were previously explored as nonvolatile memory device,[20,21] while in the current work we will

[1]Wuhan National Research Center for Optoelectronics, School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, China; [2]State Key Laboratory of Material Processing and Die & Mould Technology, School of Materials Science and Engineering, Huazhong University of Science and Technology, Wuhan, China and [3]Key Laboratory of Polar Materials and Devices, Ministry of Education, Department of Optoelectronic Science and Engineering, East China Normal University, Shanghai, China
Correspondence: Yuhui He (heyuhui@hust.edu.cn) or Xiang Shui Miao (miaoxs@hust.edu.cn)
These authors contributed equally: Yangyang Chen, Yue Zhou, Fuwei Zhuge

Received: 15 January 2019 Accepted: 31 July 2019
Published online: 21 August 2019

npj nature partner journals

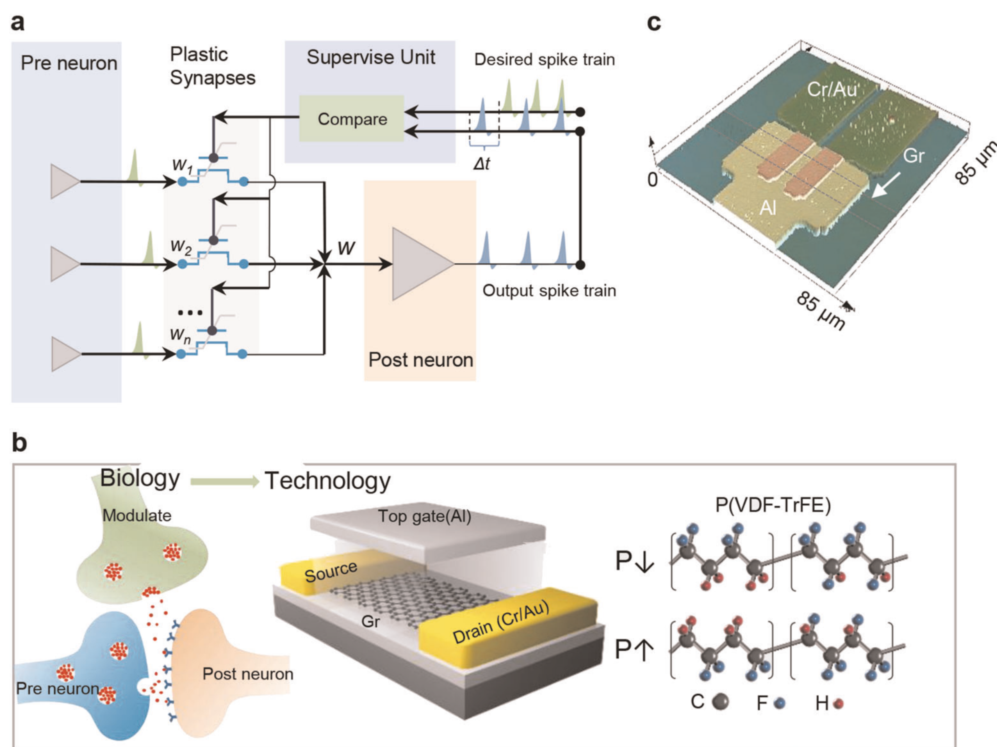

Fig. 1 Schematic view of **a** 3-terminal memristive synapses for supervised learning (SL) in spiking neural network (SNN), where **b** the field effect transistors with graphene as channels and organic ferroelectric polyvinylidene fluoride (PVDF) as gate dielectric (GrFeFET) mimic the synaptic functions. The source and drain terminals serve as the axon of the presynaptic neuron and the dendrite of the postsynaptic one, respectively, while the tuning of the channel conductance by the PVDF polarization emulates the modulation of synaptic strength. **c** Optical profiler image of the fabricated GrFeFET, where fake color by Photoshop is applied. The two brown colored pads are the source and drain regions, while the graphene region is characterized by the dash lines. A top view optical image of the fabricated device is provided in Section 1 of Supplementary Information. Dimensions of graphene channel are 5 μm long and 10 μm wide, and PVDF is 100 nm thick

demonstrate its unique potential as complementary synapses and usage in SL of SNN.

RESULTS AND DISCUSSION

Figure 2a shows the channel conductance modulated by back gate voltage $G(V_{bg})$ through 300 nm SiO_2. The left and right branches denote the hole- and electron-dominated transport, respectively. The electron and hole mobility is extracted to be $\mu = 1.7 \times 10^3$ cm^2/V s by taking the permittivity of back gate dielectric SiO_2 as $\kappa_{SiO2} = 3.9$ (estimation details are provided in the "Method" section).[21] In a back-gate sweep loop from −40 to 40 V and reversely to −40 V, the hysteresis-free conductance change in graphene indicated well-suppressed interface traps in transistor. It should be noted that the back gate does not cause memorable modulation of the channel conductance therefore not suitable for artificial synapses. The physical mechanism is that the polarization states in PVDF could hardly be changed by the back-gate voltage as the applied electric field is mostly screened by the graphene channel. On the other hand, in Fig. 2b, an obvious hysteresis window appears during the sweep of top gate voltage (V_{tg}) between −20 and 20 V due to the tuning of polarization in PVDF. It is known that the dielectric polarization of PVDF ferroelectric by gate voltage exhibits both the linear response that is proportional to external polarization field and a nonlinear component known as residual polarization when the external field is removed.[21] Depending on the upward or downward polarization in PVDF, the graphene channel was hole doped or electron doped, causing a nonvolatile shift of Dirac point. We further extract $p(E)$ curve of PVDF from Fig. 2b by adopting the theoretical description of ferroelectric FET provided in the "Method" section, and compare it with the direct electric displacement measurement of PVDF film with the same thickness alone, $D'(E)$. As shown in Fig. 2c, D and D' have very similar coercive fields $E_C \sim 50$ MV/m, which are consistent with values previously reported on PVDF.[20] Moreover, the capacitance–voltage(C–V) relationship measured from Au/PVDF/Al structure also indicates coercive fields $E_C \sim \pm 50$ MV/m as shown in Fig. 2d. These quantitative agreements strongly suggest that the hysteresis observed in the transport measurements is indeed caused by the hysteretic polarization of the ferroelectric gate dielectric. The residual polarization is estimated to be $\sim \pm 1$ μC/cm^2. It is worth reminding that compared with other tuning approaches of graphene channel conductance,[22,23] the nonvolatile nature of ferroelectric polarization causes persistent and memorable modulation of the graphene channel conductance, thus providing a feasible approach toward a 3-terminal synapse.

Figure 3a illustrates the basic principle of tuning graphene channel to be hole- or electron-conduction through the different polarization of ferroelectric layer. A negative/positive gate voltage with large amplitude over the coercive voltage will result in upward/downward polarization of the PVDF dielectric, which causes hole/electron doping of graphene channel, respectively. From the viewpoint of energy band filling, the graphene channel becomes p- or n-type conduction depending on the polarization direction of PVDF dielectric, as indicated by the right subfigures shown in Fig. 3a. Moreover, given the same positive voltage pulses, while with small magnitudes (+15 V in Fig. 3b and +8 V in Fig. 3c), the upward (downward) PVDF polarization will be decreased (increased), while the Fermi levels will shift upward in both p- and n-type graphene channels. However, the ascending of Fermi levels in p- and n-type graphene channel will lead to different results: for the former it is a reduction of the hole density, while for the latter an enhancement of electron density near the Fermi level. In other words, the conductance tuning of GrFeFET

Published in partnership with FCT NOVA with the support of E-MRS

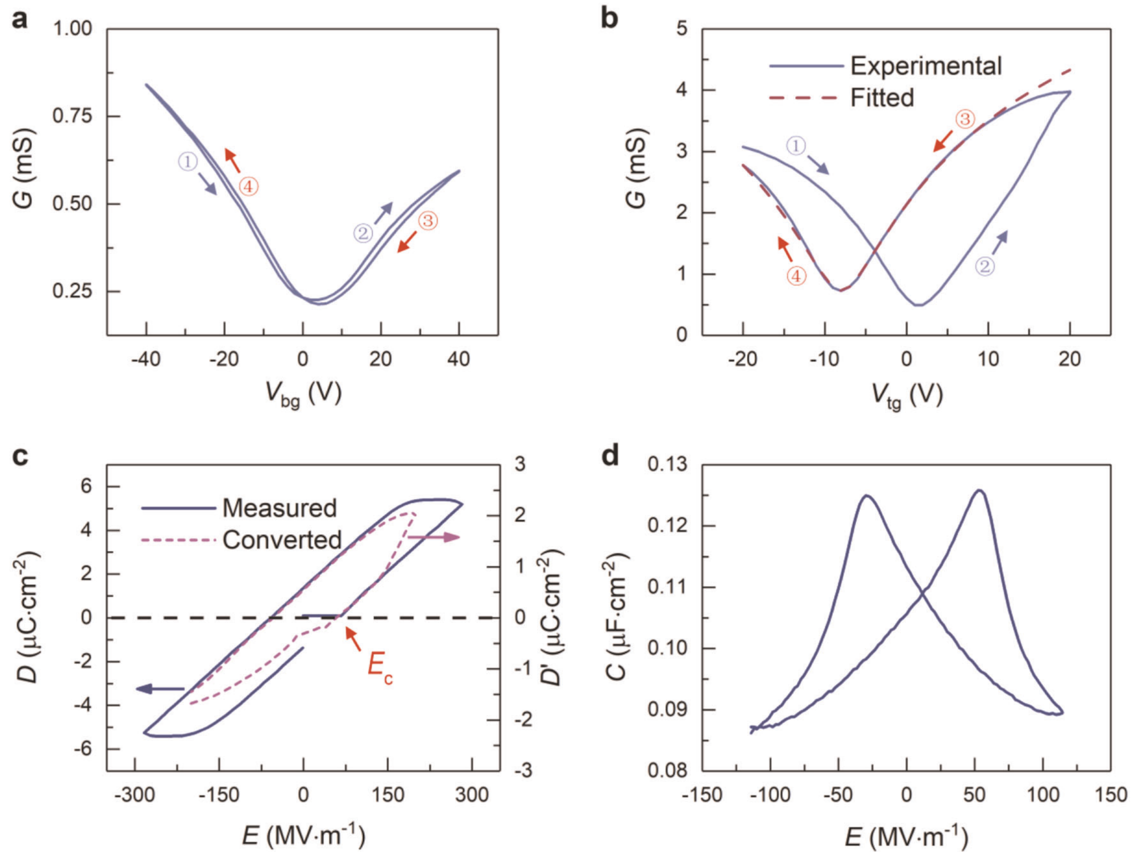

Fig. 2 Electrical properties of GrFeFET. **a** The measured source-drain conductance versus back gate voltage $G_{ds}(V_{bg})$, where the left and right branches correspond to hole- and electron-dominated transport respectively. **b** Electric hysteresis loop by sweeping top gate voltage $G_{ds}(V_{tg})$. **c** The electric displacement versus the applied electric field $D(E)$ deduced from **b**, while that obtained directly from ferroelectric measurement of PVDF is plotted with red line. **d** The measured capacitance versus the applied voltage $C(V)$ for single PVDF film with 100 nm thickness

based synapses can be depressive or potentiative depending on the initial status of graphene energy band filling and this filling is adjustable through opposite polarization of gate ferroelectric. It explains why given similar small amplitudes of positive voltage pulses on the gate, the variation trends of the measured p- and n-type graphene channel conductance become opposite as shown in Fig. 3b, c. Similar analysis can be conducted on the measurement of negative voltage pulses imposed to the gates of p- and n-type GrFeFET (−10 V in Fig. 3b and −6 V in Fig. 3c). In this way, the analog weight update of GrFeFET synapse is successfully realized as plotted in Fig. 3b, c. Compared with previous reports,[2,12] a distinct feature here is that in our GrFeFET synapse the synaptic weight update can be switched to be potentiative or depressive depending on the conduction type of channel given SET/RESET voltage pulses.

Here we point out that the synapses with positive/negative changes of weights ($\Delta w > 0$ or $\Delta w < 0$) under ordinary SET pulses should be defined as potentiative/depressive rather than as excitatory/inhibitive,[22] since in neuroscience excitatory/inhibitive synapses mean positive/negative weights ($w > 0$ or $w < 0$).[24] We further stress that the realization of both potentiative and depressive, i.e., complementary synapses, with the same material and device structures is credited to the unique characteristic of zero bandgap of graphene. Although it is usually regarded as a negative factor when trying to manage power consumption in the related devices, here the zero-bandgap feature plays a constructive role in achieving the complementary synapses. Only with this ultrasmall bandgap could a practical gate voltage tune the transition between electron- and hole-dominated conduction through ferroelectric polarization. As further illustrated in Fig. 3b and c, the electron or hole filling of graphene channel will be enhanced or reduced oppositely given similar variation of PVDF polarization field caused by gate tuning. In this way, the analog weight update of the corresponding synapse can be potentiative or depressive when imposing similar programming gate voltages. Just as the importance of complementary metal-oxide-semiconductor field effect transistors in integrated circuit design, the demonstrated GrFeFET-based complementary synapses may find promising usage in the future hardware neural network design.

We further evaluate the nonideal factors of the demonstrated analog weight update of GrFeFET synapse by formulas provided in "Method" section and the obtained quantities are listed in Table 1. They are at the same levels with those recently reported in other kinds of ferroelectric synapses.[25] A conventional convolution neural network (CNN) is then set up where both the convolutional kernels and the connections in the fully connected layers are implemented with the GrFeFET synapses. By taking these nonideal parameters into account, the simulation yields a recognition rate of 94%, when implementing MNIST tasks (details are provided in Section "GrFeFET synapse in CNN for MNIST recognition" of Supplementary Information).

In order to implement SL of SNN, we further measure the conductance tuning of GrFeFET under different amplitudes of gate voltages and initial channel conductance $\Delta G(G_0, V)$, and results are plotted in Fig. 3d, e. The directions of conductance tuning become opposite in p- and n-type GrFeFET synapses, just as expected. Moreover, saturation behaviors are observed when trying to further increase G in the presence of already large G_0 or to decrease G given small G_0. It is largely ascribed to the saturation of

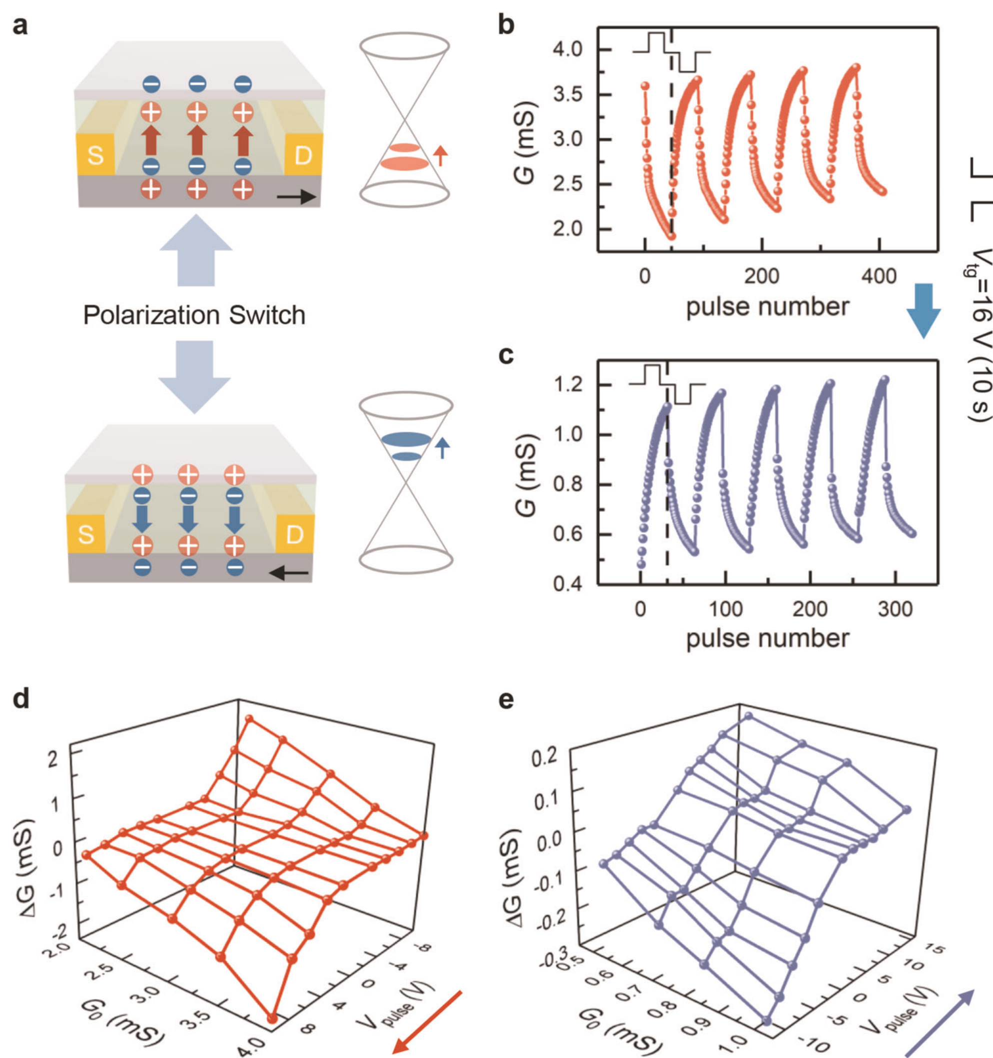

Fig. 3 Complementary synapses by tuning one GrFeFET to operate at different conduction regions. **a** Left: schematic view of tuning GrFeFET to be potentiative or depressive synapses. Upon imposing a large negative/positive gate voltage, the ferroelectric layer is polarized in upward/downward direction. Consequently, the graphene channel becomes hole/electron dominated due to different positions of Fermi levels within the graphene energy bands. Right: given the same series of positive voltage pulses on the gate, the channel conductance will be decreased/increased due to reduction/enhancement of hole/electron density caused by the corresponding change of ferroelectric polarization. **b**, **c** Analog weight update of one GrFeFET based depressive/potentiative synapses. The continuous decrease/increase of channel conductance G caused by a series of 50 (or 30) positive gate voltage pulses followed by another series of 50 (or 30) negative pulses. Here pulse width $\Delta t = 100$ ms, the source-drain voltage $V_{DS} = 0.1$ V, and six SET/RESET cycles are demonstrated. Moreover, a gate voltage with height 16 V and duration 10 s is capable of turning the hole domination to be electron domination. **d**, **e** The channel conductance change ΔG of depressive/potentiative synapses versus the magnitudes of imposed gate voltage V_{tg} and the initial conductance G_0 (pulse width $\Delta t = 500$ ms)

Table 1. The nonlinear parameters of long-term potentiation/depression (LTP/LTD), asymmetry between LTP and LTD, and cycle-to-cycle variations of depressive and potentiative GrFeFET synapses

	Nonlinear parameter		Asymmetric nonlinearity factor	Cycle-to-cycle variation	
Depressive synapse	LTP	4.17	0.11	LTP	0.035
	LTD	4.64		LTD	0.063
Potentiative synapse	LTP	2.87	0.5	LTP	0.061
	LTD	4.70		LTD	0.027

ferroelectric polarization under the applied voltage, as seen in Fig. 2. Note that the widths of imposed voltage pulses in Fig. 3d, e are different from those already shown in Fig. 3b and c, since they are set for different computing schemes of SL. The latter has been applied to *level-based* computing as seen in Section "GrFeFET synapse in CNN for MNIST recognition" of Supplementary Information, while the former will be used for *spike-based* computing[26] (detailed discussions are provided in Section "Selection of write pulse widths when training with different computing schemes" of Supplementary Information).

Besides, it is worth reminding that the potentiative and depressive behaviors of the device are not symmetric. The figure illustrates that in the hole-conduction dominated depressive region (Fig. 3b, d) the conductance and its tuning range are about two times larger than those in the electron dominated potentiative one (Fig. 3c, e). It is ascribed to the fact that graphene material is usually p-doped in the natural environment. For depressive region, it is straightforward to tune the hole conduction. On the other hand, for potentiative counterpart, first of all, a top gate voltage with quite large amplitude (~16 V) has to be imposed to induce sufficient change of PVDF ferroelectric polarization so that

electrons are attracted to, while holes are repelled from the graphene channel. After the transition of graphene channel from hole conduction to electron one, additional voltage pulses are applied to further modulate the PVDF polarization and hence induce continuous conductance changes. However, saturation of PVDF ferroelectric polarization is easily met in this situation since it has already been changed significantly during the hole-to-electron transition. Such asymmetry between hole and electron conductance tuning has also been reported in other GrFeFET devices.[22,27] As we will see, it poses challenge to the implementation of our complementary synapses in SNN tasks.

Figure 3 further demonstrates that a conductance ON/OFF ratio about 3.2 has been realized in GrFeFET device. When using as synapses, the corresponding relative weight change $\Delta w/w_0$ is about 220%, which is significantly larger than other graphene-channel based synaptic devices, where $\Delta w/w_0 \approx 12.5\%$[22] or 35%.[28] Although it benefits the learning efficiency of neural network greatly,[26] the relatively large conductance ON/OFF ratio is at the expense of using write voltage pulses with widths about hundreds or tens of milliseconds. Physically, it is ascribed to the remarkable OFF state conductance g_{OFF} caused by the zero bandgap and p-dope nature of graphene. Comparing with other memristive materials with quite small g_{OFF},[18] here a large ON state conductance g_{ON} is required to obtain the target ratio $g_{\mathrm{ON}}/g_{\mathrm{OFF}}$ in the graphene-channel devices. In order to achieve the large conductance tuning $\Delta g = g_{\mathrm{ON}} - g_{\mathrm{OFF}}$, the gate voltages with amplitudes or widths several orders larger are necessitated. It explains the 6-order slower operation speed comparing with those found in the fastest memristive synapses.[12,29] We remind that this is a common problem met by graphene-based synaptic devices,[18,22] rather than a specific issue raised by the design of complementary synapses in this work. In order to promote the operation speed, a compromise has to be made with the conductance ON/OFF ratio and further strategies are called for in the future research.

In the following, by exploiting the above conductance tuning properties of GrFeFET, an approach of using complementary GrFeFETs as synapses to implement the ReSuMe[30] is proposed as illustrated in Fig. 4a. Here ReSuMe is a widely used strategy to realize SL in SNN, since it does not resort to the conventional stochastic gradient descent method as employed by Spike Back Propagation (SpikeProp)[17,31] or the Widrow–Hoff rule as by Spike Pattern Association Neuron .[32,33] Instead, it uses window function to drive the spike timings of output neuron to the desire ones as follows[30]

$$\frac{dw}{dt} = [S_{\mathrm{d}}(t) - S_{\mathrm{o}}(t)]\left[a + \int_0^{\infty} W(\tau)S_{\mathrm{in}}(t-\tau)d\tau\right], \tag{1}$$

where $S_x(t) = \sum_f \delta(t - t_x^{(f)})$ is the spike sequence of the supervise (desire), output or input neuron with subscript $x = \mathrm{d}$, o, or in. Here f characterizes the fth spike emitted by the x neuron, a is a bias term, and $W(\tau)$ is the window function to convolute with the input. Compared with other SL methods, ReSuMe has several prominent advantages including that it is capable of learning spike sequence rather than single spikes and it is applicable to various types of neuron models. Therefore in the work, we choose to realize ReSuMe based on our complementary GrFeFETs. Figure 4a shows that the source and drain terminals of the two parallel connected GrFeFETs are for receiving spikes from the presynaptic neuron and transmitting them to the postsynaptic neuron, respectively. The gate terminals are for imposing the supervise signals to adjust the channel conductance. As illustrated in Fig. 4a, an input spike triggers a decaying waveform by convoluting the input with window function in the supervise circuit (the $\int_0^{\infty} W(\tau)S_{\mathrm{in}}(t-\tau)d\tau$ term in the above equation, as represented by the decaying waveform inside the square in the lower left corner of the circuit). The teach and output pulse signals will sample the waveform respectively according to their different timings, as depicted in the lower middle part of the supervise circuit. The resulted voltages are then fed to the gate terminals of the complementary GrFeFET synapses, respectively. Here the upper GrFeFET is a potentiative synapse, of which the weight is modified through the voltage sampled by teach signal, while the lower device is a depressive one tuned through voltage sampled by the output signal. Mathematically, the former implements $+S_{\mathrm{d}}(t)\int_0^{\infty} W(\tau)S_{\mathrm{in}}(t-\tau)d\tau$ term, while the latter does $-S_{\mathrm{o}}(t)\int_0^{\infty} W(\tau)S_{\mathrm{in}}(t-\tau)d\tau$ in the above ReSuMe expression. Note that the positive and negative signs before $S_{\mathrm{d}}(t)$ and $S_{\mathrm{o}}(t)$ are realized through the potentiative and depressive properties of the two complementary synapses: given the positive voltages sampled by teach and output signals, the conductance of the upper GrFeFET is enhanced, while that of the latter is reduced. The time chart of Fig. 4b shows an example, where the first-round output fires earlier than the desired ($t_{\mathrm{out}} < t_{\mathrm{d}}$). In this case, the amplitude of sampled voltage by the output signal is greater than that by the teach signal ($c_1 > c_2 > 0$). As results, the magnitude of conductance decrease of the lower depressive GrFeFET is greater than that of increase of the upper potentiative one. Therefore, the summing conductance of the two devices in parallel gets decreased and consequently the second-round output fires later. In this way the output timing t_{out} approaches the desired one t_{d} round-by-round. For the other case, where the initial output timing is later than the desired ($t_{\mathrm{out}} > t_{\mathrm{d}}$), similar SL is implemented by using this circuit. The above demonstration of working principle indicates that the key requirement on device properties is the symmetry of conductance tuning between the potentiative and depressive GrFeFETs. Without this symmetry, the training cannot get convergent. The mechanism is that assuming the same timings of t_{d} and t_{out}, two voltages with the same amplitudes while with opposite signs will be sampled as seen in Fig. 4b; however, different magnitudes of conductance tuning will then be induced by these two mirror voltages in the two asymmetric devices. In this situation, the summing conductance will continue to change while the ReSuMe algorithm demands no more modulation of the synaptic weight. In the real devices, as seen previously in Fig. 3b–e the symmetry between n-doped (potentiative) and p-doped (depressive) conduction of *one device* is quite difficult to obtain. Therefore, in this task two devices are employed and their conduction behaviors have been delicately tuned to be highly symmetric as illustrated in Fig. 4c. The blue and red curves represent the cycle-to-cycle conductance tuning of two GrFeFETs under a series of 50 positive top-gate voltage pulses followed by 50 negative ones. The stimulated conductance changes in the two devices are almost equal while opposite in directions, indicating nice symmetric electrical properties between the two GrFeFETs (The measured conductance tuning as functions of gate voltage amplitudes and the initial conductance for the two devices $\Delta G(V_{\mathrm{tg}}, G_0)$ are further demonstrated in Section "Conductance tuning as functions of gate voltage amplitudes and the initial conductance for two complementary GrFeFETs" of the Supplementary Information for interested readers). Figure 4d shows the converging processes of our ReSuMe circuit by using the electrical properties of the above two complementary devices where the first-round output timing $t_{\mathrm{out}} = 14$ ms, while that of desire t_{d} varies from 13 to 15 ms. Here the convergence is defined as that the relative difference between the final output timing and the desired one $|t_{\mathrm{out}} - t_{\mathrm{d}}|/|t_{\mathrm{d}}|$ is <1%. The figure indicates that convergence will be achieved with about 50 iterations at most.

Comparing with conventional approaches, in which single devices are used as synapses for SL in SNN,[17] the major advantages here are the greatly reduced auxiliary circuit and the simplified operations. As seen in ref. [17] for conventional approach not only were at circuit module level the design of neuron and synapse circuits quite complicated (e.g., three waveforms with

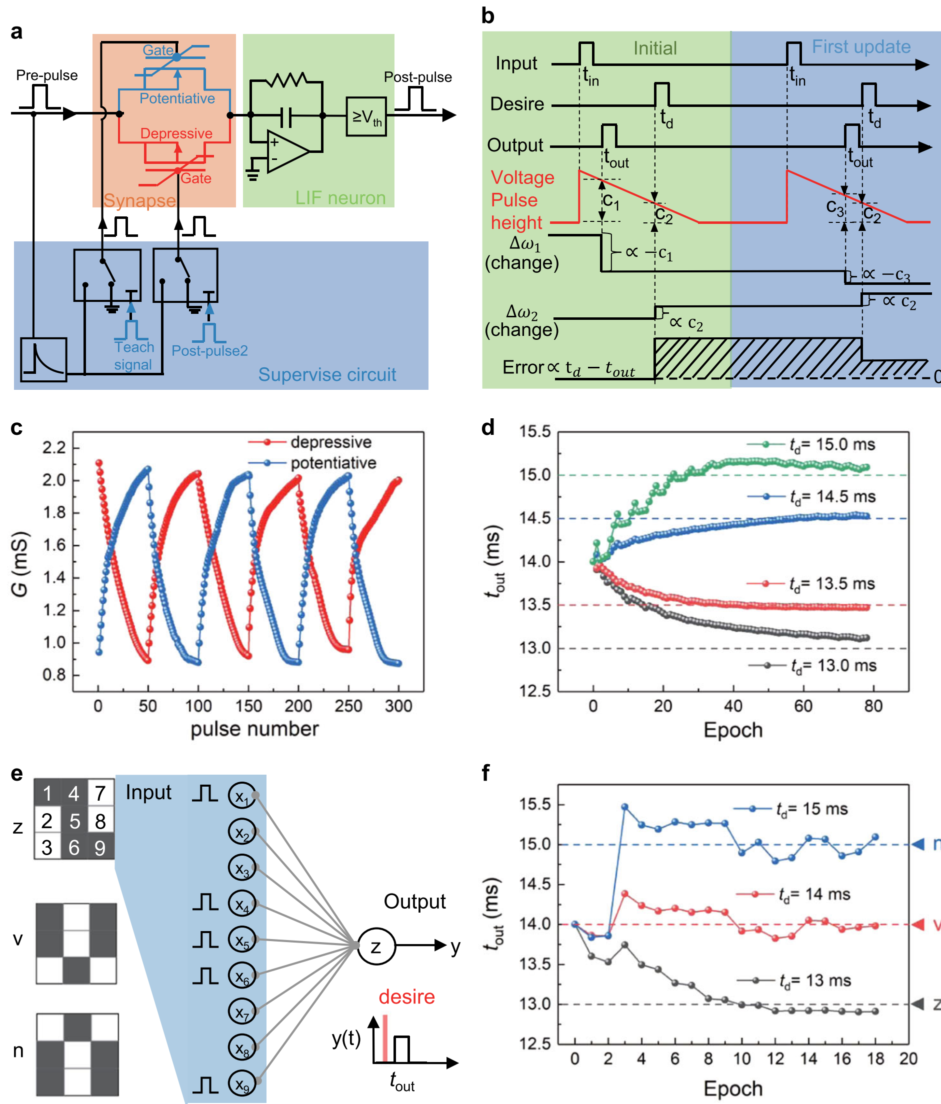

Fig. 4 The designed ReSuMe based on complementary GrFeFET synapses and its performance estimated through simulation. **a** ReSuMe module composed of complementary GrFeFET synapses, leaky integrate-and-fire (LIF) neuron, and supervise circuits. **b** Time chart of signals, where the Error is defined as $(t_d - t_{out})/(t_d + t_{out})$. Notice that since the desired time t_d keeps the same with respect to the input t_{in}, the sampled voltage amplitude c_2 and corresponding weight change of potentiative device Δw_2 are invariant during each training epoch. **c** The measured cycle-to-cycle analog weight update of two GrFeFET devices, where a series of positive voltage pulses V_{tg} with amplitudes 10 V and widths 50 ms are imposed followed by another series of V_{tg} with −8 V and 50 ms. **d** The difference between the desired timing and output one $(t_d - t_{out})$ versus the number of iterations. **e** and **f** Pattern classification with GrFeFET complementary synapse based ReSuMe learning. **e** The single-layer perceptron for classification of 3 × 3 binary images of z, v, and n, where the black/white pixels are encoded by spiking of nine input neurons, classification is represented by the different timings of the output neuron and the connection are by GrFeFET synapses. **f** The evolution of output signals, where lines with symbols are the output timings for z, v, and n inputs, while the dash lines are the desired ones

different amplitudes and durations had to be designed as a set of output spikes of neurons), but also at network level quite a few other types of circuit modules such as error detectors and analog adders were needed. On the other side, here by using complementary GrFeFET synapses both the circuit and operation have been greatly simplified as seen in Fig. 4a, b. The chip area efficiency would be drastically promoted while the power consumption would get substantially reduced due to the simplification. This improvement can also be found by comparing the present approach with that using single GrFeFET as synapse to implement ReSuMe (the latter is presented in Section "The approach of using single GrFeFET as synapse (S-approach) to ReSuMe" of Supplementary Information for interested readers).

In order to check the network-level performance of the above GrFeFET synapse, we design a SNN to implement the standard classification task of 3 × 3-pixel z, v, and n images and test through simulation as shown in Fig. 4e, f. The black-and-white images are encoded by pulses of nine input neurons, while the different timings of the output neuron infer which images are inputted[17] (design, simulation results, and comparison of different encoding approaches are discussed are provided in Section "ReSuMe to MNIST recognition with GrFeFET synapses" of Supplementary Information). SL is then conducted through the GrFeFET-based complementary synapses as designed in Fig. 4a, where the network parameters and simulation details are provided in the "Method" section. As indicated by Fig. 4f, with <15 epochs of training satisfactory convergence is achieved for the three input patterns. The demonstrated capability of quick and accurate learning is ascribed to both the power of ReSuMe algorithm and

the hardware implementation by our complementary GrFeFET synapses.

Finally, the figures of merits of our GrFeFET synapses are listed as follows: the energy consumption of each synaptic weight update operation is about 50 nJ (estimation method is provided in the "Method" section), the time step is about 50 ms and the effective area per synapse is about 50 μm^2. By analyzing these performance indexes of GrFeFET synapses and the network level simulation results, we conclude that the major advantages of using graphene as channel material are the high mobility and large conductance ON/OFF ratio that are gained through optimizing the fabrication process in our experiments. The high mobility facilitates the signal transmission through synapse and thus helps reduce the power-delay product, while the large ON/OFF ratio greatly promotes the learning capacity at the network level. However, the asymmetry between potentiative and depressive synapses as surveyed previously is outlined as one major restriction of using GrFeFET as complementary synapses. As analyzed before, this nonideal factor is a by-product of the p-dope nature of graphene material. By further improving the fabrication process the p-dope problem can be alleviated. Other strategies include trials with 2-dimensional (2D) transition metal dichalcogenide (TMD) ferroelectric devices as complementary synapses for SNN design, since several types of 2D TMD have both modest bandgaps and bipolar conduction properties.[34,35]

In summary, compared with conventional 2-terminal memristors we have found 3-terminal nonvolatile transistor appropriate to implement the synaptic plasticity required by the SL tasks in SNN, where the source/drain terminals are for transmitting spike signals from presynaptic neuron to the post one, while the gate terminal is used to impose the supervise signals. Based on the fabricated graphene–ferroelectric transistor and the measured nonvolatile and continuous tuning of channel conductance by gate voltages, we have realized complementary synapses. In these synapses, the analog weight update can be positive or negative depending on hole or electron dominance within the graphene channels. It is physically ascribed to the zero bandgap characteristic of graphene, while can be utilized to reconfigure the synapse to be potentiative or depressive. Interestingly, successful transition between these potentiative and depressive synapses have been achieved through large amplitude gate modulation of the ferroelectric polarization. The synapses have been further applied to implement remote supervise method in SNN. Through system-level simulation, we have further verified that the constructed synapses and SNN can perform classification of 3 × 3-pixel images after tens of iterations of training. In the future, concerning the proposed complementary synapses we will try to develop more complicated functions such as using the two complementary synapses to implement spike timing dependent plasticity (STDP) and anti-STDP, respectively, and dynamically interchange them, and hopefully a number of hardware architectures and the associated designs of neuromorphic computing will be accomplished.

METHODS

GrFeFET fabrication

First, the source and drain electrode regions (with 5 μm between them) were patterned by UV lithography on SiO2/p-Si (300 nm/500 um) substrate. Cr/Au (10/50 nm) electrodes were deposited through e-beam evaporation followed by a lift-off process. Commercial single layer graphene grown on copper foil with PMMA support layer was wet transferred onto the as-prepared electrodes. After removing PMMA layer by acetone, a photoresist bar with 10 μm wide was defined on the graphene between source and drain electrodes with UV lithography. The graphene within the uncovered region was removed by reactive-ion etching. By removing the residue photoresist with acetone, graphene channel with 5 μm × 10 μm size was fabricated. Note that the channel fabrication process should be finished in <2 h in order to reduce photoresist residue contamination as much as possible. After that, the sample was transferred in a glove box with argon atmosphere for top gate dielectric layer fabrication. PVDF-Trfe solution (70/30 mol%, dissolved in dimethylformamide with 3 wt%) was spin coated on graphene with film thickness of ~70 nm. Sample was annealed at 115 °C for 10 min to evaporate the solvent followed by 4 h of further annealing at 135 °C to enhance the crystallinity of the organic ferroelectric film. The aluminum electrode was then fabricated as top gate through UV lithography patterning, e-beam evaporation, and 4-h immersing in isopropanol as lift off. The fabricated device was characterized by optical profiler Olympus LEXT OLS5000 Industrial Laser Confocal Microscopes as demonstrated in Fig. 1c.

Measurement

The polarization versus electric field (*P-E*) curves of a ferroelectric capacitor with gold (Au) bottom electrode and chromium (Cr) top electrode, and thickness of ~100 nm was measured by using Radiant Inc circuit. The capacitance–voltage (*C–V*) relationship obtained from Au/PVDF/Al structure, where PVDF is 100 nm was measured by using a B1500A parameter analyzer at 10 kHz applied voltage frequency. The other measurements were performed using a Keithley 4200A-SCS parameter analyzer. All channel conductance was collected by compelling a DC bias (0.1 V) between source and drain.

Model

The carrier concentrations (electrons or holes) in the graphene channel n_{total} is estimated by[21]

$$n_{\text{total}} = \sqrt{n_0^2 + n(V_{\text{tg}})^2} \tag{2}$$

n_0 is the residual carrier concentration characterizing the density of carriers at the minimum conductivity, i.e., at Dirac point. $n(V_{\text{tg}})$ is the carrier concentration (electrons or holes) induced by the top gate voltage, measuring the Fermi level modulated away from the Dirac point. The total device resistance R_{total} is:

$$R_{\text{total}} = R_{\text{contact}} + \frac{L/W}{n_{\text{total}} e \mu} = 1/G, \tag{3}$$

where R_{contact} is the metal/graphene contact resistance, L and W are the length and width of graphene channel, and μ is the charge carrier mobility. The continuity of electric displacement field D at the PVDF/graphene interface then gives rise to the following equation[21]:

$$D = \varepsilon_0 \varepsilon_r E_{\text{PVDF}} + P(V_{\text{tg}}) = -n(V_{\text{tg}})e, \tag{4}$$

where $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m is the vacuum dielectric constant, $\varepsilon_r = 10$ is the dielectric constant of PVDF, and E_{PVDF} is the electric field within PVDF. The item $\varepsilon_0 \varepsilon_r E_{\text{PVDF}}$ represents the linear component in the dielectric response of the ferroelectrics, which is the common property in most dielectric, while $P(V_{\text{tg}})$ is the hysteretic component. Two additional equations concerning the capacitive effect of gate dielectric and the electrostatics are as below:

$$ne = C(V_{\text{tg}} - V_{\text{Dirac}}) \tag{5}$$

$$E_{\text{PVDF}} = V_{\text{tg}}/d, \tag{6}$$

where C is the capacitance of PVDF gate dielectric and V_{Dirac} is the Dirac point of graphene material. By combining the above equations, the relation between the measured conductance G and the imposed top gate voltage V_{tg} is derived as

$$G = \frac{1}{R_{total}} = 1/\left(R_{\text{contact}} + \frac{L}{We\mu\sqrt{n_0^2 + C^2(V_{\text{tg}} - V_{\text{Dirac}})}}\right) \tag{7}$$

By fitting this model to the measured top-gate transfer curves shown in Fig. 2a, parameters such as R_{contact} and μ are obtained as $R_{\text{contact}} \approx 600\,\Omega$ and $\mu \approx 1.7 \times 10^3$ cm^2/V s. Here we remind that the values of μ fitted from the left and right branches in Fig. 2a are almost the same.

The electric boundary condition at the interface between the gate dielectric PVDF and the graphene channel is described by the following equation:

$$P_{Vtg} D = \varepsilon_0 \varepsilon_r E_{\text{PVDF}} + P(V_{\text{tg}}) = -n(V_{\text{tg}})e. \tag{8}$$

By solving the above equation, *D*(*E*) is then extracted from the measured $G(V_{\text{tg}})$ shown in Fig. 2b.

Estimation of nonideal factors

The analog weight update behaviors of GrFeFET synapses shown in Fig. 3c, d are usually measured with the following nonideal factors. One is the nonlinearity factor α of long-term potentiation (LTP) and depression (LTD) processes[36]:

$$B = (G_{max} - G_{min})/(1 - e^{-P_m/A}) \tag{9}$$

$$G_{LTP} = B \cdot \left(1 - e^{-P/A}\right) + G_{min} \tag{10}$$

$$G_{LTD} = G_{max} - B \cdot \left(1 - e^{(P-P_m)/A}\right) + G_{min} \tag{11}$$

$$\alpha = 1.726/(A + 0.162), \tag{12}$$

where G_{max} and G_{min} are the maximum and minimum conductance, P_m is the maximum number of pulses required to tune the conductance from G_{min} to G_{max}, while A and B are fitting parameters.

Another is the asymmetry β between LTP and LTD[36]:

$$\beta = \left[G_{LTP}\left(\frac{N}{2}\right) - G_{LTD}\left(\frac{N}{2}\right)\right]/(G_{max} - G_{min}) \tag{13}$$

The third is the cycle-to-cycle variation σ[37]:

$$G = G_{ideal} + N(\sigma) \cdot \sqrt{n}, \tag{14}$$

where σ is the standard deviation of the conductance at different cycles obtained from the experiment, $N(\sigma)$ is the normal distribution of the variation, n is the number of pulses to be applied for each update, and G_{ideal} is the conductance when no variation is introduced.

Simulation

The synapse, neuron, and supervise circuits to execute ReSuMe are built with MATLAB Simulink. The parameters are listed in the following Table 2:

Figure 5 demonstrates the flowchart for implementing 3 × 3 pixel z, v, and n classification is as follows:

Estimation of energy consumption

The energy consumed per update operation of synaptic weight is calculated by $E_{up} \approx V_{tg} I_{tg} \Delta t$, where V_{tg} and Δt are the amplitude and duration of the imposed top gate voltage pulses, and I_{tg} is the measured leaky current through the gate terminal during the update operation. In our measurement, I_{tg} is found to be 100 nA.

Table 2. Parameters for ReSuMe simulation

	Model	Leaky integrate-and-fire
Neuron	Spike threshold	$V_{th1} = 1.60$ mV
		$V_{th2} = 11.54$ mV
	Time constant	$\tau = 40$ ms
Synapse	Minimal synaptic weight	4 mS
	Maximal synaptic weight	2 mS
Supervise circuit	Learning Windows	$U = \mp(1.5\,t + k)$, where k is a constant

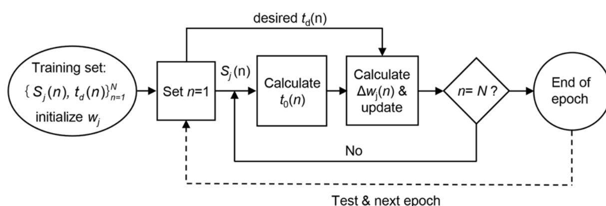

Fig. 5 Flowchart of training and test

DATA AVAILABILITY

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

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China under Grant Nos. 61841404, 51732003, and by Hubei Engineering Research Center on Microelectronics.

AUTHOR CONTRIBUTIONS

Y.C., Y.Z., and F.Z. made the same contributions; Y.H. and F.Z. conceived the idea; Y.C. fabricated the devices; Y.Z. performed the electrical measurement, circuit design, and simulation; Y.H., F.Z., B.T., and Y.L. conducted the analysis; M.Y. helped with the measurement; Y.H. and F.Z. wrote and revised the paper; X.S.M. supervised and supported the whole work.

ADDITIONAL INFORMATION

Supplementary information accompanies the paper on the *npj 2D Materials and Applications* website (https://doi.org/10.1038/s41699-019-0114-6).

Competing interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Mead, C. Neuromorphic electronic systems. *Proc. IEEE* **78**, 1629–1636 (1990).
2. Jo, S. H. et al. Nanoscale memristor device as synapse in neuromorphic systems. *Nano Lett.* **10**, 1297–1301 (2010).
3. Yu, S. et al. A low energy oxide-based electronic synaptic device for neuromorphic visual systems with tolerance to device variation. *Adv. Mater.* **25**, 1774–1779 (2013).
4. Duygu, K., Shimeng, Y. & Wong, H. S. P. Synaptic electronics: materials, devices and applications. *Nanotechnology* **24**, 382001 (2013).
5. Yang, J. J., Strukov, D. B. & Stewart, D. R. Memristive devices for computing. *Nat. Nanotechnol.* **8**, 13 (2012).
6. Ambrogio, S. et al. Novel RRAM-enabled 1T1R synapse capable of low-power STDP via burst-mode communication and real-time unsupervised machine learning. In *Proc. 2017 IEEE Symposium on VLSI Technology* 1–2. https://doi.org/10.1109/VLSIT.2016.7573432 (2016).
7. Kuzum, D., Jeyasingh, R. G. D., Lee, B. & Wong, H. S. P. Nanoelectronic programmable synapses based on phase change materials for brain-inspired computing. *Nano Lett.* **12**, 2179–2186 (2012).
8. Suri, M. et al. Bio-inspired stochastic computing using binary CBRAM synapses. *IEEE Trans. Electron Devices* **60**, 2402–2409 (2013).
9. Park, S. et al. RRAM-based synapse for neuromorphic system with pattern recognition function. In *Proc. 2012 International Electron Devices Meeting*. 10.12.11–10.12.14. https://doi.org/10.1109/IEDM.2012.6479016 (2012).
10. Wang, Z. et al. Fully memristive neural networks for pattern classification with unsupervised learning. *Nat. Electron.* **1**, 137–145 (2018).
11. Burr, G. W. et al. Experimental demonstration and tolerancing of a large-scale neural network (165,000 synapses), using phase-change memory as the synaptic weight element. In *Proc. 2017 IEEE International Electron Devices Meeting*. 29.25.21–29.25.24. https://doi.org/10.1109/IEDM.2014.7047135 (2014).
12. Yao, P. et al. Face classification using electronic synapses. *Nat. Commun.* **8**, 15199 (2017).
13. Hu, M. et al. Memristor-based analog computation and neural network classification with a dot product engine. *Adv. Mater.* **30**, 1705914 (2018).
14. Kim, T., Kim, H., Kim, J. & Kim, J. Input voltage mapping optimized for resistive memory-based deep neural network hardware. *IEEE Electron Device Lett.* **38**, 1228–1231 (2017).
15. Maass, W. Networks of spiking neurons: the third generation of neural network models. *Neural Netw.* **10**, 1659–1671 (1997).
16. Nishitani, Y., Kaneko, Y., Ueda, M., Morie, T. & Fujii, E. Three-terminal ferroelectric synapse device with concurrent learning function for artificial neural networks. *J. Appl. Phys.* **111**, 124108 (2012).
17. Nishitani, Y., Kaneko, Y. & Ueda, M. Supervised learning using spike-timing-dependent plasticity of memristive synapses. *IEEE Trans. Neural Netw. Learn. Syst.* **26**, 2999–3008 (2015).

18. Tian, B. et al. A robust artificial synapse based on organic ferroelectric olymer. *Adv. Electron. Mater.* **5**, 1800600 (2019).
19. Yang, Y. et al. Multifunctional nanoionic devices enabling simultaneous hetero-synaptic plasticity and efficient in-memory boolean logic. *Adv. Electron. Mater.* **3**, 1700032 (2017).
20. Zheng, Y. et al. Gate-controlled nonvolatile graphene-ferroelectric memory. *Appl. Phys. Lett.* **94**, 163505 (2009).
21. Zheng, Y. et al. Graphene field-effect transistors with ferroelectric gating. *Phys. Rev. Lett.* **105**, 166602 (2010).
22. Tian, H. et al. Graphene dynamic synapse with modulatable plasticity. *Nano Lett.* **15**, 8013–8019 (2015).
23. Tian, H. et al. A novel artificial synapse with dual modes using bilayer graphene as the bottom electrode. *Nanoscale* **9**, 9275–9283 (2017).
24. Gerstner, W., Kistler, W. M., Naud, R. & Paninski, L. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition* (Cambridge University Press, 2014).
25. Jerry, M. et al. Ferroelectric FET analog synapse for acceleration of deep neural network training. In *Proc. 2017 IEEE International Electron Devices Meeting (IEDM)*. 6.2.1–6.2.4. https://doi.org/10.1109/IEDM.2017.8268338 (2017).
26. Yang, J. J. Memristor crossbar arrays for analog and neuromorphic computing. https://apps.dtic.mil/docs/citations/AD1061408 (2018).
27. Raghavan, S. et al. Long-term retention in organic ferroelectric-graphene memories. *Appl. Phys. Lett.* **100**, 023507 (2012).
28. Yao, Y. et al. Reconfigurable artificial synapses between excitatory and inhibitory modes based on single-gate graphene transistors. *Adv. Electron. Mater.* **0**, 1800887 (2019).
29. Boyn, S. et al. Learning through ferroelectric domain dynamics in solid-state synapses. *Nat. Commun.* **8**, 14736 (2017).
30. Ponulak, F. & Kasinski, A. J. Supervised learning in spiking neural networks with ReSuMe: sequence learning, classification, and spike shifting. *Neural Comput.* **22**, 467–510 (2010).
31. Bohte, S. M., Kok, J. N. & La Poutré, H. Error-backpropagation in temporally encoded networks of spiking neurons. *Neurocomputing* **48**, 17–37 (2002).
32. Mohemmed, A., Schliebs, S., Matsuda, S. & Kasabov, N. Span: spike pattern association neuron for learning spatio-temporal spike patterns. *Int. J. Neural Syst.* **22**, 1250012 (2012).
33. Mohemmed, A., Schliebs, S., Matsuda, S. & Kasabov, N. Training spiking neural networks to associate spatio-temporal input–output spike patterns. *Neurocomputing* **107**, 3–10 (2013).
34. Agnihotri, P., Dhakras, P. & Lee, J. U. Bipolar junction transistors in two-dimensional WSe2 with large current and photocurrent gains. *Nano Lett.* **16**, 4355–4360 (2016).
35. Rasmussen, F. A. & Thygesen, K. S. Computational 2D materials database: electronic structure of transition-metal dichalcogenides and oxides. *J. Phys. Chem. C.* **119**, 13169–13183 (2015).
36. Chen, P. et al. Mitigating effects of non-ideal synaptic device characteristics for on-chip learning. In *Proc. 2015 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)* 194–199. https://doi.org/10.1109/ICCAD.2015.7372570 (2015).
37. Chen, P., Peng, X. & Yu, S. NeuroSim+: an integrated device-to-algorithm framework for benchmarking synaptic devices and array architectures. In *Proc. 2017 IEEE International Electron Devices Meeting (IEDM)* 6.1.1–6.1.4. https://doi.org/10.1109/IEDM.2017.8268337 (2017).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019