High Conductance Margin for Efficient Neuromorphic Computing Enabled by Stacking Nonvolatile van der Waals Transistors

Liping Xu (徐丽萍),[1,2] Hao Xiong (熊浩),[1] Zhichao Fu (付志超),[3] Menghan Deng (邓梦晗),[1] Shuiyuan Wang (王水源),[4] Jinzhong Zhang (张金中),[1,*] Liyan Shang (商丽燕),[1] Kai Jiang (姜凯),[1] Yawei Li (李亚巍),[1] Liangqing Zhu (朱亮清),[1] Liang He,[3] Zhigao Hu (胡志高),[1,5,6,†] and Junhao Chu (褚君浩)[1,5,6]

[1] *Technical Center for Multifunctional Magneto-Optical Spectroscopy (Shanghai), Engineering Research Center of Nanophotonics & Advanced Instrument (Ministry of Education), Department of Materials, School of Physics and Electronic Science, East China Normal University, Shanghai 200241, China*

[2] *Center for Advanced Electronic Materials and Devices, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

[3] *School of Computer Science and Technology, East China Normal University, Shanghai 200062, China*

[4] *ASIC & System State Key Laboratory, School of Microelectronics, Fudan University, Shanghai 200433, China*

[5] *Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China*

[6] *Shanghai Institute of Intelligent Electronics & Systems, Fudan University, Shanghai 200433, China*

(Received 14 July 2021; accepted 5 October 2021; published 26 October 2021)

High-performance artificial synaptic devices are key building blocks for developing efficient neuromorphic computing systems. However, the nonlinear and asymmetric weight update of existing devices has restricted their practical applications. Herein, floating gate nonvolatile memory (FG NVM) devices based on two-dimensional (2D) HfS_2/*h*-BN/FG-graphene heterostructures have been designed and investigated as multifunctional NVM and artificial optoelectronic synapses. Benefiting from the FG architecture, the HfS_2-based NVM device exhibits competitive performances, such as a high on:off ratio ($> 10^5$), large memory window (approximately 100 V), excellent charge retention ability ($> 10^4$ s), and robust durability ($> 10^3$ cycles). Notably, the artificial optoelectronic synapses based on HfS_2 FG NVM show an impressive large conductance margin and good linearity, owing to the ultrahigh photoresponsivity and photogain of HfS_2. The energy consumption of per spike in our artificial synapse is as low as 0.2 pJ. Therefore, a high recognition accuracy up to 91.5% of the artificial neural network on the basis of our HfS_2-based optoelectronic synapse at the system level has been achieved, which is superior to other reported 2D artificial optoelectronic synapses. This work paves the way forward for all 2D material-based memory for developing efficient optogenetics-inspired neuromorphic computing in brain-inspired intelligent systems.

DOI: 10.1103/PhysRevApplied.16.044049

I. INTRODUCTION

In recent years, the significant advancement in the era of artificial intelligence (AI) has been achieved, leading to urgent demands on high computing speed and power efficiency [1–4]. In contrast to conventional von Neumann architectures, neuromorphic computing replicates the structure of the human brain, exhibiting strong fault tolerance, high efficiency, and low energy consumption [5–8]. The huge number of neurons (approximately equal to 10^{11}) and synapses (approximately equal to 10^{15}) in the human brain conduct the information transmission and processing [9]. In particular, the synapses can process and store information simultaneously by modulating the strength of connection between neurons [10]. It is indispensable to emulate a synapse with an artificial device for the development of neuromorphic computing systems. Therefore, the implementation of synaptic devices with high on:off ratio, good data retention, and linear weight update is imperative.

Nowadays, various types of devices have been considered for implementation of highly efficient artificial synapses, such as memristors [11–15], phase change memory (PCM) [16,17], ferroelectric transistors (FeFETs) [18–21], field-effect transistors (FETs) [22–33], and electrochemical transistors [34]. Thereinto, three-terminal transistors exhibit a more linear and symmetric electrostatic gate control, due to the better controllability from the third terminal (gate) [33]. From the perspective of scaling down the limitation of silicon-based materials, atomically

*jzzhang@ee.ecnu.edu.cn
†zghu@ee.ecnu.edu.cn

thin two-dimensional (2D) van der Waals materials have attracted considerable interest as promising candidates for the next generation of nanoelectronics due to their superior electrical and optical properties [35–39]. 2D van der Waals semiconductors have been employed as channel materials to implement transistor-based synapses [5,6,23,36,37]. For example, InSe-based multibit memory exhibits large hysteresis in the transfer characteristics, high on:off ratio (approximately equal to 10^5), reliable switching operation and stable retention (10^4 s), which can be used to mimic the synaptic functionality by electrical stimuli [5]. In most of the reported synaptic devices, electrical stimuli was applied to modulate the synaptic weight. However, the processing speed of those electrical stimulated devices can be restricted, due to the bandwidth-connection-density trade-off and the considered interconnection issues [3]. Moreover, the nonlinear weight update of electronic synaptic devices is one of the major obstacles in realizing complex artificial neural networks (ANNs).

In contrast to electrical stimulus, light or photons may offer high bandwidth, ultrafast signal transmission, low power computation, and low cross-talk, resulting in the enhancement of the computational speed [7]. Therefore, many ingenious studies of optoelectronic memories and synaptic devices have demonstrated their capability in the applications of image preprocessing and neuromorphic visual systems [40–46]. Recently, fully optical-modulation synapses based on naturally oxidized black phosphorus (BP), and pyrenyl graphdiyne/graphene/PbS quantum dot have been investigated [3,41,43,44]. However, the uncontrollability of native oxidation for BP film brings about excessive device-to-device variation. Both the fully optical-modulation synapses mentioned above suffer from low dynamic range of weight (on:off ratio < 4), resulting in low recognition accuracy of ANNs in practical applications. In previous reports, the highest recognition accuracy of ANN based on artificial optoelectronic synapse is 90.8%, due to the low dynamic range of weight and large nonlinearity value [3,31,44]. Thus, the optoelectronic synapses with high on:off ratio, good data retention, and linear weight update toward neural networks are highly desired.

In this work, an artificial optoelectronic synapse is proposed on the basis of an all 2D few-layer HfS_2/*h*-BN/FG-graphene device. The HfS_2-based nonvolatile memory (NVM) shows a high on:off ratio ($> 10^5$), large memory window (approximately 100 V), excellent charge retention ability ($> 10^4$ s), and durability ($> 10^3$ cycles). Furthermore, the basic synaptic behaviors, such as pair-pulse facilitation (PPF), short-term plasticity (STP), long-term potentiation (LTP), and long-term depression (LTD) have been well mimicked by our HfS_2-based NVM device. Memory and learning behaviors are also achieved under various 405-nm UV light-pulse conditions. Specifically, the synaptic device exhibits better linearity and higher $G_{max} : G_{min}$ ratio in conductance-update characteristics by using light and voltage pulses than that observed under voltage-pulsing mode, owing to the ultrahigh photoresponsivity and photogain of HfS_2. The energy consumption per spike in our artificial synapse can be as low as 0.2 pJ. The maximum recognition accuracy for Modified National Institute of Standards and Technology (MNIST) digit patterns can be achieved up to 91.5% based on LTP and LTD curves under the light and voltage pulses. These comprehensive results pave the way for building synaptic device using all 2D materials for efficient neuromorphic computing.

II. EXPERIMENTAL METHODS

A. Fabrication of devices

To fabricate the floating gate NVM device consisting of HfS_2, *h*-BN, and graphene, standard dry-transfer methods are employed. HfS_2, *h*-BN, and graphene flakes with the thickness of a few nanometers are mechanically exfoliated from bulk crystals by polydimethysiloxane (PDMS) and vertically stacked on SiO_2(300 nm)/Si substrate. After transfer, to enhance the connection of the heterostructure, the sample is annealed at 100 °C for 15 min in an argon atmosphere. Standard electron-beam lithography (EBL) techniques are used to pattern the source and drain electrodes on a fabricated heterojunction. The thermal evaporator is used to deposit Ni/Au (2/50 nm) electrodes.

B. Characterizations

The surface topology and thickness of the exfoliated HfS_2, *h*-BN, and graphene flakes are characterized on a commercial AFM system (Dimension Icon, Bruker) in ScanAsyst modes. Raman spectra are measured by using a confocal micro-Raman spectrometer (Jobin-Yvon LabRAM HR Evolution, Horiba) with a 532-nm excitation laser line and 100x-0.9 N.A. objective lens.

The electronic and optoelectronic measurements of the floating gate NVM device are performed using a Keithley 4200-SCS semiconductor parameter analyzer. The devices are tested in a Janis ST-500 cryogenic probe station in a high-vacuum condition (10^{-6} Torr) at room temperature. All tests are carried out under dark conditions, exposure to the target light sources only. For the optoelectronic measurements, commercial light-emitting diodes with wavelengths of 405 and 637 nm (Thorlabs, Inc.) are employed. Light pulses with tunable power intensity, width, and frequency are controlled by a laser diode and temperature controller (ITC4001, Thorlabs, Inc.). The gate electrical pulses are generated by a semiconductor pulse generator unit (SPGU) module.

C. Artificial neural network simulation

The image classification simulations of handwritten digits from the MINST dataset for an ANN are carried out using the PyTorch platform. The simulator containing a three-layer (one hidden layer) perceptron neural network with 784 input, 150 hidden, and 10 output neurons is used to perform the recognition tasks. The training of the network is executed using 60 000 images from the MINST dataset. The recognition accuracy of the ANN is tested using 10 000 separate images.

III. RESULTS AND DISCUSSION

Figures 1(a) and 1(b) show the representative 3D schematic and optical micrograph of the vertically stacked FET device, which consists of HfS_2/*h*-BN/graphene layers. In this structure, multilayer HfS_2 serves as the channel, high-quality *h*-BN and graphene acts as the carrier tunneling layer and trapping layer, respectively. The degenerately doped silicon (P^{++} Si) works as the control gate. A back-gate voltage V_{CG} is applied to tune the memory characteristics. As monitored from the atomic force microscopy (AFM) images and the line profiles shown in Fig. 1(c) and Fig. S1 within the Supplemental Material [47], the average thicknesses of graphene, *h*-BN, and HfS_2 flakes are about 10, 20, and 20 nm, respectively. The crystalline qualities of the transferred HfS_2, *h*-BN, and graphene flakes are confirmed by Raman spectroscopy [Fig. 1(d)]. From the Raman spectrum of HfS_2 flakes, two typical peaks appear at 259 and 337 cm^{-1}, which are attributed to E_g (in-plane) and A_{1g} (out-of-plane) modes, respectively [35]. The characteristic peak of *h*-BN is observed at 1365 cm^{-1}, corresponding to the in-plane (E_{2g}) vibration mode. The Raman spectrum of graphene shows distinct *G* and 2D peaks at 1581 and 2722 cm^{-1}, respectively [48]. Details of the fabrication process for the HfS_2-based NVM device can be found in Sec. II. Figure 1(e) presents the output characteristic curves of a HfS_2-based NVM device for different V_{CG} values from -60 to $+60$ V. The linear relation between I_{DS} and V_{DS} represents decent electrical contacts with low contact resistance. Good Ohmic contact between the HfS_2 channel and source and drain electrodes is confirmed by a linear I_{DS}-V_{DS} curve under various V_{CG} bias. The typical transfer characteristics obtained from ± 80 V V_{CG} sweep under different V_{DS} values are shown in Fig. 1(f). The HfS_2-based FET exhibits electron doping (*n*-type) in HfS_2. The electron mobility of around 8–15 $cm^2V^{-1}s^{-1}$ and on:off current ratio over 10^5 are observed from the representative transfer curves. The effect of the channel length on the electrical transport properties for MoS_2-based FET has been investigated by Liu *et al.* [49]. For long channel devices, the on:off current ratio and field-effect mobility would remain nearly constant. In this work, the channel length is fixed at 5 μm. Thus, our HfS_2-based devices are long channel devices. Acting as the carrier tunneling layer in this floating-gate structure, the thickness of *h*-BN would affect the tunnel current of devices. The tunnel current relies exponentially on the thickness of *h*-BN, down to a monolayer thickness. The detailed investigation about the tunneling behavior dependence on thickness of *h*-BN can be found elsewhere [50].

The large memory window observed in Fig. 1(f) benefits from the floating-gate structure of the FET. For comparison, the representative transfer characteristic curves of HfS_2/SiO_2 FET structure are illustrated in Fig. 2(a). The hysteresis windows between the forward and reverse transfer characteristic curves are pretty small, as compared to the HfS_2/*h*-BN/FG-graphene FET structure. Additionally, the representative output and transfer characteristic curves of HfS_2/*h*-BN/SiO_2 FET structure are also measured and plotted in Fig. S2 within the Supplemental Material [47]. The weak electrical hysteresis windows may originate from the charge-transfer effect, owing to the interface traps between HfS_2 and dielectric layer or defects induced during the fabrication process [5]. The multiple transfer characteristic curves of the HfS_2-based FG NVM device under different V_{CG} sweeping ranges are illustrated in Fig. 2(b). The gate leakage currents I_{GS} are simultaneously plotted as a reference in Fig. 2(b). The memory hysteresis window (ΔV_{th}) can be enlarged with the increase of the $V_{CG\,max}$. Figure 2(c) presents the positive-, negative-shifted V_{th} with respect to each V_{CG} sweep range. The extracted hysteresis window (ΔV_{th}) as a function of the $V_{CG\,max}$ are presented in Fig. 2(d). The positive correlation between ΔV_{th} and the applied $V_{CG\,max}$ indicates that the charge-trapping behaviors of the HfS_2-based NVM device can be effectively modulated by control gate voltage. Furthermore, the amount of charge trapped in the FG-graphene layer as a function of V_{CG} can be estimated from the capacitance model of $n = \Delta V_{th} C_{ox}/q$, where q and C_{ox} are the elementary charge, and the capacitance between a floating gate and the control gate, respectively. The inset in Fig. 2(d) shows the calculated stored carrier density as a function of V_{CG}. For $\Delta V_{th} = 98$ V under a sweep gate bias of $V_{CG} = \pm 80$ V, the trapped carrier density is estimated to be around 7.5×10^{12} cm^{-2}. Due to the charging and emission process induced by interface traps between graphene and SiO_2, there may be a little bit of error in the calculated value of *n* in FG graphene. These results above indicate the good transport properties of our HfS_2-based NVM. Moreover, the device-to-device variation has a critical role in practical application. Therefore, we provide the transfer characteristic curves and threshold voltage shift of five representative HfS_2-based NVM devices as shown in Fig. S3 within the Supplemental Material [47]. The observed transport hysteresis loops are consistently enlarged, suggesting the boosted electrical hysteresis phenomena. The average value of the transport hysteresis loop approached 95 V, which indicates that our devices have a good repeatability.

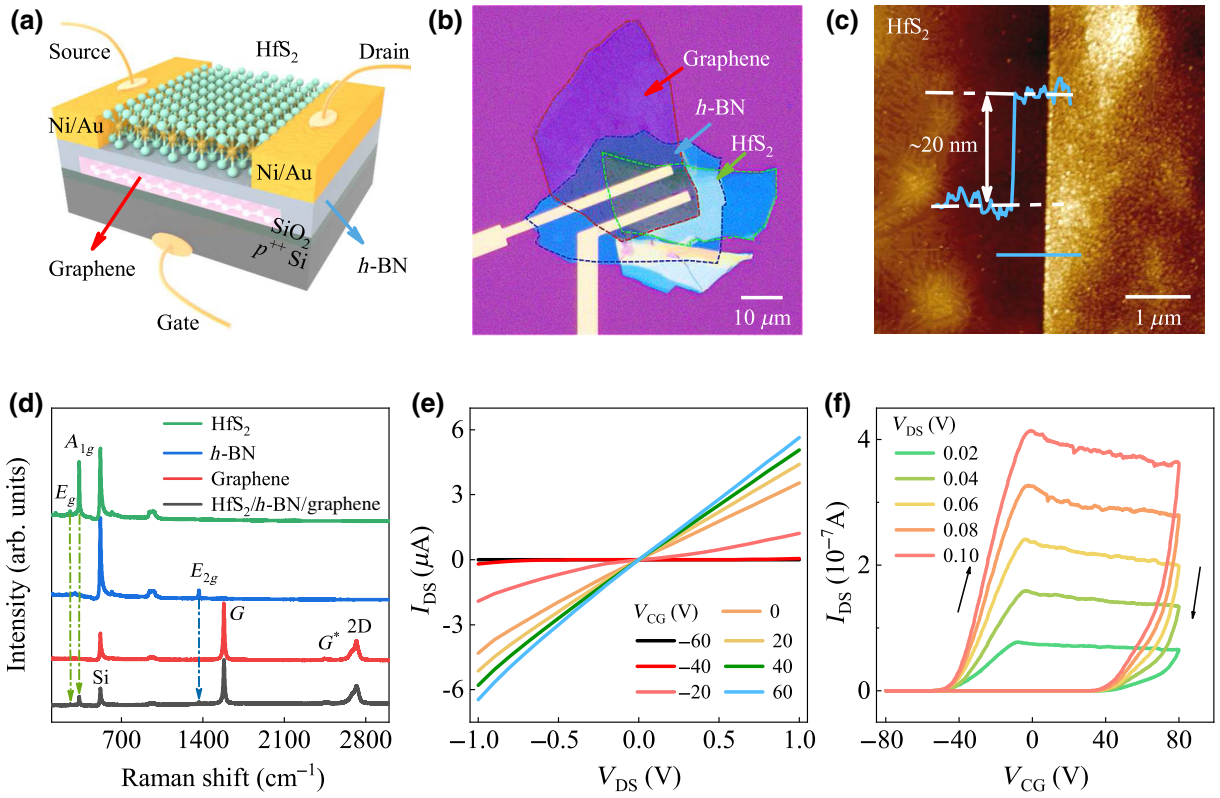

FIG. 1. (a) Three-dimensional (3D) schematic diagram of the device composing of HfS_2/*h*-BN/graphene layers. (b) Optical image of the stacked HfS_2, *h*-BN, and graphene flakes in FET with electrical contacts on Si/SiO_2 substrate. (c) Surface topography AFM image of the HfS_2 flake. Note that the thickness of the blue line is obtained from the line profile. (d) Raman spectra of the HfS_2, *h*-BN, graphene, and HfS_2/*h*-BN/graphene stacked heterostructure. (e) Output characteristic curves showing good Ohmic behavior for different control gate bias. (f) Transfer characteristic curves with a large memory window at different V_{DS} from 0.02 to 1 V.

In general, satisfactory data retention and endurance characteristics are critical to the practical application of a NVM device. To verify the memory retention characteristics, V_{CG} pulses (±80 V, 1 s) are applied. The high erase-to-program ratio over 10^5 under a V_{DS} value of 0.1 V is achieved and well maintained in the time range of 10^4 s as shown in Fig. 2(e), which can satisfy the needs of 10-year data retention in industry standard. The cycle endurance of the NVM is assessed by executing a sequential program (+80 V, 0.5 s) and erase (−80 V, 0.5 s) pulses. The program and erase states at zero gate voltage are observed and plotted as a function of cycle number as illustrated in Fig. 2(f). The erase-to-program ratio is still stable without noticeable degradation after 10^3 cycles, demonstrating the good reproducibility of the NVM device. To better understand the operation mechanism, i.e., the charge trapping and releasing process of the HfS_2/*h*-BN/FG-graphene FET, the band diagrams and the corresponding operation states under the modulation of V_{CG} are shown in Figs. 2(g)–2(i). The initial pristine state of our FET without gate pulse is shown in Fig. 2(g). The work function (ϕ) of HfS_2, graphene, and P^{++} Si is 5.27, 4.6, and 5 eV, respectively. The electronic affinity (χ) of *h*-BN and SiO_2 is 2 and 0.9 eV, respectively. Figure 2(h) defines the program operation by applying a positive V_{CG}. During the programming process, the electrons in the HfS_2 channel tunnel into FG graphene through *h*-BN based on the Fowler-Nordheim tunneling mechanism [6]. The electrons trapped in FG graphene can partially screen the electric field from control gate, leading to a positive shift of V_{th}, which correspondingly results in a lower current level, defined as the program state. While in the erase operation under a negative V_{CG} as shown in Fig. 2(i), the trapped electrons in FG graphene tunnel back to the HfS_2 channel. During the erasing process, the screen effect of electrons weakens, bringing about a negative shift of V_{th} and the high current level in the HfS_2 channel. The higher current level than the initial state is defined as the erase state. The trapping and releasing capabilities of electrons contribute to the responsible memory window of the device under different voltages.

Inspired by the the maintenance capability of the trapped charge carriers of the HfS_2-based NVM device discussed above, the basic synaptic performances of HfS_2 flash memory as an artificial synaptic device have been investigated. Synapses play a vital role in information exchange,

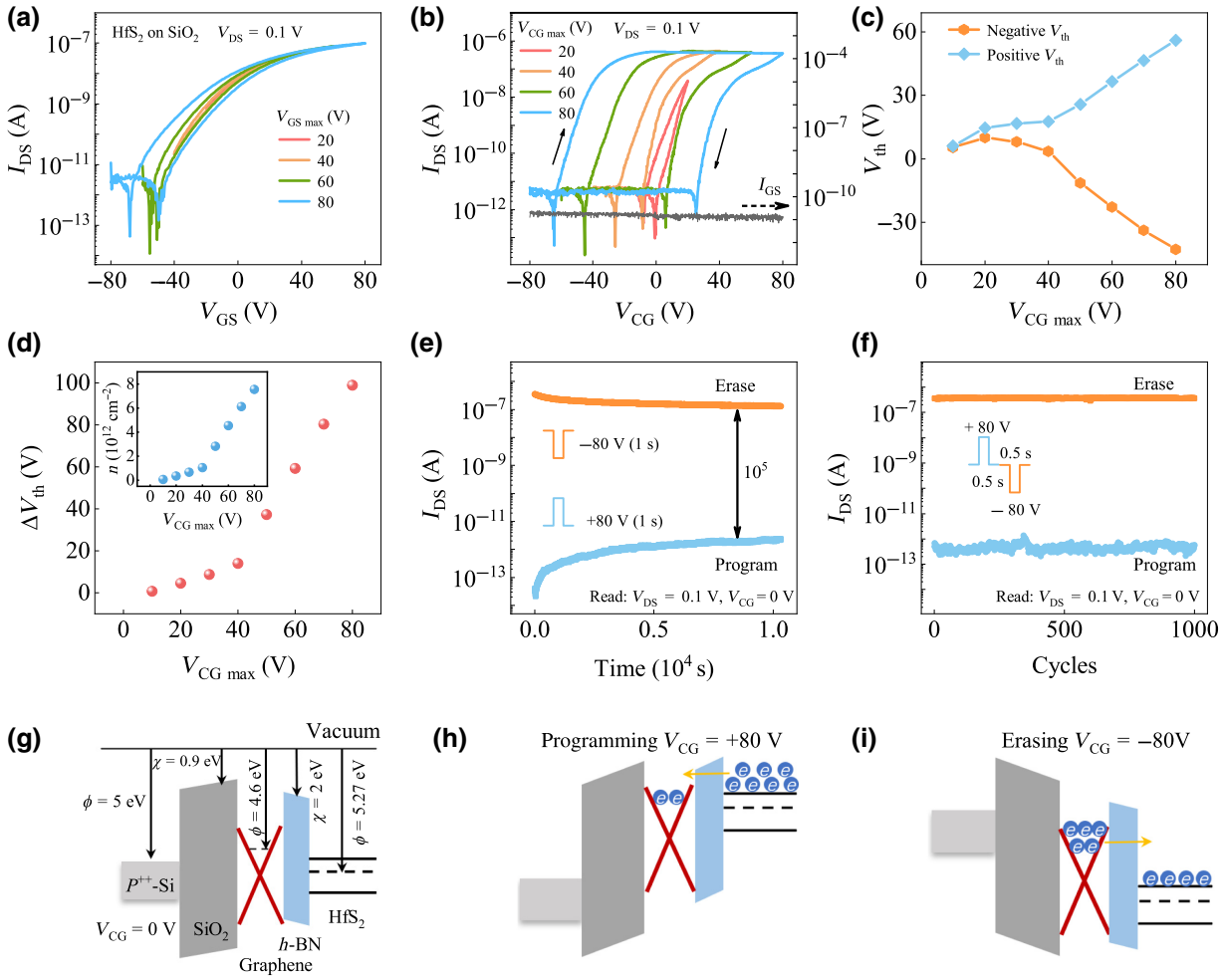

FIG. 2. (a) Transfer characteristics of the HfS_2 device on silicon. The hysteresis windows between the forward and reverse transfer characteristic curves are pretty small. (b) Transfer characteristic curves of the HfS_2-based NVM device with increasing hysteresis for different control gate-voltage sweeping ranges. Note that the directions of the hysteresis loops are indicated by the solid arrows. (c) Variation of the threshold voltage shift under various $V_{CG\,max}$. (d) Memory window extracted from transfer characteristic curves depending on $V_{CG\,max}$. The inset presents the calculated stored carrier density n as a function of V_{CG}. (e) Retention performance of the HfS_2-based NVM device after being programmed (+80 V, 1 s) and erased (−80 V, 1 s), both program and erase states are read with $V_{DS} = 0.1$ V, $V_{CG} = 0$ V. (f) Retention endurance property over 1000 cycles of the memory device. Schematic energy band diagram of nonvolatile characteristic under (g) zero, (h) positive, and (i) negative control gate bias conditions, Φ and χ are the work function and the electron affinity, respectively.

learning, and memory in human brain. In biological, the presynaptic neuron can release excitatory or inhibitory neurotransmitters upon different stimulation. The delivery of neurotransmitters from presynaptic neuron to postsynaptic neuron dominates the information transmission between two neighboring neurons. Figure 3(a) shows the schematic diagram of a biological synapse and its emulation with a three-terminal HfS_2-based NVM device. As an artificial synapse, the electrical pulse applied to the control gate is appointed as the input spike, the I_{DS} recorded at $V_{DS} = 0.1$ V as the postsynaptic current. The excitatory PSC (EPSC) and inhibitory PSC are emulated in the artificial synapse under separate negative and positive input spikes with different amplitudes, as shown in Figs. 3(b) and 3(c), respectively. When a negative input spike is applied, the PSC shows a rapid increase and a slow decay. Moreover, the EPSC can be sustained at an intermediate state for a period of time, suggesting the nonvolatility of the state. Similarly, IPSC behavior corresponding to a synaptic inhibition activity is observed by applying a positive voltage pulse. In addition, from the response of PSC to input spikes of varying amplitudes, we can find that a pulse with larger amplitude induces a greater EPSC or IPSC behavior. This behavior can be attributed to the higher trapping and releasing capabilities of electron under larger V_{CG}, which agrees well with the increasing hysteresis with V_{CG} in Fig. 2(b). Moreover, a train of spikes of varying durations is applied to mimic the spike-duration-dependent plasticity

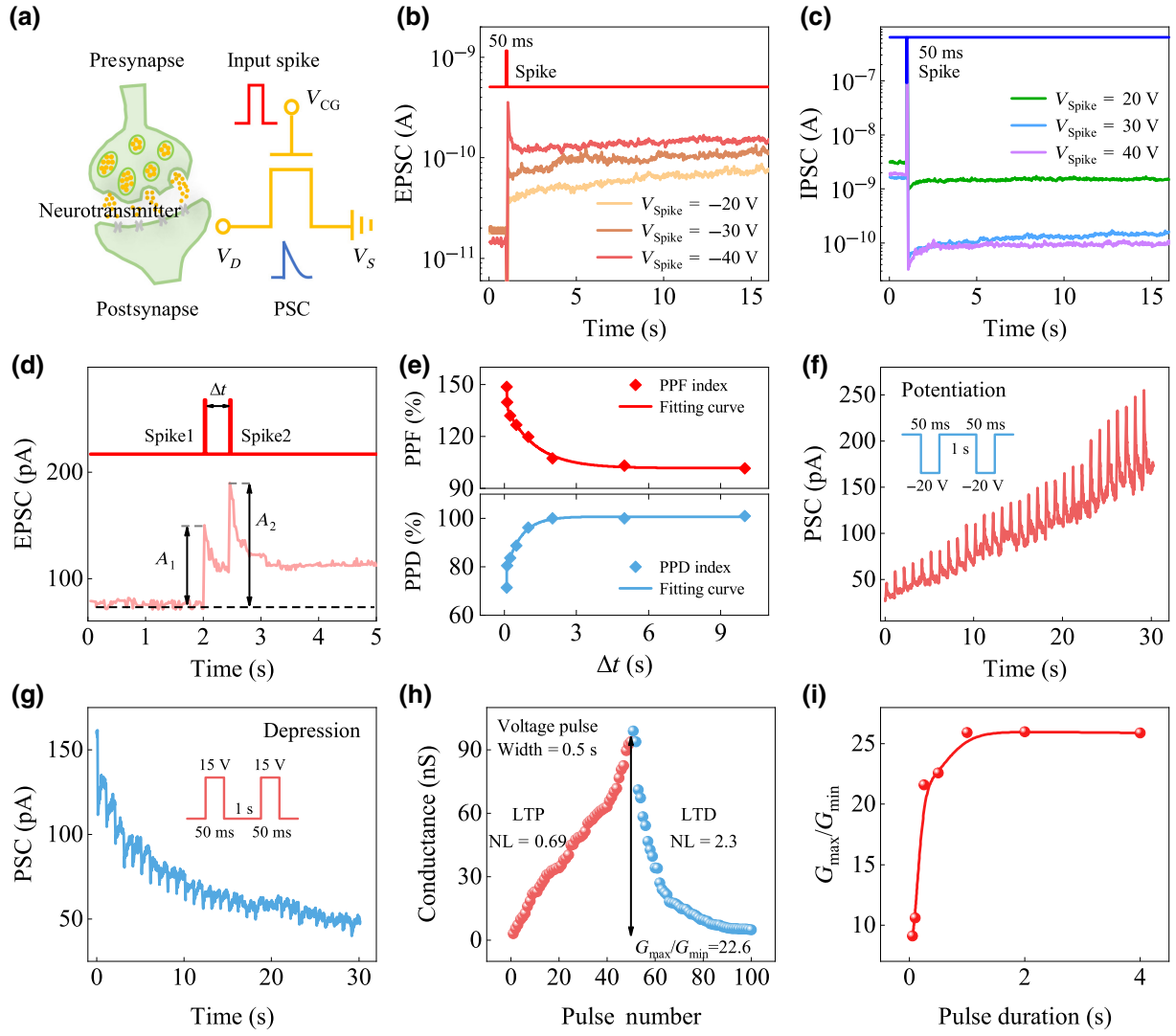

FIG. 3. (a) Schematic diagram of a biological synapse and its emulation with a three-terminal HfS_2-based NVM device. (b) Excitatory postsynaptic current (EPSC) generated by single-input spike with different amplitudes (−20, −30, and −40 V). (c) Inhibitory postsynaptic current (IPSC) triggered by positive spike with different pulse amplitudes (20, 30, and 40 V). Note that the input spike has a fixed width of 50 ms in (b) and (c). (d) EPSC behavior generated by two consecutive input spikes (−30 V, 50 ms) with an interval time of 0.5 s. (e) STP shown by the PPF and PPD index as a function of interval time between two successive input spikes (±30 V, 50 ms). (f) The dynamic long-term potentiation of PSC stimulated by a serious of input spikes (−20 V, 50 ms) with interval time of 1 s. (g) The dynamic long-term depression behavior in PSC under a series of voltage pulses (15 V, 50 ms) with an interval time of 1 s. (h) LTP and LTD synaptic behaviors of the artificial synapse in response to the application of 50 negative voltage pulses (−20 V, 100 ms) and 50 positive voltage pulses (15 V, 100 ms) with a period time of 5 s. (i) The extracted G_{max} : G_{min} ratio as a function of pulse duration.

(SDDP), as presented in Fig. S4 within the Supplemental Material [47]. The larger change in EPSC or IPSC can be induced with increasing the spike duration, demonstrating that more electrons are trapped or released for a long pulse duration.

In biology, the correlations between activity of presynaptic neuron and extent of postsynaptic membrane polarization induce variations in connection weights (i.e., synaptic plasticity). According to the retention ability, the synaptic plasticity can be divided into two categories: STP and LTP. The PPF is a representative type of STP, which is considered as a critical temporal information associated with learning and sound localization. The PPF index can be described by the following formula: PPF = $100\% \times A_2/A_1$, where A_1 and A_2 are the values of the first and second PSC, respectively. PPF behavior generated by a pair of negative input spikes (−30 V, 50 ms) with an interval time of 0.5 s is illustrated in Fig. 3(d). PPF represents a temporal synaptic activity increase when the interval time is very small. The extracted PPF and

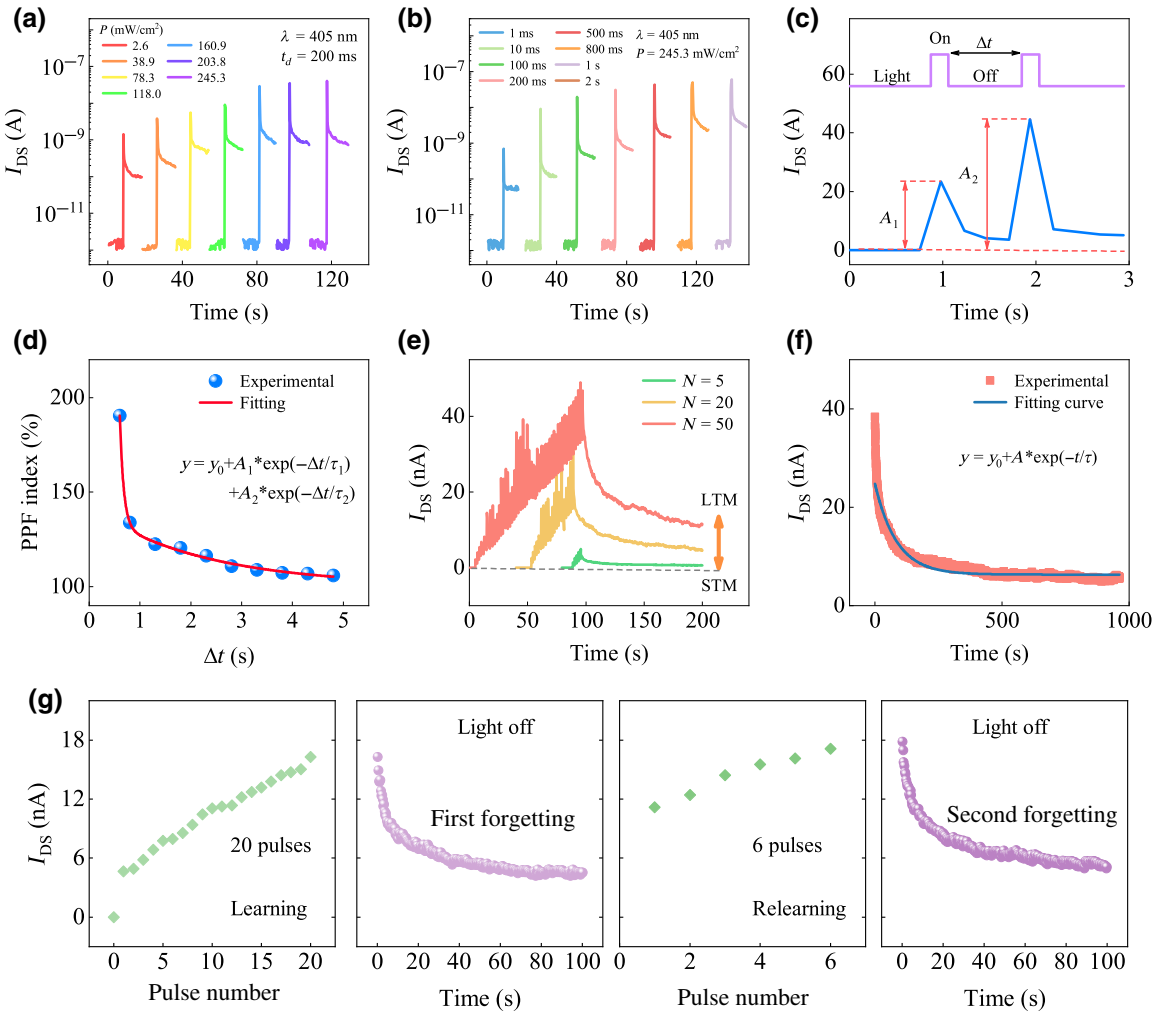

FIG. 4. (a) Transient photoresponsive characteristics of the NVM device under different light intensities from 2.6 to 245.3 mW/cm^2 for 200 ms with fixed wavelength of 405 nm. (b) Transient photoresponsive characteristics of the NVM induced by light pulses ($\lambda = 405$ nm, $P = 245.3$ mW/cm^2) with different exposure time (1 ms–2 s). (c) Typical photoresponsive characteristics of the NVM generated by a pair of light pulses (405 nm, 245.3 mW/cm^2, Δt=0.8 s). A_1 and A_2 represent the change of values at the end of the first and the second light pulse, respectively. (d) PPF index plotted as a function of interval time between light-pulse pairs, fitted by exponential curves. (e) The STM-to-LTM transition triggered by increasing the number of input light pulses. (f) Forgetting curve and fitted line using exponential relaxation model. The PSC achieved after 50 consecutive light pulses. (g) The measured "learning-experience" behavior emulated by pulsed light stimuli. Light pulse in (e)–(g): $\lambda = 405$ nm, $P = 245.3$ mW/cm^2; pulse width, 0.2 s; pulse frequency, 0.5 Hz.

paired pulse depression (PPD) index as a function of time interval between two pulses is fitted by a double exponential decay function and plotted in Fig. 3(e) [7]. The PPF index decreases from 150% to 101%, when Δt increases from 0.1 to 10 s. The calculated relaxation time constants ($\tau_1 = 20$ ms, $\tau_2 = 1.2$ s) consistent well with values in biological synapses. While the PPD index has an increasing trend from 71 to 100% with the increasing time interval. To obtain the long-term plasticity at excitatory synapses, including LTP and LTD, consecutive input spikes are applied to the control gate. Figures 3(f) and 3(g) present the evolution of PSC under 30 negative and positive consecutive pulses (amplitude = −20 V/15 V, width = 50 ms, interval = 1 s), respectively. Note that the conductance of the device is recorded with a reading voltage V_{DS} of 0.1 V. The PSC shows a persistent increase (decrease) with negative (positive) pulses. The modulated multilevel states of conductance can serve as synaptic weights for neuromorphic computing and bring about better learning capability and an improved network robustness. The LTP and LTD behaviors under a cluster of 50 identical consecutive negative and positive spikes are depicted in Fig. 3(h). Furthermore, the LTP and LTD behaviors under 50 successive potentiation and depression spikes with different pulse durations from 50 ms to 4 s are illustrated in Fig. S5 within the

Supplemental Material [47]. The observed on:off ratio between maximum conductance (G_{max}) and minimum conductance (G_{min}), i.e., G_{max}/G_{min} as a function of pulse duration is presented in Fig. 3(i). The $G_{max} : G_{min}$ ratio increases firstly and reaches saturation with increasing the pulse duration. The maximum of the $G_{max} : G_{min}$ ratio of 26 can be obtained when pulse duration is 1 s.

An artificial synaptic device with large $G_{max} : G_{min}$ ratio (> 10), multidata levels (> 32), and linear conductance modulation is essential for efficient neural network systems. The artificial synapse in this work exhibits 50 level conductance states and $G_{max} : G_{min} > 10$, demonstrating good potentiation and depression properties. The nonlinearity (NL) is extracted from the LTP and LTD characteristic curves to assess the potentiation and depression properties quantitatively. A lower NL of a synaptic device can result in a higher recognition accuracy in an ANN. The NL values for the potentiation and depression are analyzed by fitting LTP and LTD data with the following equations [44]:

$$G_{n+1} = G_n + \Delta G_p = G_n + \alpha_p e^{-\beta_p \frac{G_n - G_{min}}{G_{max} - G_{min}}},$$

$$G_{n+1} = G_n + \Delta G_d = G_n - \alpha_d e^{-\beta_d \frac{G_{max} - G_n}{G_{max} - G_{min}}},$$

where G_n and G_{n+1} are the conductance of the n^{th} and $n+1^{th}$ pulses, G_{max} and G_{min} represent the maximum and minimum conductance, α and β are the differences of conductance between two points and NL, respectively. In Fig. 3(h), extracted β_p and β_d from the fitting curves are 0.69 and 2.3, respectively. In Fig. S5 within the Supplemental Material [47], the effects of pulse width on the LTP and LTD characteristics are examined. The NL values for each pulse width are obtained to analyze the trend quantitatively as shown in Fig. S6 within the Supplemental Material [47]. The NL values show a decreasing trend firstly and increase with increasing the pulse duration from 50 ms to 4 s. The minimum NL values for LTP and LTD are 0.58 and 2.3, respectively. The large $G_{max} : G_{min}$ ratio and good linearity of our synaptic device suggest a potential in accomplishing high recognition accuracy.

In contrast to electrical signals, light can transmit without transmission loss, and can be employed as the input spike in the transistor with a light-sensitive channel. Figure 4 illustrates the basic photoresponsive characteristics of the HfS_2/*h*-BN/FG-graphene FET as an artificial optoelectronic synapse, in which the light pulse is regarded as the presynaptic input to trigger the PSC. Transient photoresponsive characteristics of the NVM device under different light intensities from 2.6 to 245.3 mW/cm^2 for 200 ms with fixed wavelength of 405 nm are shown in Fig. 4(a). Note that the PSC is recorded at a constant reading voltage of 0.1 V and gate voltage of 0 V. Light illumination with high intensity induces the large enhancement of the PSC. Figure 4(b) further illustrates the photoresponse characteristics with different pulse width from 1 ms to 2 s. Longer light pulses provide high cumulative excitation energy, leading to high photocurrent amplitudes. It is worth noting that our artificial optoelectronic synapse exhibits a clear response under a light-pulse width of 1 ms. The abrupt current rise is induced by a sudden increase of photogenerated carriers under light illumination. After the removal of the light stimuli, the PSC shows a gradual decay and back to the resting current. Therefore, the observed PSC behavior that a HfS_2-based optoelectronic transistor resembles is similar to the PSC in biological synapses. Moreover, the HfS_2-based optoelectronic transistor shows a similar trend when exposed to red light ($\lambda = 637$ nm, $P = 160$ mW/cm^2), as shown in Fig. S7 within the Supplemental Material [47].

The learning and forgetting behaviors of the human brain are dominated by its capability to handle the synaptic plasticity. PPF is a phenomenon in neuroscience where the PSC induced by the second presynaptic spike is larger than the first one. The enhancement extent is bound up with the time interval between these two consecutive spikes. The PSC peaks have been successfully observed in our light-stimulated synaptic transistor under two successive light pulses (405 nm, 245.3 mW/cm^2, 200 ms), as shown in Fig. 4(c). Obviously, the amplitude of the second PSC (A_2) triggered by the second light spike is larger than that induced by the first spike, which is analogous to the PPF behavior in the biological synapses. The observed PPF index as a function of the time interval (Δt) between two consecutive light pulses is plotted in Fig. 4(d). The PPF index decreases from 190 to 105%, when δt increases from 0.3 to 4.8 s. Note that the PPF index obtained under light stimulation is higher than that under voltage stimuli. In the fitting, the extracted relaxation time constants ($\tau_1 = 78$ ms, $\tau_2 = 2.2$ s) agree well with measured values in biological synapses.

In neurobiology, short-term and long-term memory (STM and LTM) are two typical memory behaviors. The information received with attention is converted into STM and stored in the hippocampus. The corresponding temporal weak potentiation of synaptic plasticity can only last from a few seconds to minutes. If repetitive rehearsals are provided, the STM can be transferred to LTM, which is stored into the cerebral cortex, and can persist from a few minutes to years. STM, LTM, as well as STM-to-LTM transition have been emulated using pulsed light stimuli with varied number (N), as illustrated in Fig. 4(e). The potentiation of synaptic plasticity (i.e., PSC) increases with the increase of light pulse number, transferring from STM to LTM. Similarly, as shown in Fig. S8 within the Supplemental [47], the STM-to-LTM transition can also been demonstrated under light stimuli ($\lambda = 637$ nm, $P = 160$ mW/cm^2) with the increase of pulse number (5, 20, and 50). In Fig. 4(f), we fit the forgetting curve by the formula known as the Kohlrausch law: $I(t) = I_0 + A\exp(-t/\tau)$, where I_0, A, and τ are

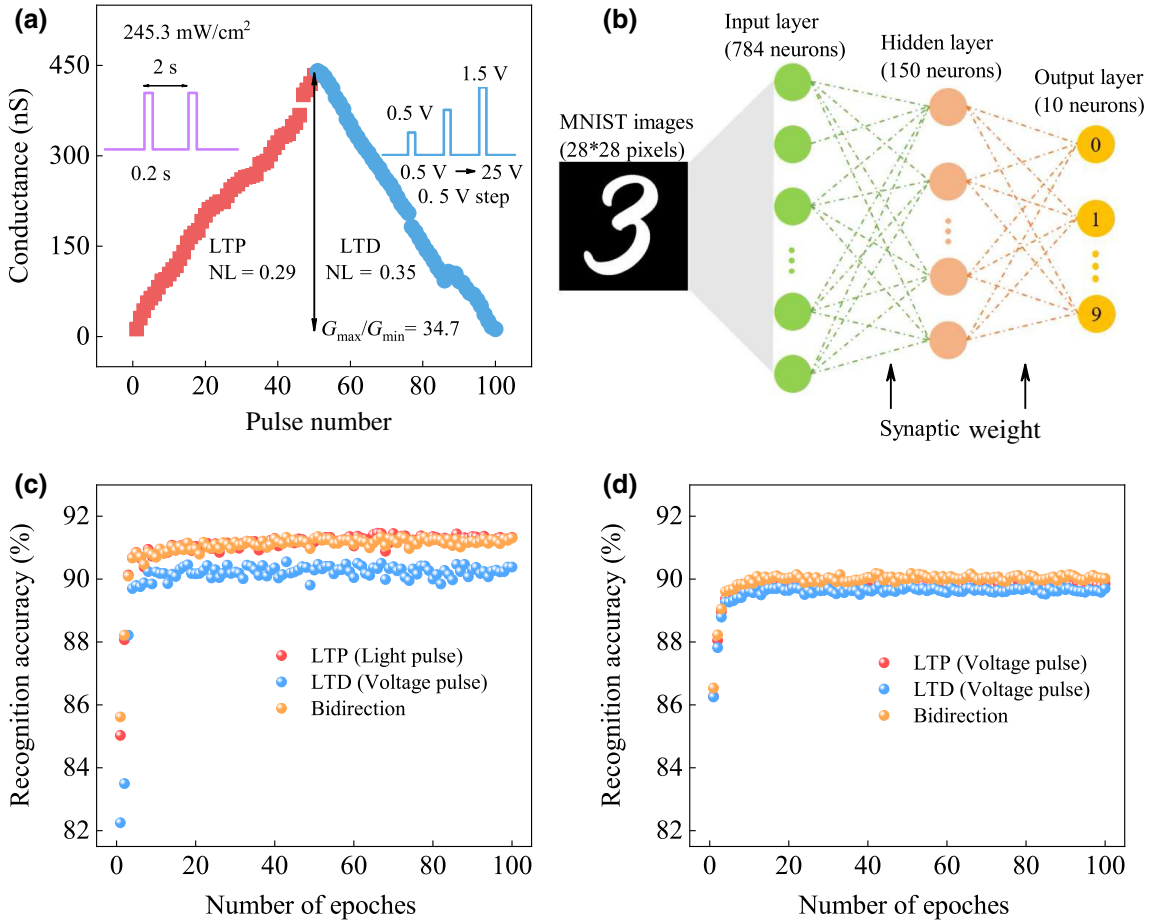

FIG. 5. (a) LTP characteristic curve measured under light pulse ($\lambda = 405$ nm, $P = 245.3$ mW/cm^2, pulse width 0.2 s, pulse frequency 0.5 Hz.) and LTD characteristic curve obtained under different voltage spike amplitudes from 0.5 to 25 V with a step of 0.5 V. (b) Schematic of a three-layer neural network for recognition tasks. (c) Recognition accuracy evolution with training epoches based on the LTP and LTD characteristic curves measured under light and voltage spikes, respectively. (d) The simulated recognition accuracy as a function of training epoches based on our artificial synapse measured under voltage pulse.

the stabilized current, prefactor, and the characteristic relaxation time, respectively. The characteristic relaxation time τ increases from 18 to 84 s, the stabilized current changes from 0.6 to 6.3 nA with pulse number (N), as shown in Fig. S9 within the Supplemental Material [47], which accord well with the consequent memory transition from STM to LTM in the human brain. Meanwhile, interesting "learning-experience" behavior of the human brain has been emulated using two sequences of successive light pulses with an interval of 100 s, as shown in Fig. 4(g). After the first pulse train (20 pulses), the synaptic weight has a notably potentiation and then shows a spontaneous decay to an intermediate level during the train interval. This behavior corresponds to the phenomenon that the information learned by the human brain is apt to lose after a period of time. In the second learning stage, to achieve the same current level (synaptic weight) as in the first train, only six pulses are required, which reflects that less time is needed for the relearning process. Moreover, the decay in synaptic weight after the relearning process is smaller than that after the first stimulating process, analogous to the learning behaviors in the human brain.

For neuromorphic computing, the linear and symmetric weight update of LTP and LTD, as well as large G_{max} : G_{min} ratio play a vital role in achieving high recognition accuracy and efficiency [3,51]. The LTP of our artificial synapse was obtained by applying 50 light pulses of width 200 ms (405 nm, $P = 245.3$ mW/cm^2), as shown in Fig. 5(a). The LTP curve has a large G_{max}/G_{min} ratio (34.7) and a small NL value of 0.29, due to the ultrahigh photoresponsivity and photogain of HfS_2 [35]. The LTD characteristic curve is investigated by applying a cluster of 50 voltage pulses with different amplitudes from 0.5 to 25 V with 0.5 V step. The NL of the LTD obtained under incremental voltage pulse scheme is improved compared with the case when voltage pulses with fixed amplitude are

applied. It is clear that our artificial synapse shows a higher conductance ratio and good linearity under the combination of electrical and optical singles, which can bring about higher recognition accuracy and efficiency in neuromorphic computing. Moreover, the energy consumption per spike in our artificial synapse is as low as 0.2 pJ, suggesting its application potential in low-power neuromorphic computing.

To fully demonstrate the potential of the HfS_2-based NVM device for neuromorphic computing, an ANN based on the experimentally measured LTP and LTD characteristic curves is simulated to perform supervised learning. As illustrated in Fig. 5(b), we design a two-layer ANN with one hidden layer using the PyTorch platform, which is an open-source Python machine learning library, for image classification from the MNIST database via back propagation [3]. The ANN generated by software consists of 784 input neurons, 150 hidden neurons, and 10 output neurons, in which the 784 input neurons and 10 output neurons correspond to the input image data with a size of 28×28 pixels and 10 kinds of Arabic digits (0-9), respectively. All neurons are connected to the neurons on the other end of corresponding layers through the synaptic device. The dynamic modulated weight connections among these neurons, are defined as LTP and LTD conduction states of HfS_2-based synaptic devices as shown in Fig. 5(a), which are extracted from experimental data under successive input spikes. During the simulation, the stochastic gradient descent (SGD) algorithm is employed for training. The detailed simulation process can refer to the investigation of Jin *et al.* [51]. The synaptic device updates their weighting values based on the HfS_2-based NVM device characteristics including the NL, $G_{max} : G_{min}$ ratio, finite number of conductance states. After training with 8000 randomly selected images per each epoch from 60 000 images of the training dataset, the recognition accuracy of ten digits is estimated by using 10 000 images from a separate testing data set, the model is updated by training with the full trainset of each epoch.

Figure 5(c) represents the recognition accuracy of the simulated ANN as a function of training epoch based on the LTP and LTD obtained under the combination of electrical and optical singles. The maximum recognition accuracies reach 91.5, 90.5, and 91.3% based on LTP only, LTD only, and LTP and LTD, respectively. The recognition accuracy is higher than previously reported ANNs with similar topology [44,51]. For comparison, the recognition accuracies of our simulated ANN based on the LTP and LTD curves observed under voltage-pulsing mode with fixed amplitude are plotted in Fig. 5(d). The estimated recognition accuracies are 89.9, 89.8, and 90.2% based on LTP only, LTD only, and LTP and LTD, respectively. Note that the recognition accuracy of the ANN based on our synaptic device can exceed 80% after only one training epoch. Our optoelectronic synapse exhibits a higher recognition accuracy under the combination of electrical and optical singles than that observed under electrical stimulation. The high recognition accuracy at the system level underscores the high potential of our HfS_2-based NVM device in designing neuromorphic computing systems. However, the synthesis of high-quality 2D materials, such as HfS_2, BP, *h*-BN, WSe_2, and MoS_2, at wafer scale is still in primary level. Therefore, it is very challenging to fabricate a large-scale array of 2D devices at the present stage. In this regard, it is vital to produce high-quality 2D materials at wafer scale with large grain sizes and uniform film thicknesses to realize high-performance fully hardware-implemented ANNs comprising 2D van der Waals synapses in the future.

IV. CONCLUSION

In conclusion, the NVM operation based on all 2D few-layer HfS_2/*h*-BN/FG-graphene FET devices is demonstrated. The NVM device exhibits reliable memory performances with a large memory window of 100 V, high on:off ratio over 10^5, stable switching cycles of 10^3, and a long retention time of 10^4 s. Notably, representative synaptic functions at the device level including PPF, STM and LTM, LTP and LTD have been emulated based on our HfS_2 NVM device. In particular, the ultrahigh photoresponsivity and photogain of the HfS_2 channel enable our devices to emulate memory and learning behavior similar to the human brain by tuning the light pulse width, intensity, and the numbers of light pulses. Moreover, the energy consumption per spike in our artificial synapse is as low as 0.2 pJ. The recognition accuracy of handwritten digits at the system level based on our HfS_2 NVM artificial synapse can be achieved up to 91.5%, which is superior to other reported 2D artificial optoelectronic synapses. The results suggest the high potential of all 2D heterostructures in developing both nonvolatile memories and efficient synaptic devices for neuromorphic computing systems.

ACKNOWLEDGMENTS

L. Xu and H. Xiong contributed equally to this work. The authors thank Dr. Mengjiao Li and Professor Peng Zhou for discussion. This work is financially supported by the National Natural Science Foundation of China (Grants No. 91833303, No. 61974043, No. 62074058, No. 62090013, and No. 61674057), the National Key Research and Development Program of China (Grant No. 2019YFB2203400), the Projects of Science and Technology Commission of Shanghai Municipality (Grants No. 21JC1402100, No. 18JC1412400, No. 18YF1407200, No. 18YF1407000, and No. 19511120100), the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning, Shanghai

Pujiang Program (20PJ1403600) and the Fundamental Research Funds for the Central Universities.

[1] P. Yao, H. Q. Wu, B. Gao, J. S. Tang, Q. T. Zhang, W. Q. Zhang, J. J. Yang, and H. Qian, Fully hardware-implemented memristor convolutional neural network, Nature **577**, 641 (2020).

[2] V. K. Sangwan and M. C. Hersam, Neuromorphic nanoelectronic materials, Nat. Nanotechnol. **15**, 517 (2020).

[3] Y. X. Hou, Y. Li, Z. C. Zhang, J. Q. Li, D. H. Qi, X. D. Chen, J. J. Wang, B. W. Yao, M. X. Yu, T. B. Lu, and J. Zhang, Large-scale and flexible optical synapses for neuromorphic computing and integrated visible information sensing memory processing, ACS Nano **15**, 1497 (2021).

[4] C. J. Wan, P. Q. Cai, M. Wang, Y. Qian, W. Huang, and X. D. Chen, Artificial sensory memory, Adv. Mater. **32**, 1902434 (2020).

[5] F. S. Yang, M. J. Li, M. P. Lee, I. Y. Ho, J. Y. Chen, H. F. Ling, Y. Z. Li, J. K. Chang, S. H. Yang, Y. M. Chang, K. C. Lee, Y. C. Chou, C. H. Ho, W. W. Li, C. H. Lien, and Y. F. Lin, Oxidation-boosted charge trapping in ultrasensitive van der waals materials for artificial synaptic features, Nat. Commun. **11**, 2972 (2020).

[6] T. Y. Wang, J. L. Meng, Z. Y. He, L. Chen, H. Zhu, Q. Q. Sun, S. J. Ding, P. Zhou, and D. W. Zhang, Ultralow power wearable heterosynapse with photoelectric synergistic modulation, Adv. Sci. **7**, 1903480 (2020).

[7] K. Wang, S. L. Dai, Y. W. Zhao, Y. Wang, C. Liu, and J. Huang, Light-stimulated synaptic transistors fabricated by a facile solution process based on inorganic perovskite quantum dots and organic semiconductors, Small **15**, 1900010 (2019).

[8] K. K. Lu, X. M. Li, Q. Q. Sun, X. C. Pang, J. Z. Chen, T. Minari, X. Y. Liu, and Y. L. Song, Solution-processed electronics for artificial synapses, Mater. Horiz. **8**, 447 (2021).

[9] S. L. Dai, Y. W. Zhao, Y. Wang, J. Y. Zhang, L. Fang, S. Jin, Y. L. Shao, and J. Huang, Recent advances in transistor-based artificial synapses, Adv. Funct. Mater. **29**, 1903700 (2019).

[10] Y. J. Park and J. S. Lee, Artificial synapses with short- and long-term memory for spiking neural networks based on renewable materials, ACS Nano **11**, 8962 (2017).

[11] P. Lin, C. Li, Z. R. Wang, Y. N. Li, H. Jiang, W. H. Song, M. Y. Rao, Y. Zhuo, N. K. Upadhyay, M. Barnell, Q. Wu, J. J. Yang, and Q. F. Xia, Three-dimensional memristor circuits as complex neural networks, Nat. Electron. **3**, 225 (2020).

[12] Z. H. Tan, R. Yang, K. Terabe, X. B. Yin, X. D. Zhang, and X. Guo, Synaptic metaplasticity realized in oxide memristive devices, Adv. Mater. **28**, 377 (2016).

[13] J. H. Yoon, Z. R. Wang, K. M. Kim, H. Q. Wu, V. Ravichandran, Q. F. Xia, C. S. Hwang, and J. J. Yang, An artificial nociceptor based on a diffusive memristor, Nat. Commun. **9**, 417 (2018).

[14] X. M. Zhang, Y. Zhuo, Q. Luo, Z. H. Wu, R. Midya, Z. R. Wang, W. H. Song, R. Wang, N. K. Upadhyay, Y. L. Fang, F. Kiani, M. Y. Rao, Y. Yang, Q. F. Xia, Q. Liu, M. Liu, and J. J. Yang, An artificial spiking afferent nerve based on mott memristors for neurorobotics, Nat. Commun. **11**, 51 (2020).

[15] F. C. Zhou, Z. Zhou, J. W. Chen, T. H. Choy, J. L. Wang, N. Zhang, Z. Y. Lin, S. M. Yu, J. F. Kang, H. S. Philip Wong, and Y. Chai, Optoelectronic resistive random access memory for neuromorphic vision sensors, Nat. Nanotechnol. **14**, 776 (2019).

[16] S. Ambrogio, P. Narayanan, H. Tsai, R. M. Shelby, I. Boybat, C. d. Nolfo, S. Sidler, M. Giordano, M. Bodini, B. C. P. Farinha, B. Killeen, C. Cheng, Y. Jaoudi, and G. W. Burr, Equivalent-accuracy accelerated neuralnetwork training using analogue memory, Nature **558**, 60 (2018).

[17] Z. G. Cheng, C. Rios, W. H. P. Pernice, C. D. Wright, and H. Bhaskaran, On-chip photonic synapse, Sci. Adv. **3**, e1700160 (2017).

[18] K. C. Kwon, Y. S. Zhang, L. Wang, W. Yu, X. J. Wang, I. Park, H. S. Choi, T. Ma, Z. Y. Zhu, B. B. Tian, C. L. Su, and K. P. Loh, In-plane ferroelectric tin monosulfide and its application in a ferroelectric analog synaptic device, ACS Nano **14**, 7628 (2020).

[19] M. K. Kim and J. S. Lee, Ferroelectric analog synaptic transistors, Nano Lett. **19**, 2044 (2019).

[20] L. Wang, X. J. Wang, Y. S. Zhang, R. L. Li, T. Ma, K. Leng, Z. Chen, I. Abdelwahab, and K. P. Loh, Exploring ferroelectric switching in α-In_2Se_3 for neuromorphic computing, Adv. Funct. Mater. **30**, 2004609 (2020).

[21] S. Y. Wang, L. Liu, L. R. Gan, H. W. Chen, X. Hou, Y. Ding, S. L. Ma, D. W. Zhang, and P. Zhou, Two-dimensional ferroelectric channel transistors integrating ultra-fast memory and neural computing, Nat. Commun. **12**, 53 (2021).

[22] S. L. Dai, X. H. Wu, D. P. Liu, Y. L. Chu, K. Wang, B. Yang, and J. Huang, Light-stimulated synaptic devices utilizing interfacial effect of organic field-effect transistors, ACS Appl. Mater. Interfaces **10**, 21472 (2018).

[23] S. G. Yi, M. U. Park, S. H. Kim, C. J. Lee, J. Kwon, G. H. Lee, and K. H. Yoo, Artificial synaptic emulators based on MoS_2 flash memory devices with double floating gates, ACS Appl. Mater. Interfaces **10**, 31480 (2018).

[24] A. J. Arnold, A. Razavieh, J. R. Nasr, D. S. Schulman, C. M. Eichfeld, and S. Das, Mimicking neurotransmitter release in chemical synapses via hysteresis engineering in MoS_2 transistors, ACS Nano **11**, 3110 (2017).

[25] L. Wang, W. G. Liao, S. L. Wong, Z. G. Yu, S. F. Li, Y. F. Lim, X. W. Feng, W. C. Tan, X Huang, L. Chen, L. Liu, J. S. Chen, X. Gong, C. X. Zhu, X. K. Liu, Y. W. Zhang, D. Z. Chi, and K. W. Ang, Artificial synapses based on multiterminal memtransistors for neuromorphic application, Adv. Funct. Mater. **29**, 1901106 (2019).

[26] M. Wang, S. H. Cai, C. Pan, C. Y. Wang, X. J. Lian, Y. Zhuo, K. Xu, T. J. Cao, X. Q. Pan, B. G. Wang, S. J. Liang, J. J. Yang, P. Wang, and F. Miao, Robust memristors based on layered two-dimensional materials, Nat. Electron. **1**, 130 (2018).

[27] S. G. Kim, S. H. Kim, J. Park, G. S. Kim, J. H. Park, K. C. Saraswat, J. Kim, and H. Y. Yu, Infrared detectable MoS_2 phototransistor and its application to artificial multilevel optic-neural synapse, ACS Nano **13**, 10294 (2019).

[28] S. Park, Y. Jeong, H. J. Jin, J Park, H. Jang, S. Lee, W. Huh, H. Cho, H. G. Shin, K. Kim, C. H. Lee, S. Choi, and S. Im, Nonvolatile and neuromorphic memory devices

using interfacial traps in two-dimensional $WSe_2/MoTe_2$ stack channel, ACS Nano **14**, 12064 (2020).

[29] X. X. Yang, J. R. Yu, J. Zhao, Y. H. Chen, G. Y. Gao, Y. F. Wang, Q. J. Sun, and Z. L. Wang, Mechanoplastic tribotronic floating-gate neuromorphic transistor, Adv. Funct. Mater. **30**, 2002506 (2020).

[30] S. Y. Wang, C. S. Chen, Z. H. Yu, Y. L. He, X. Y. Chen, Q. Wang, Y. Shi, D. W. Zhang, H. Zhou, X. R. Wang, and P. Zhou, A MoS_2-PTCDA hybrid heterojunction synapse with efficient photoelectric dual modulation and versatility, Adv. Mater. **31**, 1806227 (2019).

[31] C. Qian, S. Oh, Y. Choi, J. H. Kim, J. Sun, H. Huang, J. L. Yang, Y. L. Gao, J. H. Park, and J. H. Cho, Solar-stimulated optoelectronic synapse based on organic heterojunction with linearly potentiated synaptic weight for neuromorphic computing, Nano Energy **66**, 104095 (2019).

[32] C. Y. Wang *et al.*, Gate-tunable van der Waals heterostructure for reconfigurable neural network vision sensor, Sci. Adv. **6**, eaba6173 (2020).

[33] X. W. Feng, S. F. Li, S. L. Wong, S. W. Tong, L. Chen, P. P. Zhang, L. F. Wang, X. Y. Fong, D. Z. Chi, and K. W. Ang, Self-selective multi-terminal memtransistor crossbar array for in-memory computing, ACS Nano **15**, 1764 (2021).

[34] Y. v. d. Burgt, E. Lubberman, E. J. Fuller, S. T. Keene, G. C. Faria, S. Agarwal, M. J. Marinella, A. A. Talin, and A. Salleo, A non-volatile organic electrochemical device as a low-voltage artificial synapse for neuromorphic computing, Nat. Mater. **16**, 414 (2017).

[35] K. Xu, Z. X. Wang, F. Wang, Y. Huang, F. M. Wang, L. Yin, C. Jiang, and J. He, Ultrasensitive phototransistors based on few-layered HfS_2, Adv. Mater. **27**, 7881 (2015).

[36] S. Seo, S. H. Jo, S. Kim, J. Shim, S. Oh, J. H. Kim, K. Heo, J. W. Choi, C. Choi, S. Oh, D. Kuzum, H. S. Philip Wong, and J. H. Park, Artificial optic-neural synapse for colored and color-mixed pattern recognition, Nat. Commun. **9**, 5106 (2018).

[37] Y. Y. Shi, X. H. Liang, B. Yuan, V. Chen, H. T. Li, F. Hui, Z. Yu, F. Yuan, E. Pop, H. S. Philip Wong, and M. Lanza, Electronic synapses made of layered two-dimensional materials, Nat. Electron. **1**, 458 (2018).

[38] L. P. Xu, P. Zhang, H. N. Jiang, X. Wang, F. F. Chen, Z. G. Hu, Y. J. Gong, L. Y. Shang, J. Z. Zhang, K. Jiang, and J. H. Chu, Large-scale growth and field-effect transistors electrical engineering of atomic-layer SnS_2, Small **15**, 1904116 (2019).

[39] L. P. Xu, Z. H. Duan, P. Zhang, X. Wang, J. Z. Zhang, L. Y. Shang, K. Jiang, Y. W. Li, L. Q. Zhu, Y. J. Gong, Z. G. Hu, and J. H. Chu, Ferroelectric-modulated MoS_2 field-effect transistors as multilevel nonvolatile memory, ACS Appl. Mater. Interfaces **19**, 44902 (2020).

[40] Y. Wang, Z. Y. Lv, J. R. Chen, Z. P. Wang, Y. Zhou, L. Zhou, X. L. Chen, and S. T. Han, Photonic synapses based on inorganic perovskite quantum dots for neuromorphic computing, Adv. Mater. **30**, 1802883 (2018).

[41] T. Ahmed, S. Kuriakose, E. L. H. Mayes, R. Ramanathan, V. Bansal, M. Bhaskaran, S. Sriram, and S. Walia, Optically stimulated artificial synapse based on layered black phosphorus, Small **15**, 1900966 (2019).

[42] S. Gao, G. Liu, H. L. Yang, C. Hu, Q. L. Chen, G. D. Gong, W. H. Xue, X. H. Yi, J. Shang, and R. W. Li, An oxide Schottky junction artificial optoelectronic synapse, ACS Nano **13**, 2634 (2019).

[43] T. Ahmed, S. Kuriakose, S. Abbas, M. J. S. Spencer, M. Ataur Rahman, M. Tahir, Y. R. Lu, P. Sonar, V. Bansal, M. Bhaskaran, S. Sriram, and S. Walia, Multifunctional optoelectronics via harnessing defects in layered black phosphorus, Adv. Funct. Mater. **29**, 1901991 (2019).

[44] T. Ahmed, M. Tahir, M. X. Low, Y. Y. Ren, S. A. Tawfik, E. L. H. Mayes, S. Kuriakose, S. Nawaz, M. J. S. Spencer, H. Chen, M. Bhaskaran, S. Sriram, and S. Walia, Fully light-controlled memory and neuromorphic computation in layered black phosphorus, Adv. Mater. **33**, 2004207 (2020).

[45] Z. K.. Zhang, D. H. Meng, F. M. Bai, J. Y. Zhai, and Z. L. Wang, Photon-memristive system for logic calculation and nonvolatile photonic storage, Adv. Funct. Mater. **30**, 2002945 (2020).

[46] J. D. Gong, H. Y. Yu, X. Zhou, H. H. Wei, M. X. Ma, H. Han, S. Zhang, Y. Ni, Y. L. Li, and W. T. Xu, Lateral artificial synapses on hybrid perovskite platelets with modulated neuroplasticity, Adv. Funct. Mater. **30**, 2005413 (2020).

[47] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevApplied.16.044049 for more details about surface topography AFM images of graphene and *h*-BN flakes; transport characteristics of the device with the structure of HfS_2/*h*-BN/silicon; the repeatability of the HfS_2-based NVM; The spike-duration-dependent plasticity (SDDP), LTP and LTD behaviors, and nonlinearity under voltage spikes; transient photoresponsive characteristics and STM-to-LTM transition of the NVM induced by light pulses ($\lambda = 637$ nm, $P = 156$ mW/cm^2).

[48] B. Mukherjee, A. Zulkefli, K. Watanabe, T. Taniguchi, Y. wakayama, and S. Nakaharai, Laser-assisted multilevel non-volatile memory device based on 2D van-der-waals few-layer-ReS_2/h-BN/Graphene heterostructures, Adv. Funct. Mater. **30**, 2001688 (2020).

[49] H. Liu, A. T. Neal, and P. D. Ye, Channel length scaling of MoS_2 MOSFETs, ACS Nano **10**, 8563 (2012).

[50] L. Britnell, R. V. Gorbachev, R. Jalil, B. D. Belle, F. Schedin, M. I. Katsnelson, L. Eaves, S. V. Morozov, A. S. Mayorov, N. M. R. Peres, A. H. Castro Neto, J. Leist, A. K. Geim, L. A. Ponomarenko, and K. S. Novoselov, Electron tunneling through ultrathin boron nitride crystalline barriers, Nano Lett. **12**, 1707 (2012).

[51] T. Y. Jin, Y. Zheng, J. Gao, Y. N. Wang, E. L. Li, H. P. Chen, X. Pan, M. Lin, and W. Chen, Controlling native oxidation of HfS_2 for 2D materials based flash memory and artificial synapse, ACS Appl. Mater. Interfaces **13**, 10639 (2021).