iScience

CellPress
OPEN ACCESS

Perspective

Short-term synaptic plasticity in emerging devices for neuromorphic computing

Chao Li,[1,2,5] Xumeng Zhang,[1,3,4,*] Pei Chen,[1] Keji Zhou,[1,3,4] Jie Yu,[1,3] Guangjian Wu,[1,3,4] Du Xiang,[1,3,4,*] Hao Jiang,[1,3,4] Ming Wang,[1,3,4] and Qi Liu[1,3,4,*]

SUMMARY

Neuromorphic computing is a promising computing paradigm toward building next-generation artificial intelligence machines, in which diverse types of synaptic plasticity play an active role in information processing. Compared to long-term plasticity (LTP) forming the foundation of learning and memory, short-term plasticity (STP) is essential for critical computational functions. So far, the practical applications of LTP have been widely investigated, whereas the implementation of STP in hardware is still elusive. Here, we review the development of STP by bridging the physics in emerging devices and biological behaviors. We explore the computational functions of various STP in biology and review their recent progress. Finally, we discuss the main challenges of introducing STP into synaptic devices and offer the potential approaches to utilize STP to enrich systems' capabilities. This review is expected to provide prospective ideas for implementing STP in emerging devices and may promote the construction of high-level neuromorphic machines.

INTRODUCTION

Building intelligent machines have long been a vision, especially with the advent of the intelligent era. The emergence of edge computing, big data, and other new technologies has significantly promoted the development of the intelligent era, but the growing volume of data and the demand for real-time processing of edge tasks pose a great challenge to conventional computers. The classical computing architecture separates the data storage from the central processing unit, generating "walls" that prevent further optimization of power consumption and computational speed, particularly in data-intensive tasks. The human brain is compact and unique for cognitive tasks and computing in memory with power dissipation as low as 20 W, whereas the Fugaku supercomputer requires 28 MW. Therefore, simulating the information processing of the human brain is an effective way to achieve high-speed and power-efficient data handling. The highly efficient biological nervous system originates from the huge connection network between neurons enabling the highly parallel processing ability of brain.[1,2] As the fundamental unit for signal transmission and regulation in the neural network,[3] synapses are considered to play essential roles in the realization of brain features, which are more than analog weights.[4] It is noted that synapses are leaky memories which possess different timescales and state parameters ruling their modifications.[5,6] On the other hand, synapses are also highly stochastic since the received spikes are transmitted in an uncertain manner.[7] Basically, the functional connections in neural networks are determined by the dynamic regulation of synaptic plasticity.[8]

As a rule of thumb, synaptic plasticity can be simply divided into long-term plasticity (LTP)[9–11] and short-term plasticity (STP)[6,12–14] according to the timescale. LTP refers to the "permanent" change in synaptic connection strength, which is generally related to the formation of long-term memory underpinning the learning and memory functions.[15] On the contrary, STP denotes the temporary efficacy change (milliseconds to seconds) during calculation, possessing the time window in good agreement with many computing tasks, such as speech recognition,[16] information filter,[17] working memory,[18] and spatial orientation perception.[19] Therefore, STP provides a neural basis to achieve higher cognitive functions in the brain. LTP has been widely emulated both at device and system levels.[20–24] In detail, the synaptic devices with LTP characteristics are usually programmed into various weights to conduct vector-matrix multiplication in an array structure, resulting in the significant acceleration of the computing process in artificial neural networks (ANNs). In contrast, the implementation of STP remains limited in behavioral emulation at the device level,

[1]State Key Laboratory of Integrated Chip and System, Frontier Institute of Chip and System, Fudan University, Shanghai 200433, China

[2]Key Laboratory of Microelectronics Device & Integrated Technology, Institute of Microelectronics of Chinese Academy of Sciences, Beijing 100029, China

[3]Zhangjiang Fudan International Innovation Center, Fudan University, Shanghai 200433, China

[4]Shanghai Qi Zhi Institute, Shanghai 200232, China

[5]University of Chinese Academy of Sciences, Beijing 100049, China

*Correspondence: xumengzhang@fudan.edu.cn (X.Z.), xiang_du@fudan.edu.cn (D.X.), qi_liu@fudan.edu.cn (Q.L.)

https://doi.org/10.1016/j.isci.2023.106315

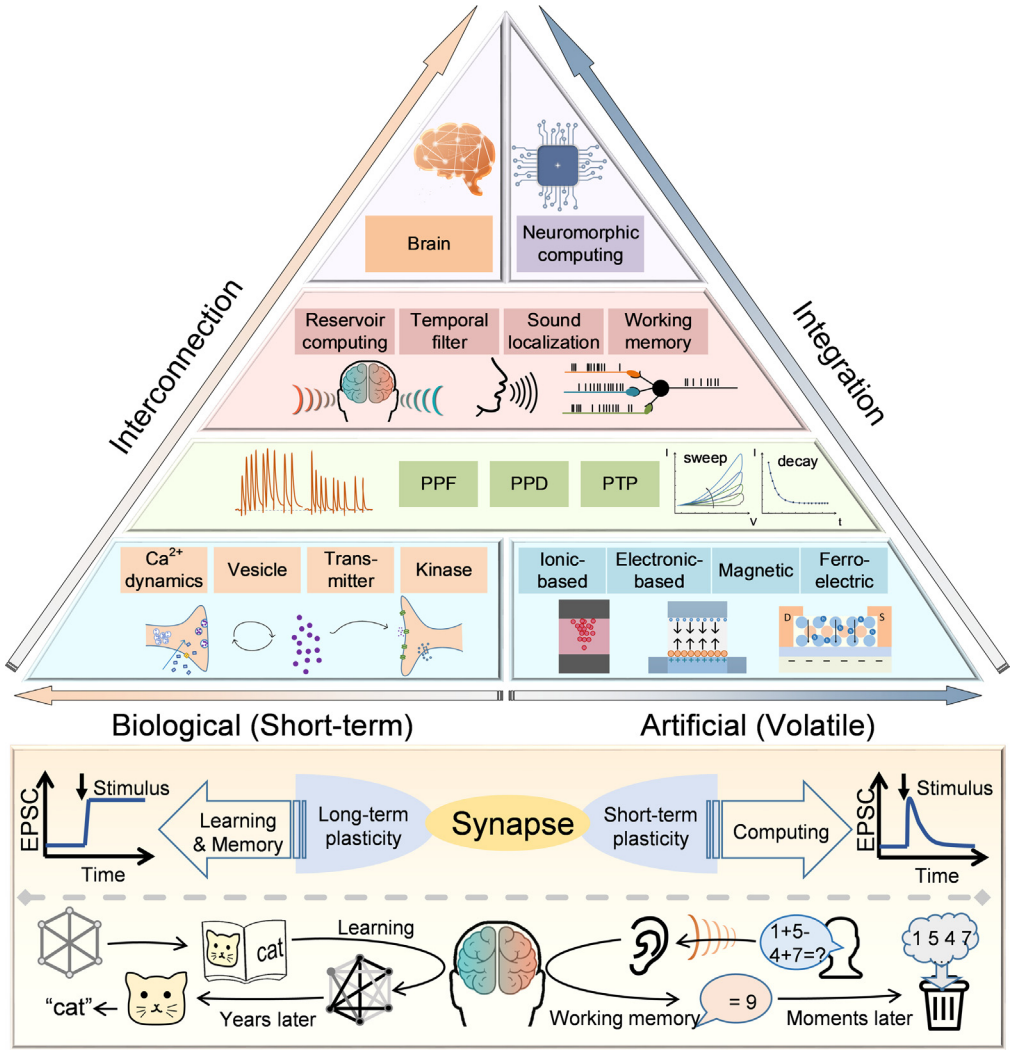

Figure 1. Functions of synaptic plasticity

Top panel: schematic diagram of the STP correspondence between biological and emerging artificial synapses, from the basic mechanisms to system implementation. Bottom panel: the behavioral and functional differences between LTP and STP in biological synapses.

whereas the prototype at the system level is still unexplored. One of the key challenges is that the STP-correlated mature computational model and network algorithms are lacking, which severely hinders the system design.

Complementary metal-oxide-semiconductor (CMOS) devices were initially used to achieve the temporary dynamic features in STP by exploiting their subthreshold characteristics.[25–27] For example, C. Bartolozzi et al. utilized diff-pair integrator synapse to reproduce the dynamics in STP, where the V_{thr} is introduced to achieve the dynamic tuning of STP timescale.[28] Although STP emulation can be realized with CMOS devices, the artificial synapses suffer from severe challenges in energy- and power-efficiency since they generally consist of several transistors and bulky capacitors. Moreover, to further control the temporal dynamics in such synaptic unit flexibly, complex bias circuits are required to adjust the working state of the subthreshold circuit.[29] Compared to the traditional CMOS devices, the emerging devices with novel physical mechanisms, e.g. ion-migration, electron trapping, ferroelectric polarization, and magnetic skyrmion motion, show intriguing potential to emulate the dynamic behaviors of synapses in a single unit with less area and energy consumption. Figure 1 schematically illustrates the correspondence between emerging artificial synapses and biological synapses from the fundamental mechanisms to system

applications. In biology, synaptic efficacy can be modulated by the interactions between Ca^{2+} and neurotransmitters, kinases, and so forth. In contrast, the conductance in artificial synapses can be controlled by tuning the movement of ions or other physics, which simulates the modulation of synaptic efficacy. The hysteresis in the *I*—*V* curves together with the conductance increment under successive excitation indicates the memory effect,[30] while the spontaneous decay of conductance is consistent with the feature of short-term property. Attributing to the intrinsic dynamics processes, such as the migration of metal ions, the changed conductance simplifies the emulation of ion-induced STP in biological synapses. Devices based on different physical mechanisms can be used as artificial synapses with distinctive features, forming the basis for mimicking various advanced short-term synaptic behaviors, such as pair-pulse facilitation (PPF), pair-pulse depression (PPD), and post-tetanic potentiation (PTP). These features could in turn enable higher-order brain functions such as sound localization,[31] associative learning,[32] and working memory.[33] Thus, implementing STP with the emerging mechanisms is the foundation for pursuing neuromorphic computing with neuroscience principles, which needs further investigation.

This article reviews the current advances in the field of short-term artificial devices and their applications in neuromorphic computing. First, an overview of the synaptic bio-kinetic process and physics for implementing artificial synapses are presented. After that, recent advances in STP implementation using physical devices are summarized from the perspective of emulating synaptic behaviors. Next, we overview the realization of the STP system by using the emerging synaptic devices. Lastly, we discuss the challenges in developing short-term devices and propose potential routes. We expect that this review could provide fundamental guidance and inspiration for implementing STP in physics and system applications, thus toward a high-order neuromorphic machine.

SHORT-TERM PLASTICITY PHYSICS IN BIOLOGICAL AND ARTIFICIAL DEVICES

In implementing short-term artificial synapses, it is essential to understand the internal correlation between biological and physical dynamics. In this section, the biochemical dynamics of STP is first discussed, followed by the introduction of some representative physical mechanisms in artificial devices which can be used for short-term implementation.

Biological short-term mechanisms and models

Figure 2A shows the schematic of a biological synapse, in which the presynaptic neuron and postsynaptic neuron are separated by a 20-40 nm gap known as synaptic cleft. Presynaptic nerve cells generally influence the excitability of postsynaptic neuron by releasing neurotransmitters that are stored in the vesicle pools in presynaptic terminals. When the presynaptic terminal receives stimuli, these transmitters are released toward the postsynaptic terminals. The postsynaptic membrane contains a large number of neurotransmitter receptors that could react to the released transmitters and mediate the postsynaptic response.[34] It is noted that the release of transmitters is regulated by the calcium ions (Ca^{2+}) in pre-synapses. The detailed working flow is described later in discussion:

(i) Action potentials (APs) reaching the terminal of a presynaptic axon could open the voltage-gated calcium channels (VGCCs) within an active zone;

(ii) The concentration of Ca^{2+} increases near the active zone, which in turn causes vesicles containing neurotransmitters to fuse with the presynaptic cell membrane and release the transmitters into the synaptic cleft;

(iii) The terminals of post-synaptic dendrites receive the transmitters and lead to transient postsynaptic membrane potential.

The above processes only induce a transient synaptic response and do not lead to any permanent change of the synaptic weights, which is therefore denoted as STP. We know that the kinetics of Ca^{2+} are abundant, which form the basis for various manifestations of STP, as schematically shown in Figure 2B. Take the excitatory synapse as an example, with a low neurotransmitter releasing probability, if the presynaptic neuron is repeatedly stimulated, the excitatory postsynaptic current (EPSC) usually increases gradually during this period. This excitatory behavior usually includes PPF and PTP. PPF is observed under two presynaptic APs with short interval, as evidenced by an increase in the postsynaptic current induced by the second action potential. During the first AP, Ca^{2+} flows in through VGCCs until the channels close. After that, the residual Ca^{2+} remaining in the terminal is gradually pumped out, leading to the reduction of the

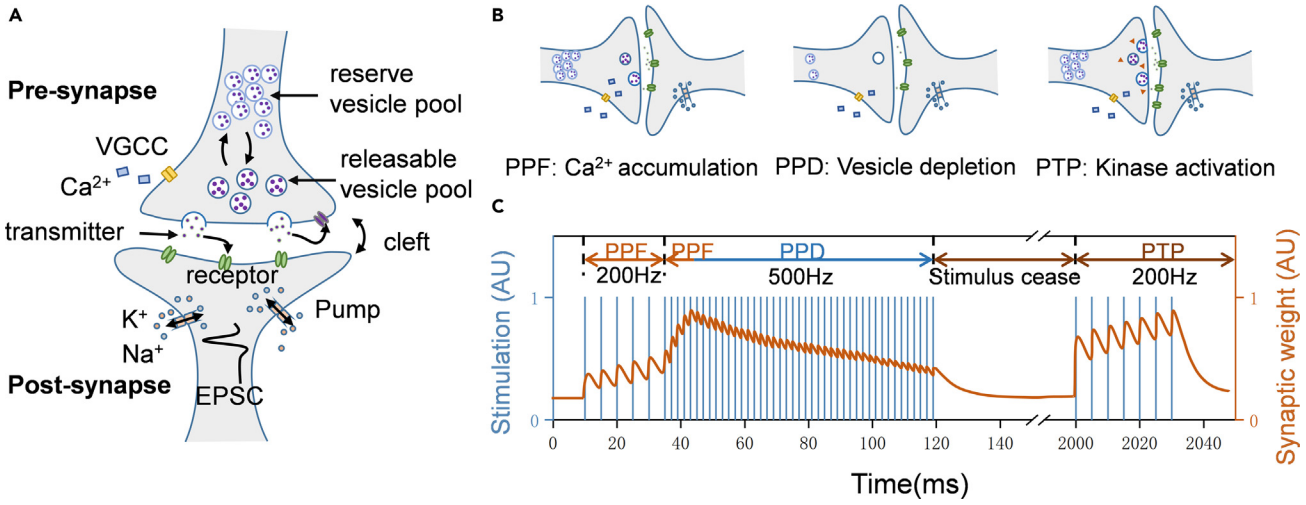

Figure 2. Biological synapse and STP behaviors

(A) Schematic of a biological synapse: spike input from presynaptic neuron causes the injection of Ca^{2+}, resulting in the release of vesicles. These vesicles are rapidly replenished by a readily releasable vesicle pool. Most vesicles are stored in a large reserve pool and can be utilized to refill the releasable pool. Released transmitters are recovered by endocytic process into releasable and reserve pool.

(B) STP behaviors, including PPF, PPD, and PTP. PPF is proposed to originate from the accumulation of calcium ions in response to multiple presynaptic stimuli, leading to increased vesicle release. PPD is associated with the depletion of vesicles, since the releasable pool cannot recover in time to cause a decrease in neurotransmitter release under high-frequency stimulation. PTP might be influenced by the activation of protein kinases, which promotes vesicle-membrane fusion and leads to longer timescale in response to tonic stimuli.

(C) Schematic diagram of STP response. PPF and PTP could strengthen the synaptic weights, while PPD inhibits the synaptic weights. PTP exhibits experience-dependent behavior in comparison to PPF.

transmitters and thus the decay of EPSC. When a second action potential arrives before the vanish of the first EPSC, the releasing probability of transmitters is enhanced by the extra Ca^{2+}, resulting in stronger response and higher EPSC, which is defined as the PPF behavior.[35] The PPF generally works within several hundred milliseconds.[36] In contrast to PPF, the generation of PTP usually requires tens or even hundreds of repeated stimuli, which can maintain for tens of seconds, with the retention time much longer than PPF (Figure 2C). PTP generally acts through enhanced vesicle fusion,[37] in which the presynaptic activity induces the residual calcium for a longer period of time.[14] In addition to enhancement, the depression effect is also commonly observed in synapses with a high probability of neurotransmitter release. Specifically, during repeated activation, the postsynaptic amplitude of the electrical potential appears to decrease. In contrast to PPF, paired-pulse depression (PPD) suppresses the postsynaptic response when a second stimulating spike closely follows the first one.[36] PPD can be clarified by multiple mechanisms, both pre and postsynaptic mechanisms have been proposed, including presynaptic depletion of vesicles, inactivation of release sites, inhibition or inactivation of calcium channels[6,38] and desensitization or receptor saturation of postsynaptic receptors.[35]

The electrophysiological phenomena associated with STP are commonly observed in experiments, whereas its computational importance is often overlooked. To elucidate brain functions, it is essential to understand the dynamical properties of neural systems and their roles in neural computations.[39] With the development of theories such as statistical physics, nonlinear dynamics, and complex systems, STP models are progressively improved in computational neurology. In the 1990s, Abbott et al.[40] and Tsodyks et al.[41–43] proposed phenomenal models of STP, mimicking the postsynaptic current generated by both PPF and PPD effects. All these models indicate that the dynamics of synaptic transmission result in complex sets of regular and irregular regimes of network activity. In these models, the factors including vesicle resources, Ca^{2+} concentration, and neurotransmitter release probability are crucial to elicit postsynaptic neural activities. The simplified STP differential model is displayed as follows[14,41]:

$$\frac{dx}{dt} = \frac{1 - x}{\tau_D} - ux\delta\left(t - t_{sp}\right) \qquad \text{(Equation 1)}$$

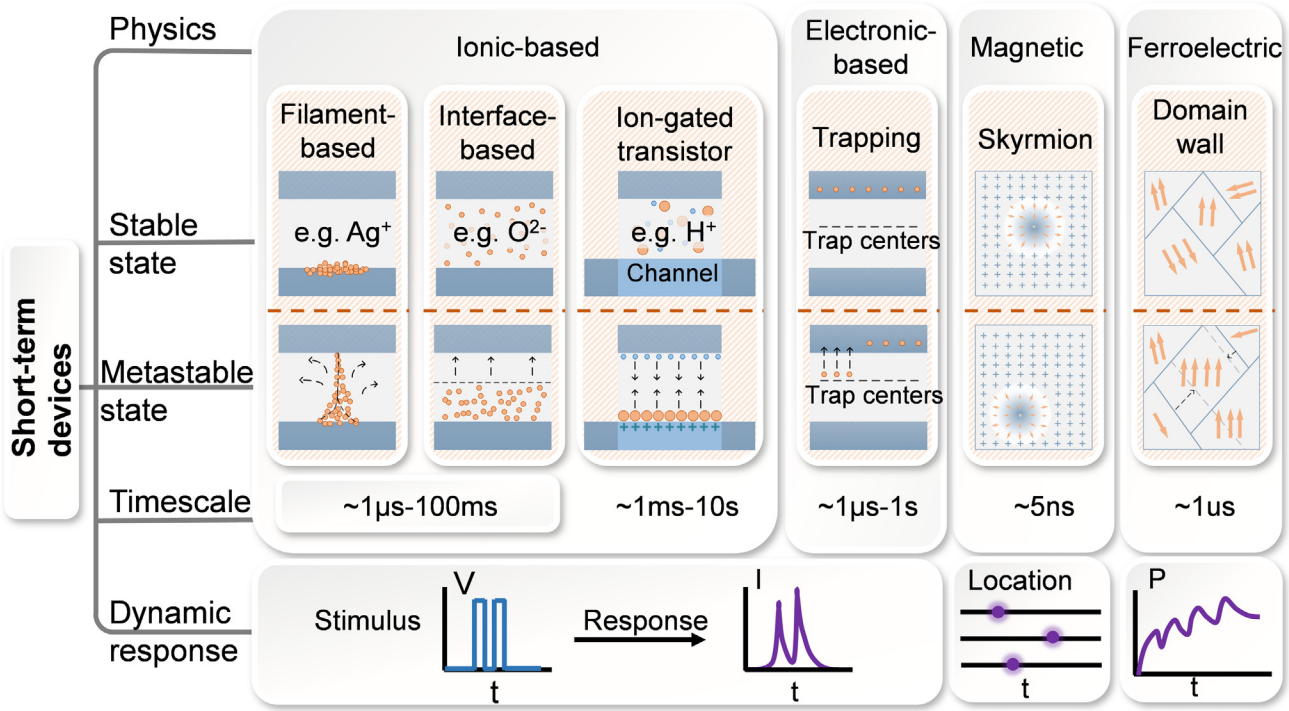

Figure 3. Physics for emerging short-term artificial synaptic devices

Short-term mechanisms are presented in six panels. The metastable state indicates the transition in the device under the applied stimulus. The dash lines represent the physical processes of returning to the stable states upon removal of the stimuli. The timescale refers to the duration of the metastable state. The dynamic response characterizes the metric change in the device under pulsed stimulation. For ionic- and electronic-based devices, the response can be expressed by the change of the current under pulses. The position of skyrmions and polarization represent the dynamic response of the magnetic- and ferroelectric-based devices respectively.

$$\frac{du}{dt} = \frac{U - u}{\tau_F} + U(1 - u)\delta(t - t_{sp}) \qquad \text{(Equation 2)}$$

where x represents normalized vesicle resources responding to synaptic efficacy, u is the normalized release probability. δ indicates the difference between time t and spike release time t_{sp}. The time constant τ_F and τ_D characterizes the time required for facilitation and depression to recover the baseline levels of 1 and U, respectively. Equations 1 and 2 correspond to the facilitation and inhibition processes, respectively, where the former mimics neurotransmitter depletion and the latter mimics Ca^{2+} influx into the presynaptic terminal to influence the release probability. Based on the mutual balance of facilitation and inhibition determined by τ, a variety of short-term behavior can be described phenomenologically. This model provides a basic reference for engineering STP devices and thus for chip design.

Short-term physical mechanisms in artificial synaptic devices

In general, the synaptic weights are described as synaptic efficacy which is represented by the conductance in artificial synapses. The performance requirements and resistance switching mechanisms of LTP are widely reviewed in emerging devices for in-memory computing,[44–49] whereas the significance of STP is rarely investigated. Herein, we summarized the fundamental mechanisms of short-term plasticity in emerging artificial synaptic devices.

The implementation of STP utilizing emerging devices generally relies on their short-term dynamic properties, that is, the capability of returning to or shifting toward the initial state spontaneously after removing the stimulus. According to the switching mechanisms, the physics for short-term dynamics could be roughly divided into four types which are ionic-, electronic-, magnetic-, and ferroelectric-based, as shown in Figure 3. In the following section, we will discuss all these different short-term devices in detail.

Ionic-based devices

The transition of ionic-based devices generally accompanies by a large number of mobile ions, including cations (e.g., metal ions) or anions (e.g., oxygen ions). Triggered by the electric field, ions can migrate inside the functional layer, which changes the device's conductance. When the activation is removed, the ions gradually degrade under thermal or gradient of concentration effects, resulting in the recovery of conductance.

Under the motivation of the electric field, ions such as Ag^{+} and O^{2-} can be driven to form the filament, as shown in the first panel in Figure 3. Spontaneous rupture of the filaments occurs following the disappearance of the stimulus. Theoretical frameworks such as the minimization of interfacial energy of atomic clusters,[50] minimization of filament surface area by surface diffusion of metal ions,[51] and local heat distribution[52,53] have been developed to clarify such spontaneous rupture. It is proposed that the timescale of the short-term plasticity in ions-based devices could be controlled by carefully designing the dielectrics, on the basis of the "easy coming, easy go" statement.

Ions in the interface-based devices accumulate at the dielectric interface under the electric field, reducing the potential barrier width and thus enhancing the tunneling current, which could effectively modulate the conductance.[54] For example, in oxide-based memristors,[55,56] the O^{2-} shift toward the bottom electrode interface through applying the electric field on the top electrode, as shown in the second panel in Figure 3. After removal of the stimulation, the ions will move to the steady state arising from the ion concentration gradient, resulting in a volatile conductance change.[57] The time constant of the interface-based devices generally ranges from 1μs to 100 ms.[55,56]

Distinct from filament- and interface-based devices, ion migration in electrolyte-gated transistors (EGTs) occurs in the electrolyte, a process that changes the channel's conductance. The EGTs possess the architecture superiority of control terminal (gate) and transduction terminals (source-drain),[49] allowing for the control and read of the conductance simultaneously. EGTs can be generally divided into two categories: electrochemical doping (ED) and electric double layer (EDL). In the ED devices, gate voltage could drive the ions, such as Li^{+}, H^{+}, OH^{-} and organic ions, in the electrolyte to inject into the channel and thus change the channel's conductance.[58] In contrast, the ions in EDL devices mainly accumulate at the interface, which makes the device recover quickly after the removal of the stimulus, as shown in the third panel in Figure 3. As reported, the EDL devices are more favorable to perform short-term dynamics with the timescale of ∼s.[8] Nevertheless, the small dynamic ranges of current EGTs limit their further applications.[59] To improve the short-term properties, the interaction between injected ions and the electrolyte materials should be carefully designed. Additionally, implementing the controlled short-term dynamics at device level is desirable, which needs further investigation.

Electronic-based devices

During the material synthesis and device fabrication process, defects including vacancies, interstitials, precipitates, dislocations, and grain boundaries might emerge either intentionally or unintentionally. These defects can act as trapping centers to capture and release carriers. Under stimulation, the trapping sites are filled with carriers, and thus the trap-assisted tunneling or thermal excitation current is strengthened, leading to an increase in device conductance, as illustrated in fourth panel in Figure 3. At the end of the excitation, the carriers in the trapping sites are gradually released or recombined, suggesting the volatile characteristics. The defects-assisted short-term plasticity is more pronounced through the introduction of nanoparticles in materials which allows for more trapping sites.[60] In comparison to ions-migration-based devices, the capture and release of carriers demand less energy, thus might be more favorable for energy-efficient applications.

In addition to thin films, two-dimensional (2D) materials in porous structures have also been investigated for short-term synapses. The interface states and defect sites in 2D materials can be modified through various physical or chemical treatments such as post-annealing or chemical doping,[61] which is beneficial for achieving STP. More intriguingly, 2D materials with suitable bandgap demonstrate outstanding light sensitivity, which greatly broadens the application area of artificial synapses and provides convenience for integrated sensory storage and computational systems. Meanwhile, it is more advantageous to scale down the size of 2D devices in comparison to the bulk counterparts, arising from their unique van der Waals layered structure, which is favorable to improve circuit density. Nevertheless, the synthesis of large-area 2D materials in high quality and uniformity needs to be addressed before their practical applications.

Magnetic devices

Magnetic-based devices, such as magnetic tunnel junctions and magnetic textures, employ both the unique magnetic and electronic properties of the electrons. When electrons flow through the magnetic materials, the majority becomes spin polarized by the magnetization orientation. Meanwhile, when a current with the spin-polarization paralleling to that of the electrons is injected into the magnetic material, the material also exhibits higher conductance if it owns parallel polarization.[62] Besides, the spin currents can drive the movement of magnetic textures.[63] Magnetic tunnel junctions can be used as embedding memory with non-volatile programming,[64] while magnetic textures such as skyrmions possess short-term characteristics in the timescale of nanoseconds.[65] Skyrmions are topologically stable particle-like spin textures in nanoscale, which can be manipulated and driven over large distances with spin torques and spin-orbit torques.[66] It is noted that the distances are potential vectors of information allowing for computing.[65,67] Particularly, the particle-like behavior of skyrmions and their thermal Brownian motion are similar to the diffusion of neurotransmitter.[68] Additionally, the displacement transition can be restrained through introducing additional potential barrier which could block the pathway of the skyrmions transition.[69] The thermal or current disturbance could deform the skyrmions, which changes the electrical response.[65] After the removal of the disturbances, Skyrmions return to their initial arrangement and morphology within several nanoseconds.[67]

Skyrmion features nanosecond response and relaxation, which is difficult to achieve in other devices with different mechanisms. However, so far, the investigation of Skyrmions-based STP is still in the stage of theoretical simulation, the experimental demonstration in hardware devices needs more exploration.

Ferroelectric devices

Ferroelectric materials intrinsically possess spontaneous electric polarization that can be reversed by an external electric field, as shown in the top right corner of Figure 3. The spontaneous polarization can modulate the interface barrier of ferroelectric tunneling junctions (FTJ) or induce carriers in the channel of ferroelectric field effect transistors (FeFET). When applying an electric field greater than the coercive voltage, the polarity of the ferroelectric materials is capable of flipping, which results in the conductance change. In contrast, a small electric field is not sufficient to completely flip the domains,[70] in which the reversible domain wall motion dominates at this point. The characteristics of domain wall motion in ferroelectric materials are depicted in the last panel in Figure 3. The polarization state of ferroelectric materials demonstrates history-dependent transient switching under pulse stimulations. For example, Li et al. reported a ferroelectric diode[71] in which the polarization-relaxation-induced increase in the Schottky barrier height leads to conductance decay. It is noted that the larger polarization flip would emerge under stronger stimulus, which gives rise to more robust retention characteristics. Therefore, the stimuli for ferroelectric materials need to be precisely controlled to achieve short-term characteristics which also limits the applications of ferroelectric devices. In comparison, anti-ferroelectric materials might be more suitable for short-term devices with spontaneous domain recovery in nature. Recently, Cao et al. constructed an anti-FeFET to emulate leaky and integration properties of neurons.[72] Such leaky nature demonstrates great potential for future short-term artificial synapses.

In summary, the rich dynamics of the emerging devices can provide easy access to emulate STP. However, there are still many problems existing in these devices for serving as the short-term synapses, such as the discrete nature of ionic-based devices, the incompatibility between EGT and CMOS circuits, the limited physical implementation of magnetic skyrmion, and the practical exploration of anti-ferroelectric materials. All these novel devices need to be further developed to satisfy the ever-growing demands.

SHORT-TERM PLASTICITY BEHAVIORAL EMULATION

Based on the above physical mechanisms, the conductance of the emerging devices can vary in the timescale from microseconds to seconds. The short-term characteristic experimentally enables the implementation of STP behaviors, including PPF, PPD, and PTP, which are three cornerstones for the computing functionalities in the network (collective) level of the nervous system.[73] In this section, recent advances in mimicking the fundamental short-term behaviors in the emerging devices are discussed.

Short-term potentiation

Short-term potentiation increases the transmission of synaptic information. PPF and PTP are two basic forms of potentiation. PPF refers to the synaptic currents evoked by pairs of stimuli, in which the second

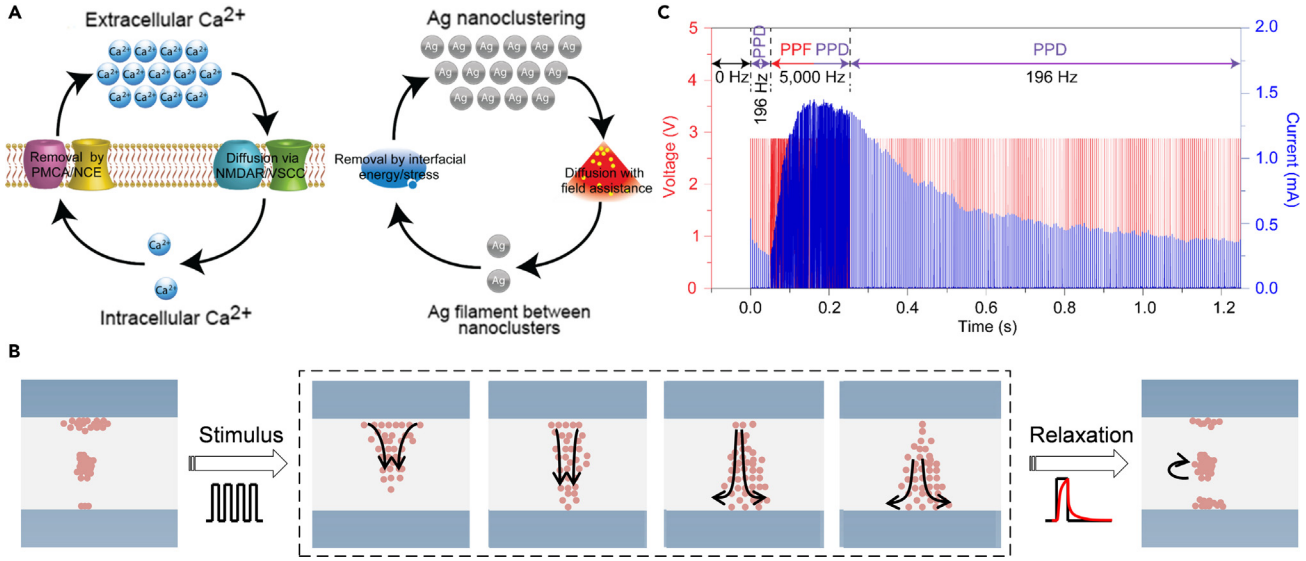

Figure 4. STP implementation in emerging devices

(A) Comparison of ion dynamics in the diffusive SiO_xN_y:Ag memristor with that in biological synapses. Left panel: Diffusion of Ca^{2+} from extracellular sources via VGCCs and receptors, and the removal of Ca^{2+} via exchanger. Right panel: Ag diffuses into the dielectric layer to form filaments under the electrical field and removes by interfacial energy or mechanical stress. Reproduced with permission from ref. [77], Springer Nature.

(B) Schematic of the switching process inside the ion migration device. Filaments are formed through ions migration under the electric field and relax to clusters after stimuli.

(C) Experimental demonstration of PPD following PPF in the diffusive SiO_xN_y:Ag memristor. Device current (blue) responses to a voltage pulse train with the same amplitudes but different frequencies. Reproduced with permission from ref. [77], Springer Nature.

response (A2) is greater than the first one (A1). The interval Δt between two stimuli typically ranges from few milliseconds to several seconds, suggesting that the PPF is able to maintain up to several seconds. Compared to PPF, PTP requires the activation of a high-frequency burst of presynaptic APs (referred to tetanus) and is delayed in its onset. For example, as shown in Figure 2C, the response to the first stimulus at 2000 ms is much stronger than that at 120 ms in the resting state. PTP typically enhances the transmitter release up to a few minutes after the train of stimuli.[74] It is noted that the PTP and PPF usually appear simultaneously during realistic stimulus trains. The PPF ratio (A_2/A_1) versus Δt has two components of facilitation that can be evaluated by a double exponential decay:[14,75,76]

$$R = 1 + C_1 \exp(-\Delta t / \tau_1) + C_2 \exp(-\Delta t / \tau_2) \quad \text{(Equation 3)}$$

where C_1 and C_2 are the initial facilitation magnitude of the rapid and slow phase, τ_1 and τ_2 are the characteristic relaxation times of the rapid and slow phase respectively. It is worth noting that the equation can also describe PPD with negative C.

PPF bionics has been investigated for more than a decade. Ohno et al. demonstrated STP in Ag_2S-based electrochemical metallization (ECM) devices.[3] The atomic bridge gradually precipitates under successive pulses, which then ruptures spontaneously once the input is removed. The device exhibits typical PPF characteristics before the completed formation of the bridge. Wang et al. clarified the mechanism of Ag diffusion device possessing dynamics similar to the biological accumulation and extrusion of Ca^{2+} (Figure 4A).[77] In detail, Ag ions can be driven by the electric field and interfacial energy, which is similar to the Ca^{2+} movement influenced by concentration gradients and exchangers. Under successive stimulation of electrical pulses, Ag ions are repelled or attracted, which produces a conductive channel with gradual solidification, as shown in the left half of Figure 4B.

In contrast to the extensive investigations in PPF, the emulation of the kinetic system of PTP is rare since its biological function is still unclear and the mechanism is not straightforward. Most works illustrate PTP by using a series of pulses, in which the PTP is expressed as the comparison of the initial and final conductance values.[78–81] For instance, Ji et al. presented an organic EGT utilizing PEDOT:Tos/PTHF as the active channel.[78] The PTP was defined as the current ratio between the 10th response and the first one, which is

however different from that of the genuine biological mechanism. The synaptic computation inspired by PTP still requires more exploration.

Short-term depression

STP plays a vital role in various neural computations and biological behaviors, which is particularly true for inhibitory synapses. The underlying biological mechanisms and related behaviors of short-term depression have been widely studied,[13,82] whereas the experimental implementation of the depression emulation is still limited. In contrast to PPF, it is difficult to realize PPD in a single device because most emerging devices exhibit positive feedback under stimulation. After the first stimulus, devices tend to be more conductive for the subsequent stimulus. Therefore, richer dynamics are required to achieve PPD.

The second half of Figure 4B illustrates that successive electrical stimulation drives Ag^+ shift toward the cathode after the formation of filament. The gradual decrease of Ag^+ near the anode leads to the breakage of the filament and the decrease of device conductivity, which is illustrated as the PPD effect after the PPF, as shown in the second half of Figure 4C. It is noted that the PPD can only be realized after the formation of filaments in Ag diffusion devices. This conductance transition emulates the switch from PPF to PPD, corresponding to the amount of available vesicle in biological synapses. In comparison, it is more advantageous to achieve PPD in three-terminal devices arising from the more controllable intrinsic properties via an extra terminal. Recently, Yang et al. proposed a versatile mechanoplastic artificial synapse based on tribotronic floating-gate MoS_2 transistors.[57] The charges are trapped in the floating gate assisted with the Au nanoparticles followed by the gradual decay through the tunneling layer in the artificial synapse, which realizes the mechanical displacement derived PPD and PPF. However, attributing to the low carrier mobility of the channel materials, the TENG-based EGT is restricted by a relatively large relaxation time of about 20 s. Xi et al. fabricated a Schottky barrier-based field-effect transistor by employing single-crystalline $NiSi_2$ contacts which forms an atomically flat interface with Si and $Hf_{0.5}Zr_{0.5}O_2$ ferroelectric layers on insulating substrates.[83] The polarization switching dynamics of the ferroelectric layer could effectively modulate the $NiSi_2$/Si Schottky barriers and the potential in the channel, which therefore changes the device conductance. Both PPF and PPD were observed in the Schottky barrier-based transistors by using the pulse with low energy of ~2 fJ, indicating its outstanding energy efficiency. Zhou et al. demonstrated an artificial synapse based on solution-processed 2D C_3N/polyvinylpyrrolidone (PVPy) matrix,[79] in which C_3N nanosheets serve as the protons capture centers. Under positive stimuli, protons accumulate in the C_3N layer and form the transport paths with proton hopping. Both PPF and PPD were successfully mimicked by changing the interval between the successive pulses.

Although huge efforts have been made to emulate STP recently, there are no certain indicators to define its standard merits. In general, the operation pulse width is expected to be $\tau < 1$ ms[2]. Scalability and formless operation are essential to achieve high-density integration with high reliability. Additionally, for practical applications in neuromorphic computing, the devices are required to have low power consumption. To reduce power consumption, the most straightforward approach is to lower the operating voltage and pulse duration. The excellent scalability of memristors is essential to lower the operating voltage. Besides, the leakage currents can be retained at a low level through interface engineering or the introduction of interlayers.[84,85] It is reported that the operating voltage of memristor can be reduced to tens of millivolts.[86] On the other hand, a faster switching speed allows for the shorter pulse duration and thus the lower power consumption. Meanwhile, faster operation speed is more favorable for enhancing the computing speed of the system, which is essential for high-speed scenarios, such as high-speed photography.[87] In addition to the power consumption, it is essential to ensure the operation voltage and relaxation time compatible with bio-signals for biological interfaces such as prosthetics and bio-detection sensors.[88] To compare the STP characteristics in different types of the emerging devices, we summarize the recent studies shown in Table 1.

STP BASED NEURAL FUNCTIONS AND HARDWARE IMPLEMENTATIONS

The previous section describes the emulation of synaptic STP at the behavioral level by using the emerging devices, including PPF, PTP and PPD. These behaviors serve as the building blocks for higher-order brain function. However, attributing to the difficulty in distinguishing the contribution of various forms of STP from the synaptic dynamics, it is challenging to study the role of individual type of STP on synaptic operations.[101] It is almost unlikely to realize high-level brain functions by employing a single synaptic behavior, since most of the functions are based on the interconnected networks. In this section, several

Table 1. Summary of recent studies using emerging devices to demonstrate short-term plasticity

Materials & Structure	Mechanism	PPF	PPD	PTP	Relaxation Time	Pulse Amplitude	Pulse Width	Features/Applications	Study
Ag/MgO/Pt	Filament	√		√	~200ms	5V	1ms	STP-to-LTP transition	Zhang et al.,[89]
Pt/SiOxNy:Ag/Pt	Filament	√	√		~30ms	0.75V	5ms	-	Wang et al.,[77]
Cu/α-Si/Pt	Filament	√			~2ms	0.8V	1ms	-	Zhang et al.,[90]
Ag_2S	Filament	√			>10s	80mV	0.5s	-	Ohno et al.,[3]
Ag/Ag:Ta_2O_5/Pt	Filament	√		√	~80us	0.6V	10us	STP-to-LTP transition	Wang et al.,[91]
TiN/Ti/TiO_2/SiOx/Si	Interface	√			~1s	3.5V	10ms	Highly nonlinear I–V characteristics	Cho and Kim,[92]
Cr/TiO/Pt	Interface	√			~200ms	1V	1ms	-	Lim et al.,[73]
ITO/TiOy/TiOx/TiN	Interface	√			>100us	1V	50us	-	She et al.,[93]
SiO_2/IGZO FET	EGT	√	√		-	0.5V	5ms	High-pass and low-pass filtering	Wan et al.,[94]
C3N/PVPy	EGT	√	√	√	~200ms	5V	10ms	-	Zhou et al.,[79]
PEDOT:Tos/PTHF	EGT	√		√	~100s	0.5V	20ms	Associative learning	Ji et al.,[78]
Au/LPE/Mxene/Si	EGT	√			~2s	1V	360ms	Associative learning	Wei et al.,[95]
Au/Electrolyte/Graphd-iyne/Si	EGT	√			~2s	1V	0.44s	Information integration	Wei et al.,[96]
MoS_2	Electronic-based	√	√		~10s	25V	50ms	Sound localization	Sun et al.,[30]
-	Skyrmion	√			<10ns	-	1ns	-	Prasad et al.,[97]
Ag/PZT/LSMO	Ferroelectric	√			<1s	2.2V	100ns	Low energy cost	Yoon et al.,[98]
Pt/BTO/SNTO	Ferroelectric	√			~10us	0.1V	50ns	-	Yoon et al.,[99]
Sb-memtransistor	Phase change	√	√	√	~5s	~20V	-	Sequential learning Combinatorial optimization	Sarwat et al.,[100]

neuromorphic functions based on one or more of the above behaviors are discussed, including temporal filtering, sound localization, associative learning, working memory and reservoir computing.

Temporal filtering

STP inherently possesses the filtering function, in which potentiation supports a high-pass filter and depression underpins a low-pass filter.[6] For potentiation, the subsequence stimulus enhances the synaptic efficacy by the residual Ca^{2+}. Thus, high-frequency input makes the synapse more active, which is beneficial for high-pass filtering. In contrast, the depression behavior in an opposite way can be used for a low-pass filter. Besides, band-pass filters can be realized in the facilitation and depression of co-existing synapses by controlling the number of vesicles and release probability.[75] Figure 5A shows three kinds of nerve signals representing these filtering functions. The bioelectrical signals in the left panel show PPD, PPF, and frequency-dependent responses, and the filtering functions are illustrated by plotting the responses as a function of stimuli frequency as shown in the right panel.[6] In spike processing, dynamic tuning of the synaptic weight gives rise to many significant pattern representation and processing capabilities.[102] For instance, in biological sensing systems, depression can be linked to habituation while facilitation is related to the sensitization, i.e., escape-related reflex.[103] Besides, filtering enables the extraction of key information for organisms, which is favorable for fast and energy-efficient information encoding in the ever-changing external world.

In hardware implementations, artificial synapses with PPF or PPD behaviors intrinsically possess rate-dependent response.[106] The lateral-coupled proton EDL transistor shown in Figure 5B[104] can serve as a high-pass filter arising from its PPF characteristics. The EPSC amplitude triggered by the presynaptic spikes increases with the frequency of the gate pulse, enabling high-frequency dynamic time filtering, as shown in Figure 5C. The equivalent filtering gain can be obtained from the ratio of the EPSC amplitudes triggered by the last spike and the first spike. The proton transistor demonstrates excellent high-pass filtering

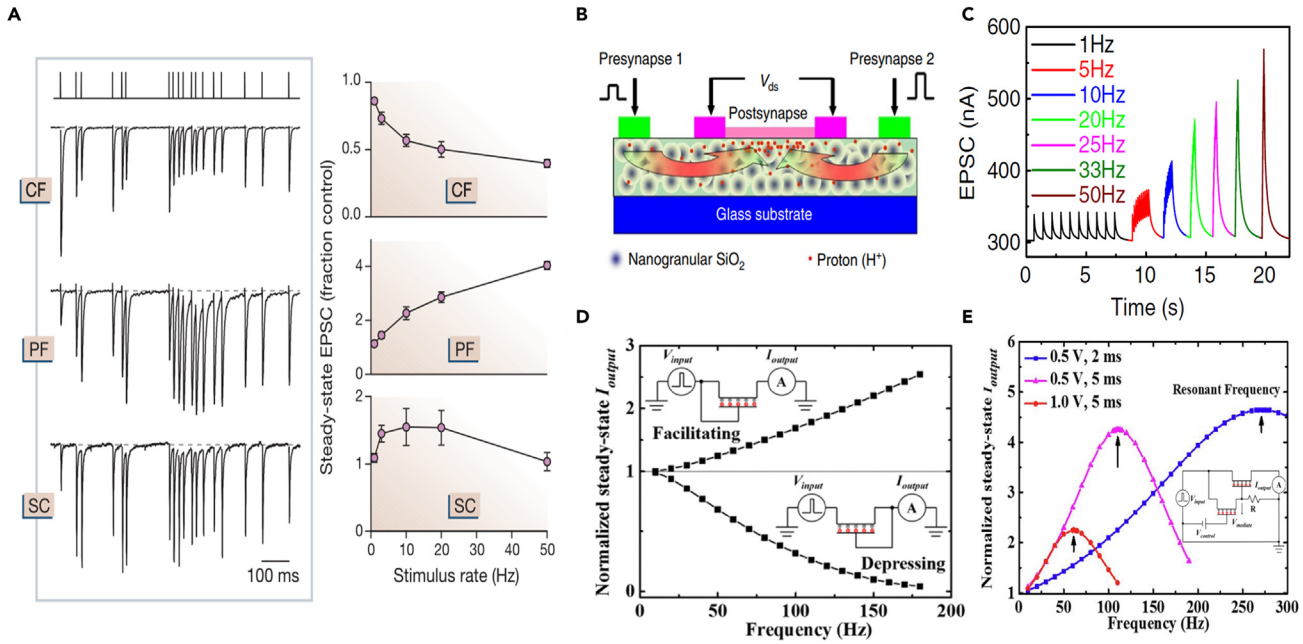

Figure 5. Temporal filtering in biology and physical devices

(A) Left panel: examples of EPSCs recorded in response to an irregular stimulus train with an average rate of 20 Hz at the climbing fiber (CF), parallel fiber (PF) and Schaffer collateral (SC) synapses. Right panel: steady-state EPSC as a function of stimulus rate. Reproduced with permission from ref. [6], Springer Nature.

(B) Indium-zinc-oxide (IZO)-based protonic/electronic hybrid transistor. Protons are laterally coupled under electric fields to produce a change in conductance. Reproduced with permission from ref. [104], Springer Nature.

(C) Frequency responses of the IZO-based EGT. The lateral coupling of protons increase with sequence inputs, and high-pass filtering is realized by the PPF effect. Reproduced with permission from ref. [104], Springer Nature.

(D) Facilitatation and depression are realized respectively in single EGT by linking the gate to input and ground. When the gate is positively stimulated, channel can induce charges. In contrast, when the gate is grounded, charges in the channel would be excluded. Reproduced with permission from ref. [105], Wiley.

(E) The analog circuit containing two EGTs connected in high- and low-pass modes is able to achieve a band-pass filter. Reproduced with permission from ref. [105], Wiley.

characteristics, whereas its relaxation time is as long as seconds, which is unfavorable for the low-pass filter. Wan et al. investigated an indium-gallium-zinc oxide (IGZO) EDL transistor,[94] in which high-pass and low-pass filtering characteristics for both fixed-rate spike train and Poisson-like rate spike train are successfully emulated. As shown in Figure 5D, the high-pass and low-pass properties are realized by different grounding methods in which the three-terminal EDL transistor is reduced to a two-terminal synaptic device with two modified connection schemes. Through combining the high- and low-pass filter, an analog circuit implementing a band-pass filter was successfully achieved using an IZO-based EGT (Figure 5E).[105] As shown in the inset of Figure 5E, the signals are firstly fed to an inhibitory connected device, in which the output serves as the gate control for the facilitated device. By adjusting the input frequency, it is possible to achieve a tunable balance between the depression and potentiation of the two devices. It is noted that the three-terminal devices have a more flexible structure to achieve different forms of filtering functions due to their richer connection modes. Although the role of dynamic filtering is not yet fully understood in practical information transfer, it is the cornerstone for achieving higher-order functionality in the hardware for the applications of neuromorphic computing.

Thanks to the rich ionic dynamics in the biological synapse, different temporal filtering states may be triggered in different scenarios. In artificial synapses, multiple dynamics are usually realized by precisely changing the control signals.[107] Multi-terminal devices offer more possibilities for regulation, which could be a promising solution to achieve various filtering characteristics through different electrical connections.

Sound localization

Sound localization is one of the most precise spatiotemporal coding functions in the brain. Sound localization relies on many complicated mechanisms, among which the interaural time difference (ITD) and

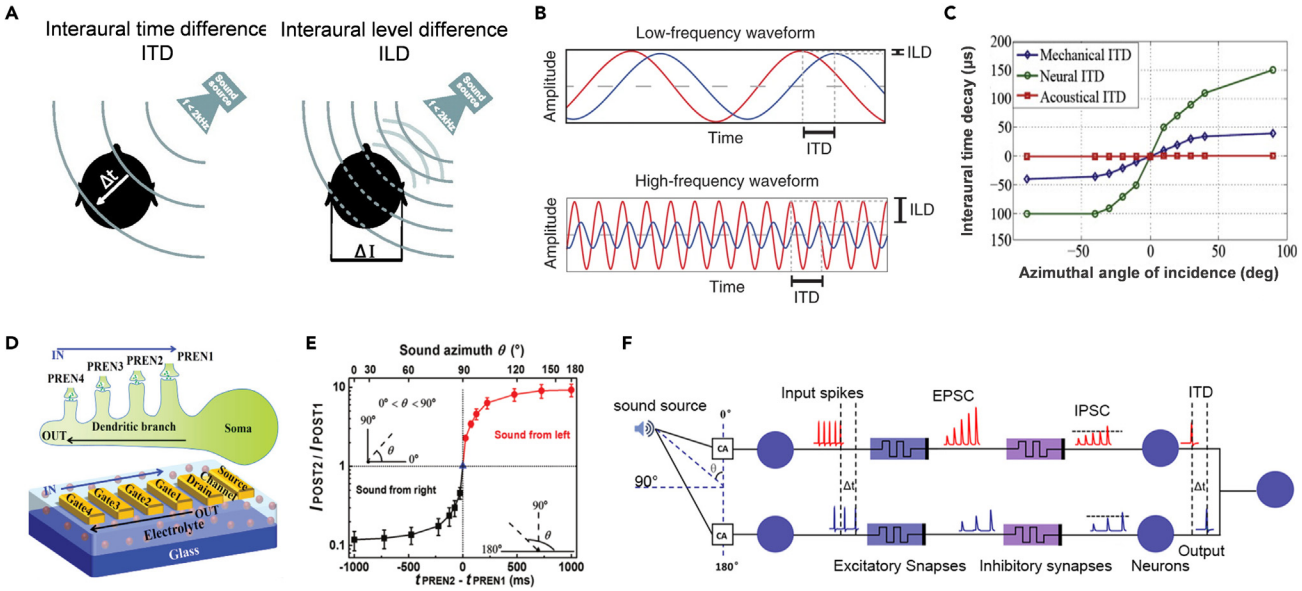

Figure 6. Sound localization in biology and physical devices

(A) Schematic image of sound location by ITD and ILD effects in human. Reproduced with permission from ref. [113], the American Physiological Society.

(B) Sample binaural signals at high and low frequencies. ITD is embodied in the phase differences and ILD in the amplitudes. Reproduced with permission from ref. [114], Springer Nature.

(C) Amplification of the ITD at the acoustic, mechanical, and neuronal levels. Reproduced with permission from ref. [110], Elsevier.

(D) A schematic structure of the neuro-transistor with multiple in-plane gates in one direction, which are regarded as presynaptic terminals on a single dendritic branch of a neuron. Reproduced with permission from ref. [115], Wiley.

(E) The ratio of the current amplitude as a function of the time interval and the sound azimuth. Reproduced with permission from ref. [115], Wiley.

(F) A sound localization system with amplifying the ITD effect. Reproduced with permission from ref. [30], American Chemical Society.

interaural level difference (ILD) are the most prominent ones.[108] Due to the stereoscopic sound perception of the creatures, sound signals are transmitted to the two ears at different times and intensities (Figure 6A). Figure 6B illustrates the ITD and ILD effects in human ears. The red and blue lines represent the sound signals in the human brain. ITD and ILD exist for both high- and low-frequency inputs. As for ITD, these signals exhibit phase-locked firing actions.[109] When the sound localization is activated, the nervous system acts as an amplifier which would magnify the time difference generated by the ITD twice in the neural circuit.[110] As shown in Figure 6C, the red line represents the acoustical signal received from the environment. The mechanical amplification expands the time differences via the cochlear structure shown as blue line. The differences could be further magnified through signal processing and adaptation at the neuronal level arising from the neuron interactions, shown as the green line. As for ILD, the improvement in sound level enhances the firing rate of cochlear nucleus neurons and thus the average frequency of excitatory postsynaptic potentials (EPSPs).[111] Although the firing intensity increases with the sound level, animals are able to maintain acoustic sensitivity over a wide dynamic range of intensities and avoid cochlear damage caused by runaway amplification in ILD mode. It is believed that PPD makes a difference in the gain control.[31,112] Although louder sounds provide higher frequency inputs to cochlear nucleus neurons, ILD is counteracted by synaptic inhibition, preserving the total synaptic input unaffected by stimulus frequency.

Most physical sound location systems are based on the ITD effect. He et al. proposed a capacitor-coupled multi-terminal oxide-based EGT for spatiotemporal information processing.[115] This device can receive inputs from multiple gates, as shown in Figure 6D. In the case that the sound signal comes from the left side, channel will receive signals from "Gate4," "Gate3," "Gate2," and "Gate1" in sequence. The device exhibits PPF behavior, which produces significant response signals while gradually enhancing the successive stimuli. In contrast, if the sound comes from the right side, the device fails to fire under progressively weakened stimuli. The coupling between devices decreases with increasing spatial distance. Using a pair of devices set in a symmetrical structure, the artificial ear will distinguish the angle by the response ratio as shown in Figure 6E. In order to clearly distinguish the ITD effect from frequency information, Sun et al. demonstrated the synaptic computation using a MoS_2-based device.[30] The highly n-doped MoS_2 transistor shows

apparent metal-insulator-transition behavior modulated by the gate voltage. Both the PPF and PPD were emulated through varying doping. Based on the frequency of selective excitatory and inhibitory synapses, a sound location system is constructed. As shown in Figure 6F, this system suppresses interference with ILD by enlarging signals using PPF and encodes ITD-related information through PPD effects. This process amplified the time difference between the two signals, and thus identifies the orientation with ITD only.

Although these devices could realize the orientation detection of signals, its practical applications should be further explored since sound localization in biology is far more complex. For example, according to the current strategy, stereo sound recognition needs at least a three-dimensional integrated system.

Associative learning

Associative learning describes the recurrence of two different stimuli and the formation of brain connections, such as conditioned reflexes and other behaviors, which is usually closely related to the interconversion of short- and long-term memory.[116] With repeated training, the two unrelated stimuli (conditioned stimulus, CS; unconditioned stimulus, US) could be coupled to each other under specific conditions in associative learning. In detail, animals can associate CS and US in artificial synapses and then respond to the CS in the same way as US. Although the biological mechanisms of associative learning are not fully understood, previous reports have verified that short-term plasticity plays an important role in associative learning processes in mammals.[117] It is proposed that associative learning enables related things to link together, which plays a crucial role in the cognitive functions of biological systems.[118,119] Pavlov's dog experiment is a typical example of conditioned reflexes that can be implemented in neuromorphic computing tasks.

In the aspect of physical emulations, the co-stimulation of electrical and optical signals is usually used as the US and CS respectively. Ji et al. fabricated two parallel PEDOT:Tos/PTHF transistors connected in series, where the associative learning is realized via coupling the response of optical and pressure signals based on PPF.[78] The transistor exhibits either STP or LTP characteristics through applying different bias voltages. More intriguingly, the LTP could be strengthened with significant non-volatility when the pressure sensor and photo-resistor response co-exist in the device, whereas triggered PPF in volatility could be observed if the operation only acts in the pressure sensor. The volatile PPF alone cannot reach the threshold to simulate salivating dogs. Once the conductivity of the non-volatile memory reaches the threshold, the superimposed PPF effect results in "drooling" under several light stimuli. In terms of the channel materials, perovskite is believed as a promising building block for sensing applications. A thin-film transistor-like synaptic device[120] using all-inorganic cesium lead bromide ($CsPbBr_3$) perovskite quantum dots (QDs) and amorphous indium gallium zinc oxide was also explored for intelligent optical computing systems. The conditioned reflexes of Pavlovian dogs to CS after training, and the time-dependent forgetting and re-response after retraining could be simulated by using the single synaptic device. In addition, flexible electronic systems have been extensively investigated for the application of short-term plasticity and associative learning,[121,122] which can be further utilized in simulating complex brain activity in bionic memory sensors, intelligent soft robots, and skinned electronic devices.

Working memory

Working memory refers to the maintenance and manipulation of temporal information after receiving signals,[123] a role which is that of DRAM in computers (Udipi et al., 2010). As a temporal interface, working memory integrates multimodal information from sensory memory and long-term memory, forming the foundation of cognition.[124] Working memory could retain information until the related actions are executed. It is noted that the short-term potentiation is critical to enhance the local connections in the brain during the maintenance process.[33] To recall the memory via performing an action, PPF serves as the promoter to reactivate the local system by adding a weak nonspecific excitatory to the whole network.[33]

Based on the function of working memory, a four-part multi-component model is proposed for different types of multimodal information processing, including central executive, visuo-spatial sketch pad, phonological loop, and links to long-term memory (knowledge or experience).[123,125] The ability to reactivate quickly after a stimulus is a key indicator for working memory in experiments. This feature is favorable for avoiding continuous delivery of high metabolic spikes and achieving information storage with low energy consumption. For instance, a planar memristor arrays based on the 2D channel was fabricated to retrieve or update information quickly.[126] Although the conductance in the array decays to the pristine

values in 10 s, the operation history is still memorized, which is beneficial for the simple recall of the programmed pixels during retrieval. This behavior emulates the aforementioned reactivation of stimulated neurons in the monostable state in the excitatory synapse, which is anti-noise. Besides, the information computing and conversion were emulated by interacting with weights loaded from "long-term memory." Through employing the four-part theoretical framework, multimodal information can be well-cognized by experience or knowledge.

Working memory is a complex concept related to storage, utilization, and reconstruction of memory. It is still difficult for hardware to emulate the whole functions because devices with different features are required to be integrated into one chip. The exploration of new devices with both tunable long- and short-term characteristics may hold promise for reconfigurable and energy-efficient memory.

Reservoir computing

Reservoir computing (RC) is a framework for computation derived from recurrent neural networks (RNN). It is hard to consider RC as a higher functional manifestation of the brain, whereas it possesses extraordinary capability to process temporal information, which has been demonstrated in various short-term devices.[127] RC contains a fixed and nonlinear "reservoir" to process information recursively, and map the information to higher dimensional computational spaces. Due to the fixed reservoir dynamics, only the readout layer requires training.[128] which makes RC favorable for low-cost training, high versatility, and fast learning. Recently, the RC network employing self-feedback nodes is experimentally demonstrated in the devices with nonlinear and short-term properties, as shown in Figure 7A. Nonlinearity maps feature from low-dimensional inputs into a high-dimensional feature space and short-term memory enables nodes to update states to handle temporal information.[128]

In physical implementation, many dynamic devices with nonlinearity and short-term memory effects (Figure 7B) have been used as reservoir nodes. For example, Ag diffusive devices[131] decaying in every frame (Figure 7C) are suitable for implementing the RC system. In the top panel, the two sequential stimulus at the beginning induces a continuous Ag filament and high conductivity in the device when the device state is analyzed in the fourth time slot. The remaining two panels are operated in a similar way with different pulse conditions. Depending on the dynamical properties of the device, the different sequences result in different device conductance. The temporal information is fully captured by the ions' dynamics. In order to integrate sensing and computing functions, ultra-wide bandgap semiconductor GaO-based photo-synapses are fabricated for in-sensor recognition of latent fingerprints. Under optical stimulus, the photo-synapses demonstrate similar response as that in the Ag diffusive electrical devices, as shown in Figure 7E. The optical response of GaO-based nodes allows the RC system to work without data transmission between sensor and reservoir, thus reducing energy consumption and data latency. In addition to memristors, RC has also been experimentally implemented by exploiting the complex dynamics of spintronic oscillators,[134] magnetic skyrmions,[67] and nanowire.[135] Through mapping information into high-dimensional space and updating states every frame, RC demonstrates the outstanding capability for recognition and temporal tasks. For instance, handwritten digit recognition task is realized by converting pixels to pulse sequences,[136] as shown in Figure 7F. On the other hand, excellent performance on the time-series prediction can be achieved in the Mackey-Glass series task,[133] as shown in Figure 7G. The advantages of low power consumption and hardware friendliness illustrate the great potential of using RC in the temporal prediction paradigm.[137–139]

Although a wide range of RC systems has been demonstrated successfully, the systems possessing multiplex time through using an individual device as a virtual node is still inefficient. Besides, to improve the recognition accuracy, complex measuring and computing methods have to be adopted, such as multiple sampling outputs (Figure 7E), and the conversion of the same information into several sequences with different frequencies (Figure 7F). Nevertheless, the problems could be solved by optimizing the device performance, including improving device-to-device uniformity for parallel input and computation, and designing devices with tunable decay constants for more dynamics in information mapping (Figure 8).

Additional functions

In addition to the above-mentioned applications, there are still various STP-related functions that are widely observed in biology but yet realized in physical devices. For example, the short-term depression in cortical synapses has been validated to provide dynamic gain control, which enables equal percentage

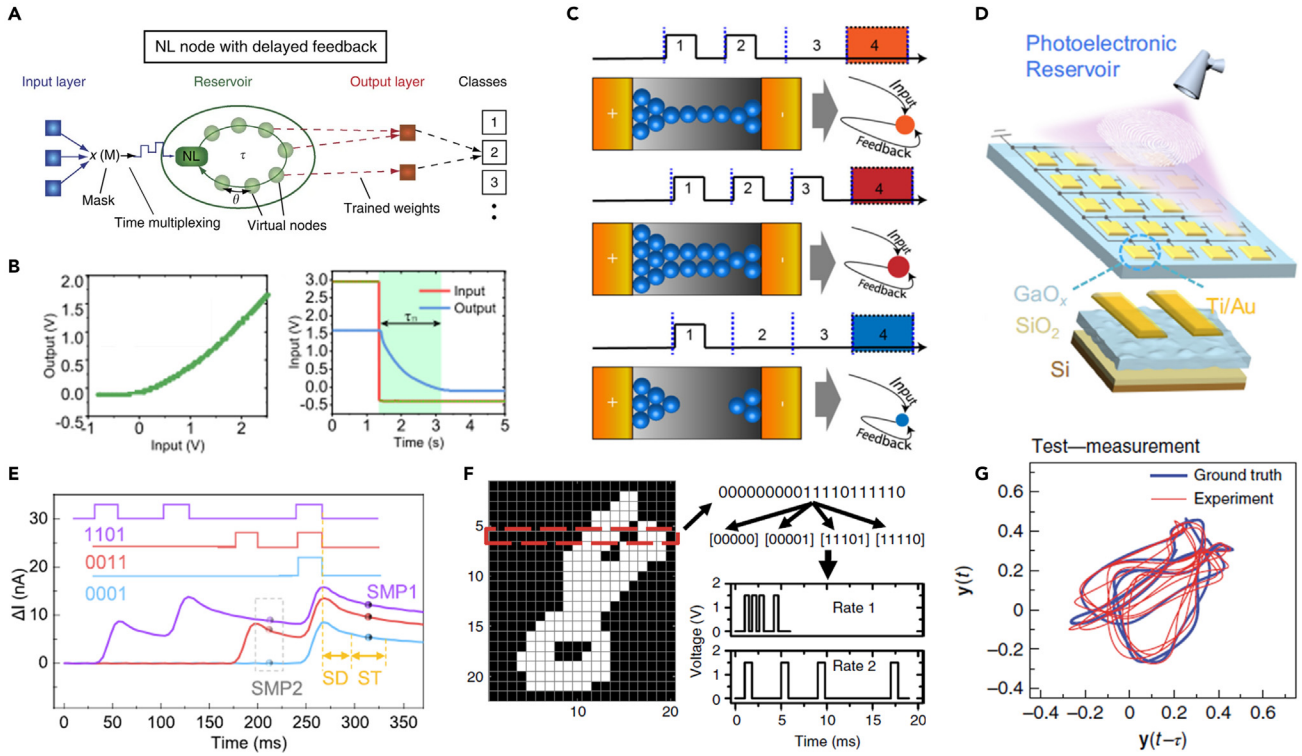

Figure 7. Reservoir computing in physical devices

(A) Schematic of a conventional RC system. Emerging devices act as virtual nodes in the RC network. Reproduced with permission from ref. [129], Springer Nature.

(B) Basic dynamic response for RC system. Nonlinearity maps information to a high-dimensional space for easy classification. Short-term memory enables the encoding of temporal information. Reproduced from ref. [130], CC BY license.

(C) Electrical synapses are able to produce history-dependent responses under different encoded sequence inputs in every frame. Reproduced from ref. [131], CC BY license.

(D) Photo-synapses are able to sense optical stimulus and update conductance according to the optical sequences. Reproduced from ref. [132], CC BY license.

(E) Processing of inputs and sampling of outputs. Sample1 (SMP1) and sample2 (SMP2) are measured as classification basis. Reproduced from ref. [132], CC BY license.

(F) Handwritten digit recognition using a memristor-based RC system. The pixels are encoded as spike sequences with different frequencies as nodes' inputs. Reproduced from ref. [128], CC BY license.

(G) Autonomous forecasting of Mackey-Glass time series using a memristor-based RC system. Reproduced with permission from ref. [133], Springer Nature.

rate changes for the inputs with different frequencies to produce equal postsynaptic responses.[40] On the other hand, thanks to the unique characteristics of potentiation and depression, synapses play a vital role in modulating the information transmission process such as burst detection[142] and decorrelation.[143] Burst-like clusters of transmissions contain more non-regular positive information which could be enhanced by facilitation. Decorrelation could filter out irregular information and eliminate redundant associations, allowing the transmission sequence to convey information more efficiently. Previous works also claim that STP is correlated with the dynamic coding of the nervous system.[144] In the early stage of adaptation, neurons fire at high frequencies, which enhances the synaptic connections persisting for a while. After that, the neurons would fire less frequently in late adaptation. Nevertheless, since the synaptic efficacy between neurons has been enhanced together with the improvement of the synchronized firing of neuronal populations, the continuous encoding of object information can still be realized. Although it is challenging to achieve many of the computational functions using STP in hardware as mature as brains, STP provides abundant valuable insights for neuromorphic computing.[19,145,146]

Inspired by STP, mature neural networks can compute in a more efficient way. For instance, Sarwat et al. demonstrate a novel phase-change transistor with both long- and short-term plasticity.[100] The synapse with mixed-plasticity is capable of coupling the non-volatile and transient weight updates, which is advantageous for important computational metrics against larger recurrent and long- and short-term memory

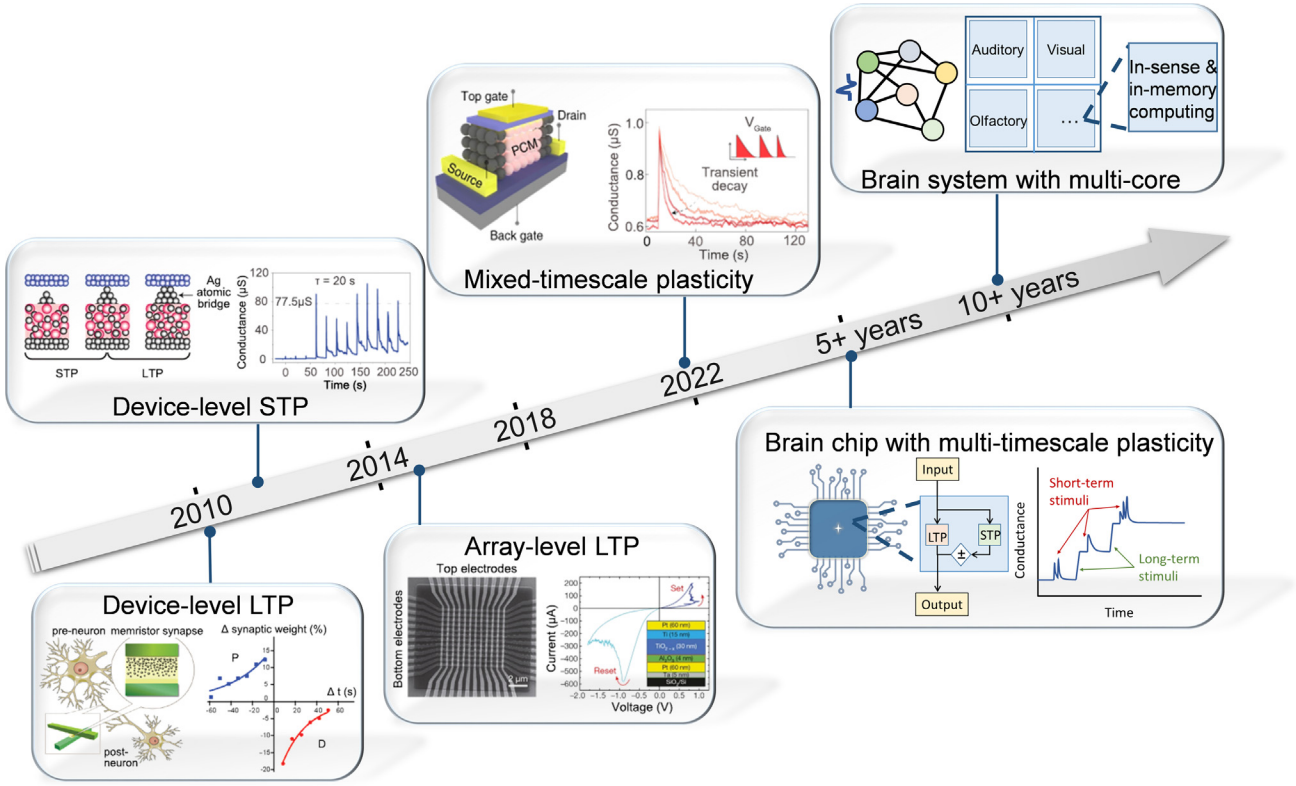

Figure 8. Roadmap with past milestones and future prospects in artificial synapses

Device-level LTP reproduced with permission from ref.[140], American Chemical Society. Device-level STP reproduced with permission from ref.[3], Springer Nature. Array-level LTP reproduced with permission from ref.[141], Springer Nature. Mixed-timescale plasticity reproduced with permission from ref.[100], Springer Nature.

neural networks. The combination of synaptic plasticity with different time scales makes it possible to achieve in-memory computing. Also, the application of STP in the continuous attractor network allows for the introduction of negative feedback to the network and achieves the effect of changing from tracking to anticipation.[145] The model can well explain the anticipated tracking behavior of rat head-facing neurons in biological experiments.

The current investigations of STP-based neuromorphic computing are still restricted by traditional algorithms. The optimization of device dynamic characteristics together with the exploration of new neuromorphic algorithms could make boost for high-efficient computing.

DISCUSSION AND OUTLOOK

STP has an important computational role far beyond the current perception, which is promising to overcome the challenges in nowadays computing systems. The rich dynamics in emerging devices make it possible to implement STP in hardware with high efficiency. Nevertheless, the computing capabilities of the STP have not been fully developed. The above-mentioned emerging devices also need further optimization. For example, two-terminal memristors are favorable for high-density integration arising from their high area efficiency. The unit area can be as low as the ideal $4F^2$ or $4F^2/N$ in a 3D structure. Nevertheless, the sneak pass issue in the crossbar array requires extra selective devices, such as transistors, which hinder the expected integration density. Thus, the investigation of area-efficient two-terminal selectors is of great significance for the implementation of high-density memristor arrays with superior performance. On the other hand, the intrinsic switching stochasticity in two-terminal memristors poses a great challenge in precisely programming the synaptic weights in artificial synapses, which severely degrades the performance of the neural networks. Therefore, the development of bio-inspired algorithms utilizing stochasticity is in high demand for practical applications. Compared to two-terminal memristors, three-terminal EGTs combine the architecture superiority of separated control and transduction terminals, making it possible for precise

control and multiple-terminal operation.[59] Moreover, the separated read/write operations of EGTs contribute to the emulation of the delayed release of neurotransmitters in biological synapses, which is important in synaptic computation. However, EGTs suffer from immature fabrication techniques and low integration density, which could be addressed through joint efforts of the material and electrical engineers. The time constant of the STP in magnetic devices is in nanoseconds, however, the devices suffer from physical implementation and controllability. Additionally, the comprehensive investigation of the synaptic dynamics in different devices is also emergent, since it is still difficult to fully apply the STP features in neuromorphic computing. We propose that the current challenges of STP exist in both the biological and physical aspects, which are discussed later in discussion.

i) Brain is a very powerful information-processing center composed of huge population of neurons and synapses. Information encoding and decoding in the brain is achieved in a statistical manner, which can be adjusted in real-time according to environmental conditions for self-learning and adaptation, therefore contributing to the strong fault tolerance.[147] The continuous perception of external information and the integrated processing of self-stimuli are realized by impulse transmission in neural networks. However, so far people still do not have a comprehensive understanding of the spatiotemporal encoding of impulse signals, which is a major obstacle to the development of artificial intelligence.

ii) Many behaviors and perceptions of organisms cannot be explained solely by short-range plasticity. The typical biological features including temporal-order learning,[148–151] information compression, and working memory are closely related to long-range plasticities such as STDP and neuronal population behavior., However, the coupling between short-term and long-term plasticity is complicated with mutual modulation between each other, which should be further explored.

After the first demonstration of the LTP device,[152] numerous works have been reported to achieve single plasticity at the device level. The realization of LTP arrays[141] has facilitated the computation of convolution and accelerated neural networks. On the other hand, the development of STP computing is much slower. It is recently proposed that the mixed-timescale devices[100] are capable of improving the conventional network on temporal information processing. Nevertheless, the implementation of a single device to achieve tunable plasticity at multi-timescale so as to design integrated brain-like computational chips is our ultimate goal. The computing system will act like a brain with in-sense and in-memory computing capabilities in the future. The current issues that we need to address are.

i) The coupling and decoupling of long- and short-term plasticity in neural networks is a complex process. Supports from dedicated algorithm engineers are in high demand.

ii) Biological synapses are capable of performing multi-timescale plasticity depending on the environment. The design of the specific single device possessing both STP and LTP is expected to achieve more efficient in-memory computing.

iii) STP computing lacks targeted application scenarios and algorithmic networks, which could be addressed through the close collaboration between the electronic and artificial intelligence engineers.

Many of the underlying neuronal mechanisms are still obscure, which requires further investigation in an interdisciplinary approach through researchers' cooperation in different fields such as brain science, electronics, physics, materials science, and computing science. Novel computing paradigms based on volatile devices should be further explored to accelerate existing AI computing systems. Integrating the unique properties of each device to build computational systems is one of the future paths for system designers. It is expected that the implementation of STP in artificial synaptic devices can significantly advance the development of neuromorphic computing.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grant Nos. 61825404, 61888102, 62104041, and 62104044, the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant No. XDB44000000, Shanghai Sailing Program under Grant No. 21YF1402600, the project of Biren Technology and MOE innovation platform.

AUTHOR CONTRIBUTIONS

The article was written through the contributions of all authors. C.L., X.Z., P.C., G.W., J.Y., K.Z., and M.W.: developing the original conception and structure of the perceptive article. Q.L., X.Z., and D.X.: conceptualization, investigation, and supervision. X.Z. and C.L.: writing original draft and image processing.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Tang, Y., Nyengaard, J.R., De Groot, D.M., and Gundersen, H.J. (2001). Total regional and global number of synapses in the human brain neocortex. Synapse *41*, 258–273. https://doi.org/10.1002/syn.1083.

2. Kuzum, D., Yu, S., and Wong, H.S.P. (2013). Synaptic electronics: materials, devices and applications. Nanotechnology *24*, 382001. https://doi.org/10.1088/0957-4484/24/38/382001.

3. Ohno, T., Hasegawa, T., Tsuruoka, T., Terabe, K., Gimzewski, J.K., and Aono, M. (2011). Short-term plasticity and long-term potentiation mimicked in single inorganic synapses. Nat. Mater. *10*, 591–595. https://doi.org/10.1038/nmat3054.

4. Josberger, E.E., Deng, Y., Sun, W., Kautz, R., and Rolandi, M. (2014). Two-terminal protonic devices with synaptic-like short-term depression and device memory. Adv. Mater. *26*, 4986–4990. https://doi.org/10.1002/adma.201400320.

5. Fuster, J.M., and Alexander, G.E. (1971). Neuron activity related to short-term memory. Science *173*, 652–654. https://doi.org/10.1126/science.173.3997.652.

6. Abbott, L.F., and Regehr, W.G. (2004). Synaptic computation. Nature *431*, 796–803. https://doi.org/10.1038/nature03010.

7. Marković, D., Mizrahi, A., Querlioz, D., and Grollier, J. (2020). Physics for neuromorphic computing. Nat. Rev. Phys. *2*, 499–510. https://doi.org/10.1038/s42254-020-0208-2.

8. He, Y., Yang, Y., Nie, S., Liu, R., and Wan, Q. (2018). Electric-double-layer transistors for synaptic devices and neuromorphic systems. J. Mater. Chem. C Mater. *6*, 5336–5352. C1038TC00530C. https://doi.org/10.1039/c8tc00530c.

9. Bi, G., and Poo, M. (2001). Synaptic modification by correlated activity: hebb's postulate revisited. Annu. Rev. Neurosci. *24*, 139–166. https://doi.org/10.1146/annurev.neuro.24.1.139.

10. Shiffrin, R.M., and Atkinson, R.C. (1969). Storage and retrieval processes in long-term memory. Psychol. Rev. *76*, 179–193. https://doi.org/10.1037/h0027277.

11. Bliss, T.V., and Collingridge, G.L. (1993). A synaptic model of memory: long-term potentiation in the hippocampus. Nature *361*, 31–39. https://doi.org/10.1038/361031a0.

12. Regehr, W.G. (2012). Short-term presynaptic plasticity. Cold Spring Harbor Perspect. Biol. *4*, a005702. https://doi.org/10.1101/cshperspect.a005702.

13. Grande, L.A., and Spain, W.J. (2005). Synaptic depression as a timing device. Physiology *20*, 201–210. https://doi.org/10.1152/physiol.00006.2005.

14. Zucker, R.S., and Regehr, W.G. (2002). Short-term synaptic plasticity. Annu. Rev. Physiol. *64*, 355–405. https://doi.org/10.1146/annurev.physiol.64.092501.114547.

15. Whitlock, J.R., Heynen, A.J., Shuler, M.G., and Bear, M.F. (2006). Learning induces long-term potentiation in the hippocampus. Science *313*, 1093–1097. https://doi.org/10.1126/science.1128134.

16. Wu, S., Wong, K.Y.M., and Tsodyks, M. (2013). Neural information processing with dynamical synapses. Front. Comput. Neurosci. *7*, 188. https://doi.org/10.3389/fncom.2013.00188.

17. Rotman, Z., Deng, P.Y., and Klyachko, V.A. (2011). Short-term plasticity optimizes synaptic information transmission. J. Neurosci. *31*, 14800–14809. https://doi.org/10.1523/jneurosci.3231-11.2011.

18. Maass, W., Natschläger, T., and Markram, H. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. Neural Comput. *14*, 2531–2560. https://doi.org/10.1162/089976602760407955.

19. Taube, J.S., and Jeffrey, S. (2007). The head direction signal: origins and sensory-motor integration. Annu. Rev. Neurosci. *30*, 181–207. https://doi.org/10.1146/annurev.neuro.29.051605.112854.

20. Seok Jeong, D., Kim, I., Ziegler, M., and Kohlstedt, H. (2013). Towards artificial neurons and synapses: a materials point of view. RSC Adv. *3*, 3169–3183. https://doi.org/10.1039/C2RA22507G.

21. Wan, Q., Sharbati, M.T., Erickson, J.R., Du, Y., and Xiong, F. (2019). Emerging artificial synaptic devices for neuromorphic computing. Adv. Mater. Technol. *4*, 1900037. https://doi.org/10.1002/admt.201900037.

22. Wang, J., and Zhuge, F. (2019). Memristive synapses for brain-inspired computing. Adv. Mater. Technol. *4*, 1800544. https://doi.org/10.1002/admt.201800544.

23. Burr, G.W., Shelby, R.M., Sebastian, A., Kim, S., Kim, S., Sidler, S., Virwani, K., Ishii, M., Narayanan, P., Fumarola, A., et al. (2017). Neuromorphic computing using non-volatile memory. Adv. Phys. X *2*, 89–124. https://doi.org/10.1080/23746149.2016.1259585.

24. Meena, J.S., Sze, S.M., Chand, U., and Tseng, T.-Y. (2014). Overview of emerging nonvolatile memory technologies. Nanoscale Res. Lett. *9*, 526. https://doi.org/10.1186/1556-276X-9-526.

25. Lont, J.B., and Guggenbuhl, W. (1992). Analog CMOS implementation of a multilayer perceptron with nonlinear synapses. IEEE Trans. Neural Network. *3*, 457–465. https://doi.org/10.1109/72.129418.

26. Ramachandran, H., Weber, S., Aamir, S.A., and Chicca, E. (2014). Neuromorphic Circuits for Short-Term Plasticity with Recovery Control.

27. Noack, M., Partzsch, J., Mayr, C.G., Hänzsche, S., Scholze, S., Höppner, S., Ellguth, G., and Schüffny, R. (2015). Switched-capacitor realization of presynaptic short-term-plasticity and stop-learning synapses in 28 nm CMOS. Front. Neurosci. *9*, 10. https://doi.org/10.3389/fnins.2015.00010.

28. Bartolozzi, C., and Indiveri, G. (2007). Synaptic dynamics in analog VLSI. Neural Comput. *19*, 2581–2603. https://doi.org/10.1162/neco.2007.19.10.2581.

29. Liu, Y.H., Zhu, L.Q., Feng, P., Shi, Y., and Wan, Q. (2015). Freestanding artificial synapses based on laterally proton-coupled transistors on chitosan membranes. Adv. Mater. *27*, 5599–5604. https://doi.org/10.1002/adma.201502719.

30. Sun, L., Zhang, Y., Hwang, G., Jiang, J., Kim, D., Eshete, Y.A., Zhao, R., and Yang, H. (2018). Synaptic computation enabled by joule heating of single-layered semiconductors for sound localization. Nano Lett. *18*, 3229–3234. https://doi.org/10.1021/acs.nanolett.8b00994.

31. Cook, D.L., Schwindt, P.C., Grande, L.A., and Spain, W.J. (2003). Synaptic depression in the localization of sound. Nature *421*, 66–70. https://doi.org/10.1038/nature01248.

32. Stokes, M.G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., and Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. Neuron *78*, 364–375.

https://doi.org/10.1016/j.neuron.2013.01.039.

33. Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. Science *319*, 1543–1546. https://doi.org/10.1126/science.1150769.

34. Hennig, M.H. (2013). Theoretical models of synaptic short term plasticity. Front. Comput. Neurosci. *7*, 45. https://doi.org/10.3389/fncom.2013.00045.

35. Xu-Friedman, M.A., and Regehr, W.G. (2004). Structural contributions to short-term synaptic plasticity. Physiol. Rev. *84*, 69–85. https://doi.org/10.1152/physrev.00016.2003.

36. Atluri, P.P., and Regehr, W.G. (1996). Determinants of the time course of facilitation at the granule cell to Purkinje cell synapse. J. Neurosci. *16*, 5661–5671. https://doi.org/10.1523/JNEUROSCI.16-18-05661.1996.

37. Stevens, C.F., and Wesseling, J.F. (1999). Augmentation is a potentiation of the exocytotic process. Neuron *22*, 139–146. https://doi.org/10.1016/S0896-6273(00)80685-6.

38. Xu, J., and Wu, L.G. (2005). The decrease in the presynaptic calcium current is a major cause of short-term depression at a calyx-type synapse. Neuron *46*, 633–645. https://doi.org/10.1016/j.neuron.2005.03.024.

39. Herz, A.V.M., Gollisch, T., Machens, C.K., and Jaeger, D. (2006). Modeling single-neuron dynamics and computations: a balance of detail and abstraction. Science *314*, 80–85. https://doi.org/10.1126/science.1127240.

40. Abbott, L.F., Varela, J.A., Sen, K., and Nelson, S.B. (1997). Synaptic depression and cortical gain control. Science *275*, 220–224. https://doi.org/10.1126/science.275.5297.221.

41. Markram, H., and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. Nature *382*, 807–810. https://doi.org/10.1038/382807a0.

42. Tsodyks, M.V., Markram, H., and Markram. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc. Natl. Acad. Sci. USA *94*, 719–723. https://doi.org/10.1073/pnas.94.2.719.

43. Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. Neural Comput. *10*, 821–835. https://doi.org/10.1162/089976698300017502.

44. Chen, F., Zhou, Y., Zhu, Y., Zhu, R., Guan, P., Fan, J., Zhou, L., Valanoor, N., von Wegner, F., Saribatir, E., et al. (2021). Recent progress in artificial synaptic devices: materials, processing and applications. J. Mater. Chem. C Mater. *9*, 8372–8394. https://doi.org/10.1039/d1tc01211h.

45. Lee, G.H., Baek, J.H., Ren, F., Pearton, S.J., Lee, G., and Kim, J. (2021). Artificial neuron and synapse devices based on 2D materials. Small *17*, e2100640. https://doi.org/10.1002/smll.202100640.

46. Park, Y., Kim, M.K., and Lee, J.S. (2020). Emerging memory devices for artificial synapses. J. Mater. Chem. C Mater. *8*, 9163–9183. https://doi.org/10.1039/d0tc01500h.

47. Zhu, J., Zhang, T., Yang, Y., and Huang, R. (2020). A comprehensive review on emerging artificial neuromorphic devices. Appl. Phys. Rev. *7*, 011312. https://doi.org/10.1063/1.5118217.

48. He, Y., Zhu, L., Zhu, Y., Chen, C., Jiang, S., Liu, R., Shi, Y., and Wan, Q. (2021). Recent progress on emerging transistor-based neuromorphic devices. Adv. Intell. Syst. *3*, 2000210. https://doi.org/10.1002/aisy.202000210.

49. Bian, H., Goh, Y.Y., Liu, Y., Ling, H., Xie, L., and Liu, X. (2021). Stimuli-responsive memristive materials for artificial synapses and neuromorphic computing. Adv. Mater. *33*, e2006469. https://doi.org/10.1002/adma.202006469.

50. Shukla, N., Ghosh, R.K., Grisafe, B., and Datta, S. (2017). Fundamental Mechanism behind Volatile and Non-volatile Switching in Metallic Conducting Bridge RAM.

51. Wang, W., Wang, M., Ambrosi, E., Bricalli, A., Laudato, M., Sun, Z., Chen, X., and Ielmini, D. (2019). Surface diffusion-limited lifetime of silver and copper nanofilaments in resistive switching devices. Nat. Commun. *10*, 81. https://doi.org/10.1038/s41467-018-07979-0.

52. Sakellaropoulos, D., Bousoulas, P., Papakonstantinopoulos, C., Kitsios, S., and Tsoukalas, D. (2021). Impact of active electrode on the synaptic properties of SiO-based forming-free conductive bridge memory. IEEE Trans. Electron. Dev. *68*, 1598–1603. https://doi.org/10.1109/TED.2021.3057841.

53. Bousoulas, P., Sakellaropoulos, D., Papakonstantinopoulos, C., Kitsios, S., Arvanitis, C., Bagakis, E., and Tsoukalas, D. (2020). Investigating the origins of ultra-short relaxation times of silver filaments in forming-free SiO2-based conductive bridge memristors. Nanotechnology *31*, 454002. https://doi.org/10.1088/1361-6528/aba3a1.

54. Wang, Y.-F., Lin, Y.-C., Wang, I.T., Lin, T.-P., and Hou, T.-H. (2015). Characterization and modeling of nonfilamentary Ta/TaOx/TiO2/Ti analog synaptic device. Sci. Rep. *5*, 10150. https://doi.org/10.1038/srep10150.

55. Hu, R., Li, X., Tang, J., Li, Y., Zheng, X., Gao, B., Qian, H., and Wu, H. (2021). Investigation of resistive switching mechanisms in Ti/TiOx/Pd-based RRAM devices. Adv. Electron. Mater *8*, 2100827. https://doi.org/10.1002/aelm.202100827.

56. Park, S.-O., Jeong, H., Park, J., Bae, J., and Choi, S. (2022). Experimental demonstration of highly reliable dynamic memristor for artificial neuron and neuromorphic computing. Nat. Commun. *13*, 2888. https://doi.org/10.1038/s41467-022-30539-6.

57. Yang, X., Yu, J., Zhao, J., Chen, Y., Gao, G., Wang, Y., Sun, Q., and Wang, Z.L. (2020). Mechanoplastic tribotronic floating-gate neuromorphic transistor. Adv. Funct. Mater. *30*, 2002506. https://doi.org/10.1002/adfm.202002506.

58. Rivnay, J., Inal, S., Salleo, A., Owens, R.M., Berggren, M., and Malliaras, G.G. (2018). Organic electrochemical transistors. Nat. Rev. Mater. *3*, 17086. https://doi.org/10.1038/natrevmats.2017.86.

59. Zhu, Y., Zhu, Y., Mao, H., He, Y., Jiang, S., Zhu, L., Chen, C., Wan, C., and Wan, Q. (2022). Recent advances in emerging neuromorphic computing and perception devices. J. Phys. D Appl. Phys. *55*, 053002. https://doi.org/10.1088/1361-6463/ac2868.

60. Zhang, S.-R., Zhou, L., Mao, J.-Y., Ren, Y., Yang, J.-Q., Yang, G.-H., Zhu, X., Han, S.-T., Roy, V.A.L., and Zhou, Y. (2019). Artificial synapse emulated by charge trapping-based resistive switching device. Adv. Mater. Technol. *4*, 1800342. https://doi.org/10.1002/admt.201800342.

61. Wang, S., Liu, X., Xu, M., Liu, L., Yang, D., and Zhou, P. (2022). Two-dimensional devices and integration towards the silicon lines. Nat. Mater. *21*, 1225–1239. https://doi.org/10.1038/s41563-022-01383-2.

62. Wang, Z., Wu, H., Burr, G.W., Hwang, C.S., Wang, K.L., Xia, Q., and Yang, J.J. (2020). Resistive switching materials for information processing. Nat. Rev. Mater. *5*, 173–195. https://doi.org/10.1038/s41578-019-0159-3.

63. Grollier, J., Querlioz, D., Camsari, K.Y., Everschor-Sitte, K., Fukami, S., and Stiles, M.D. (2020). Neuromorphic spintronics. Nat. Electron. *3*, 360–370. https://doi.org/10.1038/s41928-019-0360-9.

64. Jarollahi, H., Onizawa, N., Gripon, V., Sakimura, N., Sugibayashi, T., Endoh, T., Ohno, H., Hanyu, T., and Gross, W.J. (2014). A nonvolatile associative memory-based context-driven search engine using 90 nm CMOS/MTJ-hybrid logic-in-memory architecture. IEEE J. Emerg. Sel. Top. Circuits Syst. *4*, 460–474. https://doi.org/10.1109/JETCAS.2014.2361061.

65. Pinna, D., Bourianoff, G., and Everschor-Sitte, K. (2020). Reservoir computing with random skyrmion textures. Phys. Rev. Appl. *14*, 054020. https://doi.org/10.1103/PhysRevApplied.14.054020.

66. Woo, S., Litzius, K., Krüger, B., Im, M.-Y., Caretta, L., Richter, K., Mann, M., Krone, A., Reeve, R.M., Weigand, M., et al. (2016). Observation of room-temperature magnetic skyrmions and their current-driven dynamics in ultrathin metallic ferromagnets. Nat. Mater. *15*, 501–506. https://doi.org/10.1038/nmat4593.

67. Prychynenko, D., Sitte, M., Litzius, K., Krüger, B., Bourianoff, G., Kläui, M., Sinova, J., and Everschor-Sitte, K. (2018). Magnetic skyrmion as a nonlinear resistive element: a potential building block for reservoir computing. Phys. Rev. Appl. *9*, 014034. https://doi.org/10.1103/PhysRevApplied.9.014034.

68. Pinna, D., Abreu Araujo, F., Kim, J.-V., Cros, V., Querlioz, D., Bessière, P., Droulez, J., and Grollier, J. (2018). Skyrmion gas manipulation for probabilistic computing. Phys. Rev. Appl. *9*, 064018. https://doi.org/10.1103/PhysRevApplied.9.064018.

69. Huang, Y., Kang, W., Zhang, X., Zhou, Y., and Zhao, W. (2017). Magnetic skyrmion-based synaptic devices. Nanotechnology *28*, 08LT02. https://doi.org/10.1088/1361-6528/aa5838.

70. Dutta, S., Saha, A., Panda, P., Chakraborty, W., Gomez, J., Khanna, A., Gupta, S., Roy, K., and Datta, S. (2019). Biologically plausible ferroelectric quasi-leaky integrate and fire neuron. In In 2019 Symposium on VLSI Technology (IEEE), pp. T140–T141.

71. Li, W., Fan, Z., Huang, Q., Rao, J., Cui, B., Chen, Z., Lin, Z., Yan, X., Tian, G., Tao, R., et al. (2023). Polarization-dominated internal timing mechanism in a ferroelectric second-order memristor. Phys. Rev. Appl. *19*, 014054. https://doi.org/10.1103/PhysRevApplied.19.014054.

72. Spalla, D., Cornacchia, I.M., Treves, A., Lu, J., Wang, Y., Yang, Y., Sun, Y., Wei, W., Wang, J., and Xu, H. (2021). Compact artificial neuron based on anti-ferroelectric transistor. Elife *10*, e69499. https://doi.org/10.7554/eLife.69499.

73. Lim, H., Kim, I., Kim, J.S., Seong Hwang, C., and Jeong, D.S. (2013). Short-term memory of TiO2-based electrochemical capacitors: empirical analysis with adoption of a sliding threshold. Nanotechnology *24*, 384005. https://doi.org/10.1088/0957-4484/24/38/384005.

74. Purves, D., Augustine, G.J., Fitzpatrick, D., Hall, W., LaMantia, A.-S., and White, L. (2019). Neurosciences (De Boeck Supérieur).

75. Dittman, J.S., Kreitzer, A.C., and Regehr, W.G. (2000). Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. J. Neurosci. *20*, 1374–1385. https://doi.org/10.1523/JNEUROSCI.20-04-01374.2000.

76. Lin, Y., Zeng, T., Xu, H., Wang, Z., Zhao, X., Liu, W., Ma, J., and Liu, Y. (2018). Transferable and flexible artificial memristive synapse based on WOx Schottky junction on arbitrary substrates. Adv. Electron. Mater. *4*, 1800373. https://doi.org/10.1002/aelm.201800373.

77. Wang, Z., Joshi, S., Savel'ev, S.E., Jiang, H., Midya, R., Lin, P., Hu, M., Ge, N., Strachan, J.P., Li, Z., et al. (2017). Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. Nat. Mater. *16*, 101–108. https://doi.org/10.1038/nmat4756.

78. Ji, X., Paulsen, B.D., Chik, G.K.K., Wu, R., Yin, Y., Chan, P.K.L., and Rivnay, J. (2021). Mimicking associative learning using an ion-trapping non-volatile synaptic organic electrochemical transistor. Nat. Commun. *12*, 2480. https://doi.org/10.1038/s41467-021-22680-5.

79. Zhou, L., Yang, S., Ding, G., Yang, J.Q., Ren, Y., Zhang, S.R., Mao, J.Y., Yang, Y., Zhou, Y., and Han, S.T. (2019). Tunable synaptic behavior realized in C3N composite based memristor. Nano Energy *58*, 293–303. https://doi.org/10.1016/j.nanoen.2019.01.045.

80. Chang, T., Jo, S.H., Lu, W., and Lu. (2011). Short-term memory to long-term memory transition in a nanoscale memristor. ACS Nano *5*, 7669–7676. https://doi.org/10.1021/nn202983n.

81. Park, Y., and Lee, J.-S. (2017). Artificial synapses with short-and long-term memory for spiking neural networks based on renewable materials. ACS Nano *11*, 8962–8969. https://doi.org/10.1021/acsnano.7b03347.

82. Anwar, H., Li, X., Bucher, D., and Nadim, F. (2017). Functional roles of short-term synaptic plasticity with an emphasis on inhibition. Curr. Opin. Neurobiol. *43*, 71–78. https://doi.org/10.1016/j.conb.2017.01.002.

83. Xi, F., Han, Y., Liu, M., Bae, J.H., Tiedemann, A., Grützmacher, D., and Zhao, Q.T. (2021). Artificial synapses based on ferroelectric Schottky barrier field-effect transistors for neuromorphic applications. ACS Appl. Mater. Interfaces *13*, 32005–32012. https://doi.org/10.1021/acsami.1c07505.

84. Li, B., Liu, Y., Wan, C., Liu, Z., Wang, M., Qi, D., Yu, J., Cai, P., Xiao, M., Zeng, Y., and Chen, X. (2018). Mediating short-term plasticity in an Artificial memristive synapse by the orientation of silica mesopores. Adv. Mater. *30*, 1706395. https://doi.org/10.1002/adma.201706395.

85. Xiao, M., Yeow, T., Nguyen, V.H., Muñoz-Rojas, D., Musselman, K.P., Duley, W.W., and Zhou, Y.N. (2019). Ultrathin TiOx interface-mediated ZnO-nanowire memristive devices emulating synaptic behaviors. Adv. Electron. Mater. *5*, 1900142. https://doi.org/10.1002/aelm.201900142.

86. Hu, L., Fu, S., Chen, Y., Cao, H., Liang, L., Zhang, H., Gao, J., Wang, J., and Zhuge, F. (2017). Ultrasensitive memristive synapses based on lightly oxidized sulfide films. Adv. Mater. *29*, 1606927. https://doi.org/10.1002/adma.201606927.

87. Gallego, G., Delbrück, T., Orchard, G., Bartolozzi, C., Taba, B., Censi, A., Leutenegger, S., Davison, A.J., Conradt, J., Daniilidis, K., and Scaramuzza, D. (2022). Event-based vision: a survey. IEEE Trans. Pattern Anal. Mach. Intell. *44*, 154–180. https://doi.org/10.1109/TPAMI.2020.3008413.

88. Wang, Z., Wang, L., Nagai, M., Xie, L., Yi, M., and Huang, W. (2017). Nanoionics-enabled memristive devices: strategies and materials for neuromorphic applications. Adv. Electron. Mater. *3*, 1600510. https://doi.org/10.1002/aelm.201600510.

89. Zhang, Y., Zhong, S., Song, L., Ji, X., and Zhao, R. (2018). Emulating dynamic synaptic plasticity over broad timescales with memristive device. Appl. Phys. Lett. *113*, 203102. https://doi.org/10.1063/1.5052556.

90. Zhang, X., Liu, S., Zhao, X., Wu, F., Wu, Q., Wang, W., Cao, R., Fang, Y., Lv, H., Long, S., et al. (2017). Emulating short-term and long-term plasticity of bio-synapse based on Cu/a-Si/Pt memristor. IEEE Electron. Device Lett. *38*, 1208–1211. https://doi.org/10.1109/LED.2017.2722463.

91. Wang, Y., Zhang, Z., Xu, M., Yang, Y., Ma, M., Li, H., Pei, J., and Shi, L. (2019). Self-doping memristors with equivalently synaptic ion dynamics for neuromorphic computing. ACS Appl. Mater. Interfaces *11*, 24230–24240. https://doi.org/10.1021/acsami.9b04901.

92. Cho, H., and Kim, S. (2020). Short-term memory dynamics of TiN/Ti/TiO2/SiOx/Si resistive random access memory. Nanomaterials *10*, 1821. https://doi.org/10.3390/nano10091821.

93. She, Y., Wang, F., Zhao, X., Zhang, Z., Li, C., Pan, H., Hu, K., Song, Z., and Zhang, K. (2021). Oxygen vacancy-dependent synaptic dynamic behavior of TiOx-based transparent memristor. IEEE Trans. Electron. Dev. *68*, 1950–1955. https://doi.org/10.1109/ted.2021.3056333.

94. Wan, X., Yang, Y., Feng, P., Shi, Y., and Wan, Q. (2016). Short-term plasticity and synaptic filtering emulated in electrolyte-gated IGZO transistors. IEEE Electron. Device Lett. *37*, 299–302. https://doi.org/10.1109/led.2016.2517080.

95. Wei, H., Yu, H., Gong, J., Ma, M., Han, H., Ni, Y., Zhang, S., and Xu, W. (2021). Redox MXene artificial synapse with bidirectional plasticity and hypersensitive responsibility. Adv. Funct. Mater. *31*, 2007232. https://doi.org/10.1002/adfm.202007232.

96. Wei, H., Shi, R., Sun, L., Yu, H., Gong, J., Liu, C., Xu, Z., Ni, Y., Xu, J., and Xu, W. (2021). Mimicking efferent nerves using a graphdiyne-based artificial synapse with multiple ion diffusion dynamics. Nat. Commun. *12*, 1068. https://doi.org/10.1038/s41467-021-21319-9.

97. Prasad, N., Pramanik, T., Banerjee, S.K., and Register, L.F. (2020). Realizing both short- and long-term memory within a single magnetic tunnel junction based synapse. J. Appl. Phys. *127*, 093904. https://doi.org/10.1063/1.5142418.

98. Yoon, C., Lee, J.H., Lee, S., Jeon, J.H., Jang, J.T., Kim, D.H., Kim, Y.H., and Park, B.H. (2017). Synaptic plasticity selectively activated by polarization-dependent energy-efficient ion migration in an ultrathin ferroelectric tunnel junction. Nano Lett. *17*, 1949–1955. https://doi.org/10.1021/acs.nanolett.6b05308.

99. Li, J., Ge, C., Du, J., Wang, C., Yang, G., and Jin, K. (2020). Reproducible ultrathin ferroelectric domain switching for high-performance neuromorphic computing. Adv. Mater. *32*, 1905764. https://doi.org/10.1002/adma.201905764.

100. Sarwat, S.G., Kersting, B., Moraitis, T., Jonnalagadda, V.P., and Sebastian, A. (2022). Phase-change memtransistive synapses for mixed-plasticity neural

computations. Nat. Nanotechnol. *17*, 507–513. https://doi.org/10.1038/s41565-022-01095-3.

101. Deng, P.Y., and Klyachko, V.A. (2011). The diverse functions of short-term plasticity components in synaptic computations. Commun. Integr. Biol. *4*, 543–548. https://doi.org/10.4161/cib.4.5.15870.

102. He, Y., Jiang, S., Chen, C., Wan, C., Shi, Y., and Wan, Q. (2021). Electrolyte-gated neuromorphic transistors for brain-like dynamic computing. J. Appl. Phys. *130*, 190904. https://doi.org/10.1063/5.0069456.

103. Fortune, E.S., and Rose, G.J. (2000). Short-term synaptic plasticity contributes to the temporal filtering of electrosensory information. J. Neurosci. *20*, 7122–7130. https://doi.org/10.1523/JNEUROSCI.20-18-07122.2000.

104. Zhu, L.Q., Wan, C.J., Guo, L.Q., Shi, Y., and Wan, Q. (2014). Artificial synapse network on inorganic proton conductor for neuromorphic systems. Nat. Commun. *5*, 3158. https://doi.org/10.1038/ncomms4158.

105. Wan, X., He, Y., Nie, S., Shi, Y., and Wan, Q. (2018). Biological band-pass filtering emulated by oxide-based neuromorphic transistors. IEEE Electron. Device Lett. *39*, 1764–1767. https://doi.org/10.1109/LED.2018.2869095.

106. Xu, C., Zhao, M.-x., Poo, M.-m., and Zhang, X.-h. (2008). GABAB receptor activation mediates frequency-dependent plasticity of developing GABAergic synapses. Nat. Neurosci. *11*, 1410–1418. https://doi.org/10.1038/nn.2215.

107. Kumar, S., Wang, X., Strachan, J.P., Yang, Y., and Lu, W.D. (2022). Dynamical memristors for higher-complexity neuromorphic computing. Nat. Rev. Mater. *7*, 575–591. https://doi.org/10.1038/s41578-022-00434-z.

108. Masterton, B., Diamond, I.T., Harrison, J.M., and Beecher, M.D. (1967). Medial superior olive and sound localization. Science *155*, 1696–1697. https://doi.org/10.1126/science.155.3770.1696-a.

109. Cheng, L., Gao, L., Zhang, X., Wu, Z., Zhu, J., Yu, Z., Yang, Y., Ding, Y., Li, C., Zhu, F., et al. (2022). A bioinspired configurable cochlea based on memristors. Front. Neurosci. *16*, 982850. https://doi.org/10.3389/fnins.2022.982850.

110. Hindo, T., and Chakrabartty, S. (2013). Chapter 2 - noise exploitation and adaptation in neuromorphic sensors. In Engineered Biomimicry, A. Lakhtakia and R.J. Martín-Palma, eds. (Elsevier), pp. 37–58. https://doi.org/10.1016/B978-0-12-415995-2.00002-7.

111. Warchol, M.E., and Dallos, P. (1990). Neural coding in the chick cochlear nucleus. J. Comp. Physiol. *166*, 721–734. https://doi.org/10.1007/BF00240021.

112. Kuba, H., Koyano, K., and Ohmori, H. (2002). Synaptic depression improves coincidence detection in the nucleus laminaris in brainstem slices of the chick embryo. Eur. J. Neurosci. *15*, 984–990. https://doi.org/10.1046/j.1460-9568.2002.01933.x.

113. Grothe, B., Pecka, M., and McAlpine, D. (2010). Mechanisms of sound localization in mammals. Physiol. Rev. *90*, 983–1012. https://doi.org/10.1152/physrev.00026.2009.

114. Francl, A., and McDermott, J.H. (2022). Deep neural network models of sound localization reveal how perception is adapted to real-world environments. Nat. Human Behav. *6*, 111–133. https://doi.org/10.1038/s41562-021-01244-z.

115. He, Y., Nie, S., Liu, R., Jiang, S., Shi, Y., and Wan, Q. (2019). Spatiotemporal information processing emulated by multiterminal neuro-transistor networks. Adv. Mater. *31*, e1900903. https://doi.org/10.1002/adma.201900903.

116. Lamprecht, R., and LeDoux, J. (2004). Structural plasticity and memory. Nat. Rev. Neurosci. *5*, 45–54. https://doi.org/10.1038/nrn1301.

117. Ferguson, G.D., Wang, H., Herschman, H.R., and Storm, D.R. (2004). Altered hippocampal short-term plasticity and associative memory in synaptotagmin IV (−/−) mice. Hippocampus *14*, 964–974. https://doi.org/10.1002/hipo.20013.

118. Moon, K., Park, S., Jang, J., Lee, D., Woo, J., Cha, E., Lee, S., Park, J., Song, J., Koo, Y., and Hwang, H. (2014). Hardware implementation of associative memory characteristics with analogue-type resistive-switching device. Nanotechnology *25*, 495204. https://doi.org/10.1088/0957-4484/25/49/495204.

119. Crow, T. (2004). Pavlovian conditioning of hermissenda: current cellular, molecular, and circuit perspectives. Learn. Mem. *11*, 229–238. https://doi.org/10.1101/lm.70704.

120. Subramanian Periyal, S., Jagadeeswararao, M., Ng, S.E., John, R.A., and Mathews, N. (2020). Halide perovskite quantum dots photosensitized-amorphous oxide transistors for multimodal synapses. Adv. Mater. Technol. *5*, 2000514. https://doi.org/10.1002/admt.202000514.

121. Li, N., He, C., Wang, Q., Tang, J., Zhang, Q., Shen, C., Tang, J., Huang, H., Wang, S., Li, J., et al. (2022). Gate-tunable large-scale flexible monolayer MoS2 devices for photodetectors and optoelectronic synapses. Nano Res. *15*, 5418–5424. https://doi.org/10.1007/s12274-022-4122-z.

122. Huang, W., Wang, Y., Zhang, Y., Zhu, J., Liu, D., Wang, J., Fan, L., Qiu, R., and Zhang, M. (2022). Intrinsically stretchable carbon nanotube synaptic transistors with associative learning ability and mechanical deformation response. Carbon *189*, 386–394. https://doi.org/10.1016/j.carbon.2021.12.081.

123. Baddeley, A. (2012). Working memory: theories, models, and controversies. Annu. Rev. Psychol. *63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422.

124. Nairne, J.S., and James, S. (2002). Remembering over the short-term: the case against the standard model. Annu. Rev. Psychol. *53*, 53–81. https://doi.org/10.1146/annurev.psych.53.100901.135131.

125. Baddeley, A. (2010). Working memory. Curr. Biol. *20*, R136–R140. https://doi.org/10.1016/j.cub.2009.12.014.

126. Ji, X., Hao, S., Lim, K.G., Zhong, S., and Zhao, R. (2021). Artificial working memory constructed by planar 2D channel memristors enabling brain-inspired hierarchical memory systems. Adv. Intell. Syst. *4*, 2100119. https://doi.org/10.1002/aisy.202100119.

127. Cao, J., Zhang, X., Cheng, H., Qiu, J., Liu, X., Wang, M., and Liu, Q. (2022). Emerging dynamic memristors for neuromorphic reservoir computing. Nanoscale *14*, 289–298. https://doi.org/10.1039/D1NR06680C.

128. Du, C., Cai, F., Zidan, M.A., Ma, W., Lee, S.H., and Lu, W.D. (2017). Reservoir computing using dynamic memristors for temporal information processing. Nat. Commun. *8*, 2204. https://doi.org/10.1038/s41467-017-02337-y.

129. Appeltant, L., Soriano, M.C., Van der Sande, G., Danckaert, J., Massar, S., Dambre, J., Schrauwen, B., Mirasso, C.R., and Fischer, I. (2011). Information processing using a single dynamical node as complex system. Nat. Commun. *2*, 468. https://doi.org/10.1038/ncomms1476.

130. Liang, X., Zhong, Y., Tang, J., Liu, Z., Yao, P., Sun, K., Zhang, Q., Gao, B., Heidari, H., Qian, H., and Wu, H. (2022). Rotating neurons for all-analog implementation of cyclic reservoir computing. Nat. Commun. *13*, 1549. https://doi.org/10.1038/s41467-022-29260-1.

131. Midya, R., Wang, Z., Asapu, S., Zhang, X., Rao, M., Song, W., Zhuo, Y., Upadhyay, N., Xia, Q., and Yang, J.J. (2019). Reservoir computing using diffusive memristors. Adv. Intell. Syst. *1*, 1900084. https://doi.org/10.1002/aisy.201900084.

132. Zhang, Z., Zhao, X., Zhang, X., Hou, X., Ma, X., Tang, S., Zhang, Y., Xu, G., Liu, Q., and Long, S. (2022). In-sensor reservoir computing system for latent fingerprint recognition with deep ultraviolet photo-synapses and memristor array. Nat. Commun. *13*, 6590. https://doi.org/10.1038/s41467-022-34230-8.

133. Moon, J., Ma, W., Shin, J.H., Cai, F., Du, C., Lee, S.H., and Lu, W.D. (2019). Temporal data classification and forecasting using a memristor-based reservoir computing system. Nat. Electron. *2*, 480–487. https://doi.org/10.1038/s41928-019-0313-3.

134. Torrejon, J., Riou, M., Araujo, F.A., Tsunegi, S., Khalsa, G., Querlioz, D., Bortolotti, P., Cros, V., Yakushiji, K., Fukushima, A., et al. (2017). Neuromorphic computing with nanoscale spintronic oscillators. Nature *547*, 428–431. https://doi.org/10.1038/nature23011.

135. Milano, G., Pedretti, G., Montano, K., Ricci, S., Hashemkhani, S., Boarino, L., Ielmini, D., and Ricciardi, C. (2022). In materia reservoir computing with a fully memristive architecture based on self-organizing nanowire networks. Nat. Mater. *21*, 195–202. https://doi.org/10.1038/s41563-021-01099-9.

136. Zheng, Q., Zhu, X., Mi, Y., Yuan, Z., and Xia, K. (2020). Recurrent neural networks made of magnetic tunnel junctions. AIP Adv. *10*, 025116. https://doi.org/10.1063/1.5143382.

137. Gartside, J.C., Stenning, K.D., Vanstone, A., Holder, H.H., Arroo, D.M., Dion, T., Caravelli, F., Kurebayashi, H., and Branford, W.R. (2022). Reconfigurable training and reservoir computing in an artificial spin-vortex ice via spin-wave fingerprinting. Nat. Nanotechnol. *17*, 460–469. https://doi.org/10.1038/s41565-022-01091-7.

138. Liu, K., Zhang, T., Dang, B., Bao, L., Xu, L., Cheng, C., Yang, Z., Huang, R., and Yang, Y. (2022). An optoelectronic synapse based on α-In2Se3 with controllable temporal dynamics for multimode and multiscale reservoir computing. Nat. Electron. *5*, 761–773. https://doi.org/10.1038/s41928-022-00847-2.

139. Zhong, Y., Tang, J., Li, X., Liang, X., Liu, Z., Li, Y., Xi, Y., Yao, P., Hao, Z., Gao, B., et al. (2022). A memristor-based analogue reservoir computing system for real-time and power-efficient signal processing. Nat. Electron. *5*, 672–681. https://doi.org/10.1038/s41928-022-00838-3.

140. Jo, S.H., Chang, T., Ebong, I., Bhadviya, B.B., Mazumder, P., and Lu, W. (2010). Nanoscale memristor device as synapse in neuromorphic systems. Nano Lett. *10*, 1297–1301. https://doi.org/10.1021/nl904092h.

141. Prezioso, M., Merrikh-Bayat, F., Hoskins, B.D., Adam, G.C., Likharev, K.K., and Strukov, D.B. (2015). Training and operation of an integrated neuromorphic network based on metal-oxide memristors. Nature *521*, 61–64. https://doi.org/10.1038/nature14441.

142. Lisman, J.E. (1997). Bursts as a unit of neural information: making unreliable synapses reliable. Trends Neurosci. *20*, 38–43. https://doi.org/10.1016/s0166-2236(96)10070-9.

143. Goldman, M.S., Maldonado, P., and Abbott, L.F. (2002). Redundancy reduction and sustained firing with stochastic depressing synapses. J. Neurosci. *22*, 584–591. https://doi.org/10.1523/JNEUROSCI.22-02-00584.2002.

144. Xiao, L., Zhang, D.-k., Li, Y.-q., Liang, P.-j., and Wu, S. (2013). Adaptive neural information processing with dynamical electrical synapses. Front. Comput. Neurosci. *7*, 36. https://doi.org/10.3389/fncom.2013.00036.

145. Fung, C., Wong, K., and Wu, S. (2012). Delay compensation with dynamical synapses. Adv. Neural Inf. Process. Syst.

146. Spalla, D., Cornacchia, I.M., and Treves, A. (2021). Continuous attractors for dynamic memories. Elife *10*, e69499. https://doi.org/10.7554/eLife.69499.

147. Kornijcuk, V., Kavehei, O., Lim, H., Seok, J.Y., Kim, S.K., Kim, I., Lee, W.-S., Choi, B.J., and Jeong, D.S. (2014). Multiprotocol-induced plasticity in artificial synapses. Nanoscale *6*, 15151–15160. https://doi.org/10.1039/c4nr03405h.

148. Leibold, C., Gundlfinger, A., Schmidt, R., Thurley, K., Schmitz, D., and Kempter, R. (2008). Temporal compression mediated by short-term synaptic plasticity. Proc. Natl. Acad. Sci. USA *105*, 4417–4422. https://doi.org/10.1073/pnas.0708711105.

149. Reifenstein, E.T., Bin Khalid, I., and Kempter, R. (2021). Synaptic learning rules for sequence learning. Elife *10*, e67171. https://doi.org/10.7554/eLife.67171.

150. Romani, S., and Tsodyks, M. (2015). Short-term plasticity based network model of place cells dynamics. Hippocampus *25*, 94–105. https://doi.org/10.1002/hipo.22355.

151. Leibold, C., and Bendels, M.H.K. (2009). Learning to discriminate through long-term changes of dynamical synaptic transmission. Neural Comput. *21*, 3408–3428. https://doi.org/10.1162/neco.2009.12-08-929.

152. Jo, S.H., Chang, T., Ebong, I., Bhadviya, B.B., Mazumder, P., and Lu, W. (2010). Nanoscale memristor device as synapse in neuromorphic systems. Nano Lett. *10*, 1297–1301. https://doi.org/10.1021/nl904092h.