Short-Term Plasticity and Long-Term Potentiation in Magnetic Tunnel Junctions: Towards Volatile Synapses

Abhronil Sengupta* and Kaushik Roy

School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA

(Received 30 October 2015; revised manuscript received 27 December 2015; published 26 February 2016)

Synaptic memory is considered to be the main element responsible for learning and cognition in humans. Although traditionally nonvolatile long-term plasticity changes are implemented in nanoelectronic synapses for neuromorphic applications, recent studies in neuroscience reveal that biological synapses undergo metastable volatile strengthening followed by a long-term strengthening provided that the frequency of the input stimulus is sufficiently high. Such "memory strengthening" and "memory decay" functionalities can potentially lead to adaptive neuromorphic architectures. In this paper, we demonstrate the close resemblance of the magnetization dynamics of a magnetic tunnel junction (MTJ) to short-term plasticity and long-term potentiation observed in biological synapses. We illustrate that, in addition to the magnitude and duration of the input stimulus, the frequency of the stimulus plays a critical role in determining long-term potentiation of the MTJ. Such MTJ synaptic memory arrays can be utilized to create compact, ultrafast, and low-power intelligent neural systems.

DOI: 10.1103/PhysRevApplied.5.024012

I. INTRODUCTION

With significant research efforts being directed to the development of neurocomputers based on the functionalities of the brain, a seismic shift is expected in the domain of computing based on the traditional von Neumann model. BrainScaleS [1], SpiNNaker [2], and the IBM TrueNorth [3] are instances of recent flagship neuromorphic projects that aim to develop brain-inspired computing platforms suitable for recognition (image, video, speech), classification, and mining problems. While Boolean computation is based on sequential fetch, decode, and execute cycles, such neuromorphic computing architectures are massively parallel and event-driven and are potentially appealing for pattern recognition tasks and cortical brain simulations. To that end, researchers propose various nanoelectronic devices where the underlying device physics offer a mapping to the neuronal and synaptic operations performed in the brain. The main motivation behind the usage of such non–von Neumann post-CMOS technologies as neural and synaptic devices stems from the fact that the significant mismatch between the CMOS transistors and the underlying neuroscience mechanisms results in significant area and energy overhead for a corresponding hardware implementation. A very popular instance is the simulation of a cat's brain on IBM's Blue Gene supercomputer where the power consumption is reported to be of the order of approximately a few megawatts [4]. While the power required to simulate the human brain rises significantly as we proceed along the hierarchy in the animal kingdom, actual power consumption in the mammalian brain is just a few tens of watts.

*asengup@purdue.edu

In a neuromorphic computing platform, synapses form the pathways between neurons, and their strength modulates the magnitude of the signal transmitted between the neurons. The exact mechanisms that underlie the "learning" or "plasticity" of such synaptic connections are still under debate. Meanwhile, researchers attempt to mimic several plasticity measurements observed in biological synapses in nanoelectronic devices like phase change memories [5], Ag-Si memristors [6] and spintronic devices [7], etc. However, the majority of the research focuses on nonvolatile plasticity changes of the synapse in response to the spiking patterns of the neurons it connects corresponding to long-term plasticity [8], and the volatility of human memory has been largely ignored. As a matter of fact, neuroscience studies performed in Refs. [9,10] demonstrate that synapses exhibit an inherent learning ability where they undergo volatile plasticity changes and ultimately undergo long-term plasticity conditionally based on the frequency of the incoming action potentials. Such volatile or metastable synaptic plasticity mechanisms can lead to neuromorphic architectures where the synaptic memory can adapt itself to a changing environment, since sections of the memory that have been not receiving frequent stimulus can now be erased and utilized to memorize more frequent information. Hence, it is necessary to include such volatile memory transition functionalities in a neuromorphic chip in order to leverage from the computational power that such metastable synaptic plasticity mechanisms have to offer.

Figure 1(a) demonstrates the biological process involved in such volatile synaptic plasticity changes. During the transmission of each action potential from the preneuron to the postneuron through the synapse, an influx of ionic species like Ca^{2+}, Na^{+}, and K^{+} causes the release of

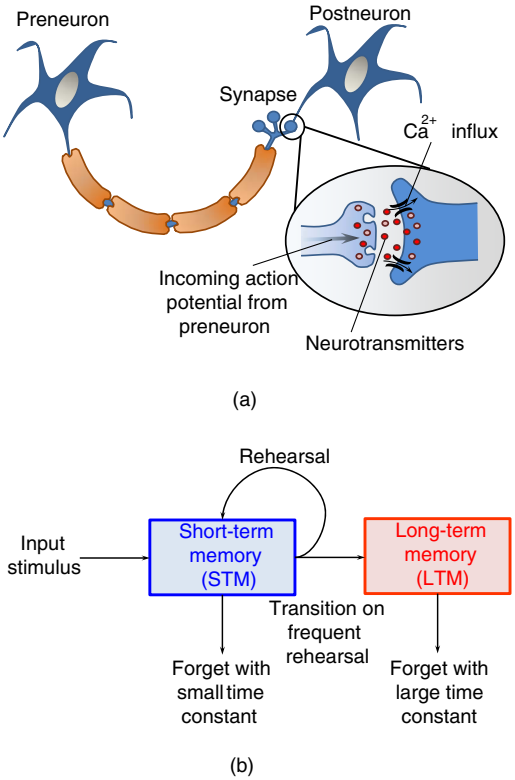

FIG. 1. (a) A synapse is a junction joining the preneuron to the postneuron. An incoming action potential from the preneuron results in the influx of ionic elements like Ca^{2+} which, in turn, results in the release of neurotransmitters at the synaptic junction. This causes STP, while frequent action potentials result in LTP. (b) Such STP and LTP mechanisms can be related to the psychological model of human memory, where memory transitions from a temporary STM to a LTM based on the frequency of rehearsal of the input stimulus.

neurotransmitters from the pre- to the postneuron. This results in temporary strengthening of the synaptic strength. However, in the absence of the action potential, the ionic species concentration settles down to its equilibrium value, and the synapse strength diminishes. This phenomenon is termed short-term plasticity (STP) [9]. However, if the action potentials occur frequently, the concentration of the ions does not get enough time to settle down to the equilibrium concentration, and this buildup of concentration eventually results in long-term strengthening of the synaptic junction. This phenomenon is termed long-term potentiation (LTP). While STP is a metastable state and lasts for a very small time duration, LTP is a stable synaptic state which can last for hours, days, or even years [10]. A similar discussion is valid for the case where there is a long-term reduction in synaptic strength with frequent stimulus, and then the phenomenon is referred to as long-term depression.

Such STP and LTP mechanisms are often correlated to the short-term memory (STM) and long-term memory (LTM) models proposed by Atkinson and Shiffrin [11,12] [Fig. 1(b)]. This psychological model partitions the human memory into a STM and a LTM. On the arrival of an input stimulus, information is first stored in the STM. However, upon frequent rehearsal, information gets transferred to the LTM. While the "forgetting" phenomena occurs at a fast rate in the STM, information can be stored for a much longer duration in the LTM.

In order to mimic such volatile synaptic plasticity mechanisms, a nanoelectronic device is required that is able to undergo metastable resistance transitions depending on the frequency of the input and also transition to a long-term stable resistance state on frequent stimulations. Hence, a competition between synaptic memory reinforcement or strengthening and memory loss is a crucial requirement for such nanoelectronic synapses. In the next section, we describe the mapping of the magnetization dynamics of a nanomagnet to such volatile synaptic plasticity mechanisms observed in the brain.

II. FORMALISM

Let us first describe the device structure and principle of operation of a MTJ [13–15] as shown in Fig. 2(a). The device consists of two ferromagnetic layers separated by a tunneling oxide barrier (TB). The magnetization of one of the layers is magnetically "pinned," and hence it is termed as the pinned layer (PL). The magnetization of the other layer, denoted as the "free layer" (FL), can be manipulated by an incoming spin current $\mathbf{I}_s$. The MTJ structure exhibits two extreme stable conductive states—the low-conductive "antiparallel" orientation (AP), where PL and FL magnetizations are oppositely directed, and the high-conductive "parallel" orientation (P), where the magnetizations of the two layers are in the same direction.

Let us consider that the initial state of the MTJ synapse is in the low-conductive AP state. Considering the input stimulus (current) to flow from terminal $T2$ to terminal $T1$, electrons flow from terminal $T1$ to $T2$ and get spin polarized by the PL of the MTJ. Subsequently, these spin-polarized electrons try to orient the FL of the MTJ parallel to the PL. It is worth noting here that the spin polarization of incoming electrons in the MTJ is analogous to the release of neurotransmitters in a biological synapse.

The STP and LTP mechanisms exhibited in the MTJ due to the spin polarization of the incoming electrons can be explained by the energy profile of the FL of the MTJ. Let the angle between the FL magnetization $\hat{\mathbf{m}}$ and the PL magnetization $\hat{\mathbf{m}}_P$ be denoted by θ. The FL energy as a function of θ is shown in Fig. 2(a), where the two energy minima points ($\theta = 0°$ and $\theta = 180°$) are separated by the energy barrier E_B. During the transition from the AP state to the P state, the FL has to transition from $\theta = 180°$ to $\theta = 0°$. Upon the receipt of an input stimulus, the FL magnetization proceeds "uphill" along the energy profile [from initial point 1 to point 2 in Fig. 2(a)]. However, since point 2 is a metastable state, it starts going "downhill" to point 1, once the stimulus is removed. If the input stimulus is not frequent

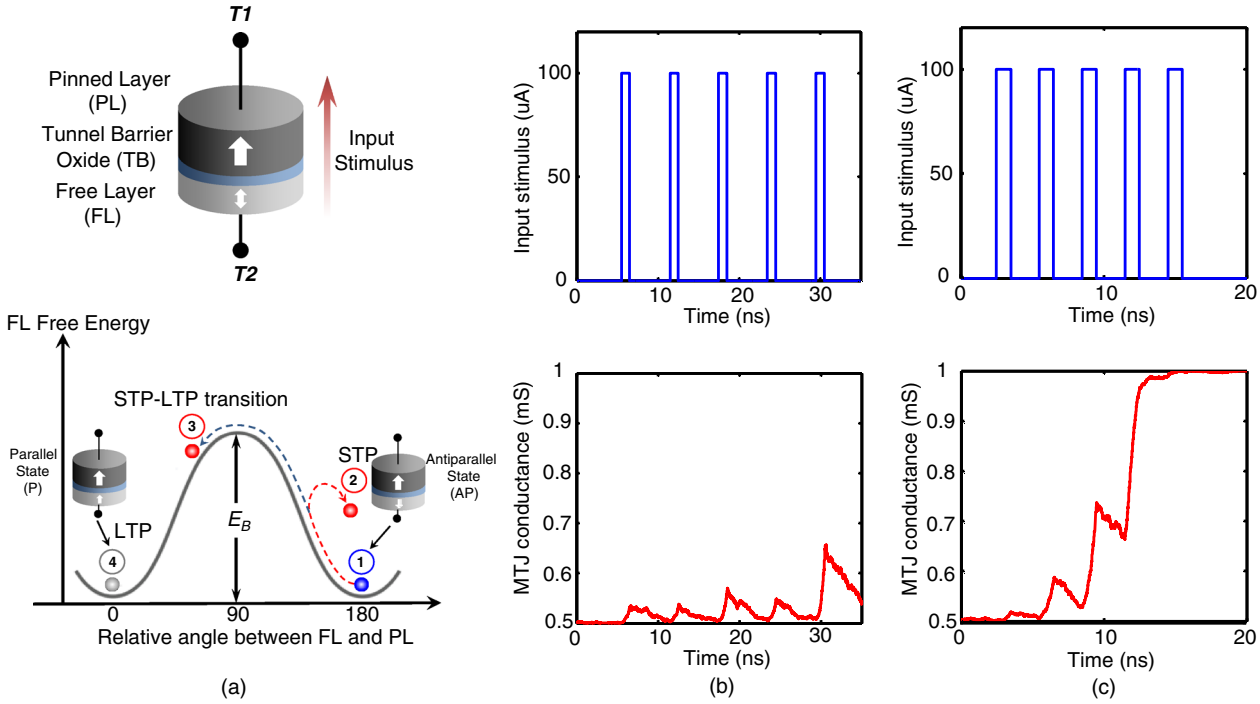

FIG. 2. (a) A MTJ structure consists of a FL separated from a PL by a TB. Initially, the MTJ synapse is in the low-conductive AP state. On receiving an input stimulus, it transitions to the high-conductive P state conditionally depending on the time interval between the inputs. The STP-LTP behavior can be explained from the energy landscape of the FL. (b) STP behavior exhibited in the MTJ synapse. The MTJ structure is an elliptic disk of volume $(\pi/4) \times 40 \times 40 \times 1.5$ nm^3 with a saturation magnetization of $M_s = 1000$ KA/m and damping factor $\alpha = 0.0122$. The AP and P conductances of the MTJ are taken to be 0.5 and 1 mS, respectively. The input stimulus is taken to be 100 μA in magnitude (assuming $\eta = 50\%$) and 1 ns in duration. The time interval between the pulses is taken to be 6 ns. (c) The MTJ synapse undergoes a LTP transition incrementally when the interval between the pulses is reduced to 3 ns.

enough, the FL tries to stabilize back to the AP state after each stimulus. However, if the stimulus is frequent, the FL will not get sufficient time to reach point 1 and ultimately will be able to overcome the energy barrier [point 3 in Fig. 2(a)]. It is worth noting here that, on crossing the energy barrier at $\theta = 90°$, it becomes progressively difficult for the MTJ to exhibit STP and switch back to the initial AP state. This is in agreement with the psychological model of human memory, where it becomes progressively difficult for the memory to "forget" information during a transition from STM to LTM. Hence, once it has crossed the energy barrier, it starts transitioning from the STP to the LTP state [point 4 in Fig. 2(a)]. The stability of the MTJ in the LTP state is dictated by the magnitude of the energy barrier. The lifetime of the LTP state is exponentially related to the energy barrier [16]. For instance, for an energy barrier of 31.44 KT used in this work, the LTP lifetime is approximately 12.4 h, while the lifetime can be extended to around 7 y by engineering a barrier height of 40 KT. The lifetime can be varied by varying the energy barrier or, equivalently, the volume of the MTJ.

The STP-LTP behavior of the MTJ can also be explained from the magnetization dynamics of the FL described by the Landau-Lifshitz-Gilbert (LLG) equation with an additional term to account for the spin-momentum torque according to Slonczewski [17]:

$$\frac{d\hat{\mathbf{m}}}{dt} = -\gamma(\hat{\mathbf{m}} \times \mathbf{H}_{\text{eff}}) + \alpha\left(\hat{\mathbf{m}} \times \frac{d\hat{\mathbf{m}}}{dt}\right) + \frac{1}{qN_s}(\hat{\mathbf{m}} \times \mathbf{I}_s \times \hat{\mathbf{m}}), \quad (1)$$

where $\hat{\mathbf{m}}$ is the unit vector of FL magnetization, $\gamma = (2\mu_B\mu_0/\hbar)$ is the gyromagnetic ratio for an electron, α is Gilbert's damping ratio, $\mathbf{H}_{\text{eff}}$ is the effective magnetic field including the shape anisotropy field for elliptic disks calculated using Ref. [18], $N_s = (M_s V/\mu_B)$ is the number of spins in a free layer of volume V (M_s is the saturation magnetization, and μ_B is the Bohr magneton), and $\mathbf{I}_s = \eta\mathbf{I}_Q$ is the spin current generated by the input stimulus $\mathbf{I}_Q$ (η is the spin-polarization efficiency of the PL). Thermal noise is included by an additional thermal field [19] $\mathbf{H}_{\text{thermal}} = \sqrt{(\alpha/1+\alpha^2)(2K_BT_K/\gamma\mu_0 M_s V\delta_t)}G_{0,1}$, where $G_{0,1}$ is a Gaussian distribution with zero mean and unit standard deviation, K_B is the Boltzmann constant, T_K is the temperature, and δ_t is the simulation time step. Equation (1) can be reformulated by simple algebraic manipulations as

$$\frac{1+\alpha^2}{\gamma}\frac{d\hat{\mathbf{m}}}{dt} = -(\hat{\mathbf{m}} \times \mathbf{H}_{\text{eff}}) - \alpha(\hat{\mathbf{m}} \times \hat{\mathbf{m}} \times \mathbf{H}_{\text{eff}}) + \frac{1}{qN_s}(\hat{\mathbf{m}} \times \mathbf{I}_s \times \hat{\mathbf{m}}). \quad (2)$$

Hence, in the presence of an input stimulus, the magnetization of the FL starts changing due to integration of the input. However, in the absence of the input, it starts leaking back due to the first two terms in the right-hand side of the above equation.

It is worth noting here that, like traditional semiconductor memories, the magnitude and duration of the input stimulus definitely have an impact on the STP-LTP transition of the synapse. However, the frequency of the input is a critical factor in this scenario. Even though the total flux through the device is the same, the synapse conditionally changes its state if the frequency of the input is high. We verify that this functionality is exhibited in MTJs by performing LLG simulations (including thermal noise). The conductance of the MTJ as a function of θ can be described by

$$G = G_{\mathrm{P}}\cos^2\left(\frac{\theta}{2}\right) + G_{\mathrm{AP}}\sin^2\left(\frac{\theta}{2}\right) \tag{3}$$

where G_{P} (G_{AP}) is the MTJ conductance in the P (AP) orientation. As shown in Fig. 2(b), the MTJ conductance undergoes metastable transitions (STP) and is not able to undergo LTP when the time interval of the input pulses is large (6 ns). However, on frequent stimulations with a time interval such as 3 ns, the device undergoes a LTP transition incrementally. Figures 2(b) and 2(c) illustrate the competition between memory reinforcement and memory decay in a MTJ structure that is crucial to implement STP and LTP in the synapse.

III. RESULTS AND DISCUSSIONS

We demonstrate simulation results to verify the STP and LTP mechanisms in a MTJ synapse depending on the time interval between stimulations. The device simulation parameters are obtained from experimental measurements [20] and are shown in Table I.

The MTJ is subjected to ten stimulations, each stimulation being a current pulse of magnitude 100 μA and 1 ns in duration. As shown in Fig. 3, the probability of a LTP transition and average device conductance at the end of each stimulation increases with a decrease in the time interval between the stimulations. The dependence on the stimulation time interval can be further characterized by measurements corresponding to paired-pulse facilitation (PPF: synaptic plasticity increase when a second stimulus follows a previous similar stimulus) and posttetanic potentiation (PTP: progressive synaptic plasticity increment when a large number of such stimuli are received successively) [21,22]. Figure 4 depicts such PPF (after second stimulus) and PTP (after tenth stimulus) measurements for the MTJ synapse with variation in the stimulation interval. The measurements closely resemble measurements performed in frog neuromuscular junctions [21], where PPF measurements reveal that there is a small synaptic

TABLE I. Device simulation parameters.

Parameters	Value
Free layer area	$(\pi/4) \times 40 \times 40$ nm^2
Free layer thickness	1.5 nm
Saturation magnetization M_S	1000 KA/m [20]
Gilbert damping factor α	0.0122 [20]
Energy barrier E_B	$31.44K_BT$
Spin-polarization strength of PL, η	0.5
MTJ conductance	0.5–1 mS
Pulse magnitude	100 μA
Pulse width t_{PW}	1 ns
Temperature T_K	300 K

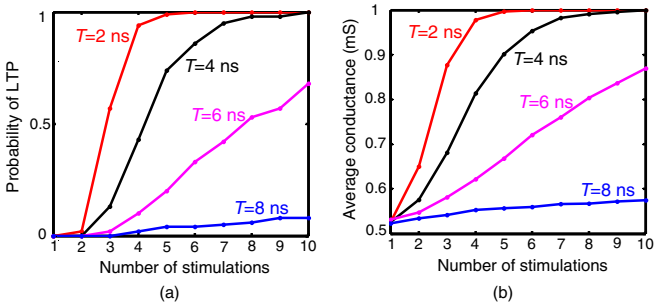

FIG. 3. (a) Stochastic LLG simulations with thermal noise performed to illustrate the dependence of the stimulation interval on the probability of a LTP transition for the MTJ. The MTJ is subjected to ten stimulations, each stimulation being a current pulse of magnitude 100 μA and 1 ns in duration. However, the time interval between the stimulations is varied from 2 to 8 ns. While the probability of LTP is 1 for a time interval of 2 ns, it is very low for a time interval of 8 ns, at the end of the ten stimulations. (b) The average MTJ conductance plotted at the end of each stimulation. As expected, the average conductance increases faster with a decrease in the stimulation interval. The results are averaged over 100 LLG simulations.

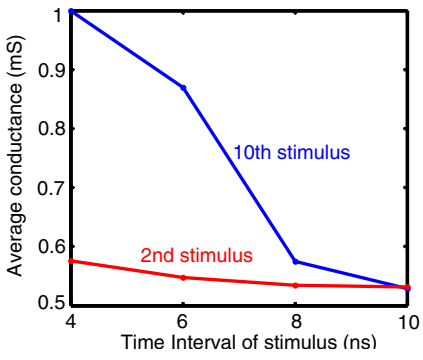

FIG. 4. PPF (average MTJ conductance after second stimulus) and PTP (average MTJ conductance after tenth stimulus) measurements in a MTJ synapse with variation in the stimulation interval. The results are in qualitative agreement with PPF and PTP measurements performed in frog neuromuscular junctions [21,22].

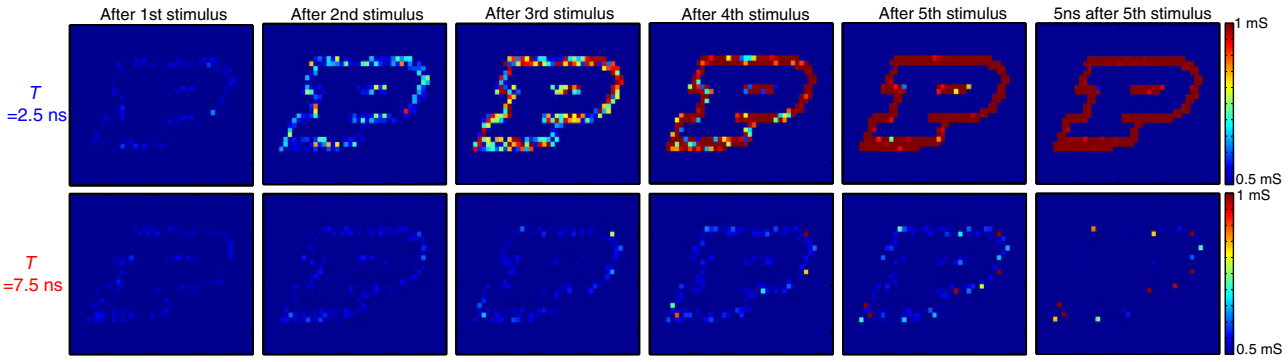

FIG. 5. STM and LTM transition exhibited in a 34×43 MTJ memory array. The input stimulus is a binary image of the Purdue University logo, where a set of five pulses (each of magnitude 100 μA and 1 ns in duration) is applied for each on pixel. While the array transitions to LTM progressively for frequent stimulations at an interval of $T = 2.5$ ns, it "forgets" the input pattern for stimulation for a time interval of $T = 7.5$ ns.

conductivity increase when the stimulation rate is frequent enough, while PTP measurements indicate a LTP transition on frequent stimulations with a fast decay in synaptic conductivity on a decrement in the stimulation rate. Hence, the stimulation rate indeed plays a critical role in the MTJ synapse to determine the probability of a LTP transition.

The psychological model of STM and LTM utilizing such MTJ synapses is further explored in a 34×43 memory array (Fig. 5). The array is stimulated by a binary image of the Purdue University logo, where a set of five pulses (each of magnitude 100 μA and 1 ns in duration) is applied for each on pixel. The snapshots of the conductance values of the memory array after each stimulus are shown for two different stimulation intervals of 2.5 and 7.5 ns, respectively. While the memory array attempts to remember the displayed image right after the stimulation, it fails to transition to LTM for the case $T = 7.5$ ns, and the information is eventually lost 5 ns after the stimulation. However, information gets transferred to LTM progressively for $T = 2.5$ ns. It is worth noting here that the same amount of flux is transmitted through the MTJ in both cases. The simulation not only provides a visual depiction of the temporal evolution of a large array of MTJ conductances as a function of stimulus but also provides inspiration for the realization of adaptive neuromorphic systems exploiting the concepts of STM and LTM. Readers interested in the practical implementation of such arrays of spintronic devices are referred to Ref. [23].

IV. CONCLUSIONS

The contributions of this work over state-of-the-art approaches may be summarized as follows. This is a theoretical demonstration of STP and LTP mechanisms in a MTJ synapse. We demonstrate the mapping of neurotransmitter release in a biological synapse to the spin polarization of electrons in a MTJ and perform extensive simulations to illustrate the impact of stimulus frequency on the LTP probability in such a MTJ structure. There are recent proposals of other emerging devices that can exhibit such STP-LTP mechanisms like Ag_2S synapses [24] and WO_X memristors [22,25]. However, it is worth noting here that input stimulus magnitudes are usually in the range of volts (1.3 V in Ref. [22] and 80 mV in Ref. [24]) and stimulus durations are of the order of a few milliseconds (1 ms in Ref. [22] and 0.5 s in Ref. [24]). In contrast, similar mechanisms can be exhibited in MTJ synapses at much lower energy consumption (by stimulus magnitudes of a few hundred microamperes and duration of a few nanoseconds). We believe that this work may stimulate proof-of-concept experiments to realize such MTJ synapses that can potentially pave the way for future ultralow-power intelligent neuromorphic systems capable of adaptive learning.

ACKNOWLEDGMENTS

The work is supported in part by Center for Spintronic Materials, Interfaces, and Novel Architectures (C-SPIN), a MARCO and DARPA sponsored StarNet center, the Semiconductor Research Corporation, the National Science Foundation, Intel Corporation, and the National Security Science and Engineering Faculty Fellowship.

[1] J. Schemmel, J. Fieres, and K. Meier, in *Proceedings of the IEEE International Joint Conference on Neural Networks, IJCNN 2008* (IEEE, New York, 2008), pp. 431–438.
[2] X. Jin, M. Lujan, L. A. Plana, S. Davies, S. Temple, and S. Furber, Modeling spiking neural networks on SpiNNaker, Comput. Sci. Eng. **12**, 91 (2010).
[3] P. Merolla, J. Arthur, F. Akopyan, N. Imam, R. Manohar, and D. S. Modha, in *Proceedings of the 2011 IEEE Custom Integrated Circuits Conference (CICC)* (IEEE, New York, 2011), pp. 1–4.
[4] S. Adee, IBM unveils a new brain simulator, IEEE Spectrum (2009), http://spectrum.ieee.org/computing/hardware/ibm-unveils-a-new-brain-simulator.

[5] B. L. Jackson *et al.*, Nanoscale electronic synapses using phase change devices, ACM J. Emerging Technol. Comput. Syst. **9**, 12 (2013).
[6] S. H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, Nanoscale memristor device as synapse in neuromorphic systems, Nano Lett. **10**, 1297 (2010).
[7] A. Sengupta, Z. Al Azim, X. Fong, and K. Roy, Spin-orbit torque induced spike-timing dependent plasticity, Appl. Phys. Lett. **106**, 093704 (2015).
[8] G.-q. Bi and M.-m. Poo, Synaptic modification by correlated activity: Hebb's postulate revisited, Annu. Rev. Neurosci. **24**, 139 (2001).
[9] R. S. Zucker and W. G. Regehr, Short-term synaptic plasticity, Annu. Rev. Physiol. **64**, 355 (2002).
[10] S. Martin, P. Grimwood, and R. Morris, Synaptic plasticity and memory: An evaluation of the hypothesis, Annu. Rev. Neurosci. **23**, 649 (2000).
[11] R. C. Atkinson, R. M. Shiffrin, K. Spence, and J. Spence, in *Advances in Research and Theory* (Academic, New York, 1968), pp. 89–195.
[12] R. Lamprecht and J. LeDoux, Structural plasticity and memory, Nat. Rev. Neurosci. **5**, 45 (2004).
[13] M. Julliere, Tunneling between ferromagnetic films, Phys. Lett. **54A**, 225 (1975).
[14] M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen Van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas, Giant Magnetoresistance of (001) Fe/(001) Cr Magnetic Superlattices, Phys. Rev. Lett. **61**, 2472 (1988).
[15] G. Binasch, P. Grünberg, F. Saurenbach, and W. Zinn, Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange, Phys. Rev. B **39**, 4828 (1989).
[16] L. Sun, Y. Hao, C.-L. Chien, and P. C. Searson, Tuning the properties of magnetic nanowires, IBM J. Res. Dev. **49**, 79 (2005).
[17] J. C. Slonczewski, Conductance and exchange coupling of two ferromagnets separated by a tunneling barrier, Phys. Rev. B **39**, 6995 (1989).
[18] M. Beleggia, M. De Graef, Y. Millev, D. Goode, and G. Rowlands, Demagnetization factors for elliptic cylinders, J. Phys. D **38**, 3333 (2005).
[19] W. Scholz, T. Schrefl, and J. Fidler, Micromagnetic simulation of thermally activated switching in fine particles, J. Magn. Magn. Mater. **233**, 296 (2001).
[20] C.-F. Pai, L. Liu, Y. Li, H. Tseng, D. Ralph, and R. Buhrman, Spin transfer torque devices utilizing the giant spin hall effect of tungsten, Appl. Phys. Lett. **101**, 122404 (2012).
[21] K. Magleby, The effect of repetitive stimulation on facilitation of transmitter release at the frog neuromuscular junction, J. Physiol. **234**, 327 (1973).
[22] T. Chang, S.-H. Jo, and W. Lu, Short-term memory to long-term memory transition in a nanoscale memristor, ACS Nano **5**, 7669 (2011).
[23] H. Noguchi *et al.*, in *Proceedings of the 2015 IEEE International Solid-State Circuits Conference (ISSCC)* (IEEE, New York, 2015), pp. 1–3.
[24] T. Ohno, T. Hasegawa, T. Tsuruoka, K. Terabe, J. K. Gimzewski, and M. Aono, Short-term plasticity and long-term potentiation mimicked in single inorganic synapses, Nat. Mater. **10**, 591 (2011).
[25] R. Yang, K. Terabe, Y. Yao, T. Tsuruoka, T. Hasegawa, J. K. Gimzewski, and M. Aono, Synaptic plasticity and memory functions achieved in a WO_{3-x}-based nanoionics device by using the principle of atomic switch operation, Nanotechnology **24**, 384003 (2013).