CHIN. PHYS. LETT. Vol. 37, No. 7 (2020) 078501

Voltage-Driven Adaptive Spintronic Neuron for Energy-Efficient Neuromorphic Computing

Ya-Bo Chen(陈亚博)[1], Xiao-Kuo Yang(杨晓阔)[1*], Tao Yan(闫涛)[2], Bo Wei(危波)[1], Huan-Qing Cui(崔焕卿)[1], Cheng Li(李成)[3], Jia-Hao Liu(刘嘉豪)[3], Ming-Xu Song(宋明旭)[1], and Li Cai(蔡理)[1]

[1]*Department of Basic Sciences, Air Force Engineering University, Xi'an 710051, China*
[2]*School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150006, China*
[3]*College of Computer, National University of Defense Technology, Changsha 410005, China*

(Received 17 April 2020; accepted 25 April 2020; published online 21 June 2020)

A spintronics neuron device based on voltage-induced strain is proposed. The stochastic switching behavior, which can mimic the firing behavior of neurons, is obtained by using two voltage signals to control the in-plane magnetization of a free layer of magneto-tunneling junction. One voltage signal is used as the input, and the other voltage signal can be used to tune the activation function (Sigmoid-like) of spin neurons. Therefore, this voltage-driven tunable spin neuron does not necessarily use energy-inefficient Oersted fields and spin-polarized current. Moreover, a voltage-control reading operation is presented, which can achieve the transition of activation function from Sigmoid-like to ReLU-like. A three-layer artificial neural network based on the voltage-driven spin neurons is constructed to recognize the handwritten digits from the MNIST dataset. For the MNIST handwritten dataset, the design achieves 97.75% recognition accuracy. The present results indicate that the voltage-driven adaptive spintronic neuron has the potential to realize energy-efficient well-adapted neuromorphic computing.

PACS: 85.75.−d, 85.80.Jm, 87.18.Sn DOI: 10.1088/0256-307X/37/7/078501

Neuromorphic computing refers to the utilization of hardware to emulate the functionalities of biological neurons and synapses.[1−3] It can overcome the limitations that traditional von Neumann machines cannot perform complex tasks, such as real-time image recognition.[4] An artificial neuron with activation function $y = F(x)$ is at the heart of any artificial neural network (ANN) for neuromorphic computing.[5] However, the energy consumption and circuit area of neuromorphic computing based on CMOS technology remain highly inefficient.[6] Recently, artificial neurons based on spintronic devices combine the advantages of high working speeds, low power consumption, and high integration capability of spintronic devices with the excellent processing capabilities of biological neurons.[7,8] To date, various spin neuron schemes such as utilizing spin-torque nano-oscillators,[9,10] magneto-tunneling junction (MTJ) with stochastic magnetization switching behavior in a free layer,[11−16] spintronic memristors,[17] and spin-torque diodes (STDs)[18] have already been proposed and verified by experiments. The energy dissipation of spintronic neuron is caused by the energy loss of the clock generating system. The driven methods of the spin neurons mentioned above are based on current-induced spin orbit torque (SOT), spin transfer torque (STT), or magnetic field. Current control magnetization has been proven to be energy-inefficient due to Joule effect.[19] Spin neuron with step activation function proposed in Ref. [20], which dissipates orders of smaller magnitude energy than the current-driven spin neuron with Sigmoid function. However, the derivative of the step function is 0 in most places (except $x = 0$), which is not conducive to updating parameters in the training process. Therefore, this spin neuron cannot be applied to large-scale neural network devices. A spin neuron driven by a mixed-mode (Voltage + STT) clock is proposed in 2019,[21] which can work as current-driven spin neuron, but it has the lower energy dissipation than current-driven spin neuron. Although this design greatly reduces the energy consumption, there is still energy-inefficient spin-polarized current. Therefore, it is necessary to design a clocking scheme with lower energy consumption to fire the stochastic switching behavior of an MTJ, so as to realize low-power spin neurons.

In this letter, we propose an energy-efficient voltage-driven spin neuron with sigmoid-like (S-type) activation function, which is implemented by the MTJ that has two discrete resistance values depending on the two magnetization orientations of the free layer. It is driven by the two-step voltage clocking scheme we designed. Energy dissipation per neuron in our design under the same specifications is much lower than the previous designs in theory. Moreover, it is worth noting that the control voltage of writing operation can tune the S-type activation function of spin neuron. The activation function type can be changed to ReLU-like function by controlling the voltage of reading operation, which improve the scalability and adaptability of the spin neuron. Furthermore, an ANN based on voltage-driven spin neurons is designed to perform handwritten digits recognition.

The voltage-driven stochastic magnetization

Supported by the National Natural Science Foundation of China under Grants Nos. 61804184 and 11975311, the Natural Science Basic Research Plan in Shaanxi Province of China under Grant No. 2020JQ470, and the Foundation of Independent Scientific Research under Grant Nos. YNJC19070501, YNJC19070502, and YNJC19070504.
*Corresponding author. Email: yangxk0123@163.com
© 2020 Chinese Physical Society and IOP Publishing Ltd

switching scheme in the free layer at the room temperature is shown in Fig. 1.(a). The magnetic layer and two electrode pairs are fabricated on the piezoelectric layer in our design. The Terfenol-D layer (magnetic layer) has a damping coefficient α of 0.1 and a saturation magnetization M_s of 8×10^5 A/m,[22] whose size is 160 nm ×120 nm ×12 nm. PMN-PT ($Pb(Mg_{1/3}Nb_{2/3})O_3$-$PbTiO_3$) layer as the piezoelectric layer (thickness $t_p = 400$ nm), which has $d_{31} = -3000$ pm/V and $d_{32} = 1000$ pm/V.[23] $U_1 = U_2 = 50$ mV of voltages would generate 30 MPa stress according to $\sigma = Y(d_{31} - d_{32})U/[(1+v)t_p]$,[11] where $Y = 200$ GPa is Young's modulus and $v = 0.3$ is Poisson's ratio. The magnetization dynamics of the magnetic layer can be obtained by solving the Landau–Lifshitz–Gilbert equation[24]

$$\frac{d\boldsymbol{M}}{dt} = -\gamma \boldsymbol{M} \times \boldsymbol{H}_{\text{eff}} - \frac{\alpha\gamma}{M_s}[\boldsymbol{M} \times (\boldsymbol{M} \times \boldsymbol{H}_{\text{eff}})], \quad (1)$$

where α is the damping coefficient, M_s is the saturation magnetization, γ is the gyromagnetic ratio, and $\boldsymbol{H}_{\text{eff}}$ is the effective field:[25]

$$H_{\text{eff}} = -\frac{1}{\mu_0 V}\frac{dE_{\text{total}}}{dM}, \quad (2)$$

where $\mu_0 = 4\pi\times10^{-7}$ is the vacuum permeability and V is the volume of single nanomagnet. The total energy E_{total} includes shape anisotropy energy, stress anisotropy energy and thermal fluctuations. The shape anisotropy energy is[26]

$$E_{\text{shape anisotropy}} = \frac{\mu_0 M_s^2 V}{2}[N_{dx}(\cos\theta\sin\phi)^2 + N_{dy}(\sin\theta\sin\phi)^2 + N_{dz}(\cos\phi)^2], \quad (3)$$

where θ is the polar angle and ϕ is the azimuth angle. N_{dx}, N_{dy}, and N_{dz} are the demagnetization factors of elliptical nanomagnets, which can be calculated by employing Ref. [27]. The stress anisotropy energy is given as follows:[28]

$$E_{\text{stress anisotropy}} = -\frac{3}{2}\lambda_s \sigma V \sin^2\theta \sin^2\phi. \quad (4)$$

Substituting Eq. (4) into Eq. (2), we can obtain the stress anisotropic field $\boldsymbol{H}_{\text{stress1}}$ generated by voltage V_1 and V_2 as follows:

$$H_{\text{stress1}}(t) = \frac{3\lambda_s}{\mu_0 M_s}\sigma(t)\sin\theta\sin(\phi-\beta), \quad (5)$$

$$H_{\text{stress2}}(t) = \frac{3\lambda_s}{\mu_0 M_s}\sigma(t)\sin\theta\sin(\phi+\beta), \quad (6)$$

where $3\lambda_s/2 = 6\times10^{-4}$ is the saturation magnetostriction, σ is the stress applied in the magnetic layer and β is the angle between the line joining the centers of the electrodes in one pair and the easy axis of nanomagnet in Fig. 1(a). The field $h(t)$ is related to thermal fluctuation energy $E_{\text{thermal fluctuations}}$, which can be expressed as[29]

$$h(t) = \sqrt{\frac{2\alpha kTf}{\gamma\mu_0 M_s V}}G_{(0,1)}(t), \quad (7)$$

where $k = 1.38\times10^{-23}$ J/K is the Boltzmann constant, $T = 300$ K is the room temperature, f is the frequency of thermal noise oscillations at room temperature, and $G_{(0,1)}(t)$ is the standard Gaussian distribution. The two-step switching clocking scheme is shown in Fig. 1(b). A voltage applied across the piezoelectric layer generates a mechanical strain in it, which is mostly transferred to the magnetostrictive layer by elastic coupling and produces a stress. This stress will cause the magnetization of the magnetostrictive layer (nanomagnet) to rotate, whose feasibility has been proven by experiments.[22,30] The inverse piezoelectric effect of piezoelectric materials and the inverse magnetostrictive effect of magnetostrictive materials are the physical essence of this clock. In the first step, a voltage V_1 is applied to the electrode pair (blue), which rotates the magnetization away from the long axis, as shown in Figs. 1(c) and 1(d). In the second step, the V_1 is withdrawn and a voltage V_2 is applied to the electrode pair (yellow) at the same time. Magnetization will be stochastically aligned to either one of the easy axis, as shown in Figs. 1(c) and 1(d). Note that the magnetization will not be stochastically rotated when we activated only one pair of electrodes, which will return to the initial state after the voltage is withdrawn.

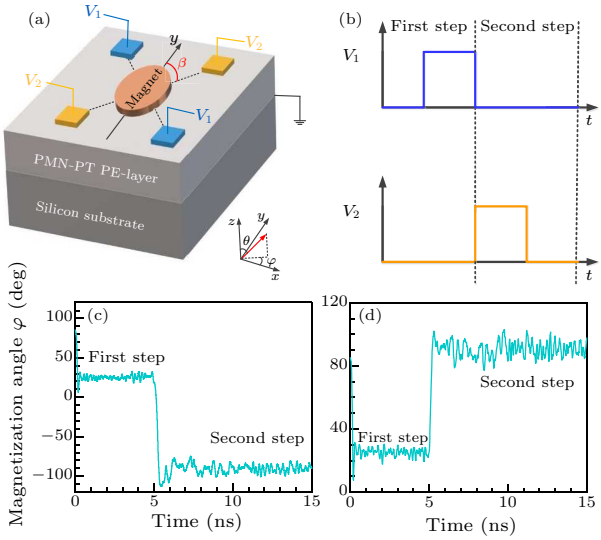

Fig. 1. (a) Voltage-driven stochastic magnetization switching scheme, where θ is the polar angle and ϕ is the azimuth angle, β is the angle between the line joining the centers of the electrodes in one pair and the easy axis of nanomagnet. (b) Two-step switching clocking consisting of two voltage signals: V_1 is the control voltage and V_2 is the input voltage V_{in}. After applying the same voltage clocking, the free layer stochastically appears to be (c) 180° magnetization switching, or (d) magnetization remains the initial state.

Since the free layer has the in-plane magnetization, and the voltage-driven 180° in-plane magnetiza-

tion switching will change the resistance of the MTJ, we choose to study the switching probability (probability of resistance change) that appears under the various input voltage. To obtain the 180° magnetization switching probability at a certain input voltage V_2, 1000 simulations in the presence of noise were carried out. The 180° switching probability can be calculated by $P = N_{180°\,\text{switching}}/1000$, where $N_{180°\,\text{switching}}$ is the number of achieving 180° magnetization switching. By changing input voltage V_2 and repeating the measurement, we can obtain the relationship between the input voltage V_2 and 180° magnetization switching probability, which is the activation characteristics of the free layer, as shown in Fig. 2(a). The switching probability increases with increasing input voltage V_2 and reaches a saturation value finally. Fitting the input-output relationship (activation function) of the device by utilizing shifted sigmoid-like function $F(x) = 1/\{1 + \exp[-a(x - b)]\}$. Encouragingly, when slope $a = 1.33$ and offset $b = 12.71$, the fitting function (blue line) has the same shape as these probability dots (orange dots), as shown in Fig. 2(a). Thereby, the fitting function obtained is the activation function of devices. To show the accuracy of the data, we repeated the above simulation for all probability dots in Fig. 2(a) to obtain standard deviation of these probability dots, as shown in Fig. 2(b). Then, we randomly select 10 groups of data, and the fitting functions of the 10-group data are consistent with the fitting function obtained in Fig. 2(a). These results indicate that the relationship between switching probability and input voltage V_2 of the device we obtained is correct and has very high reliability.

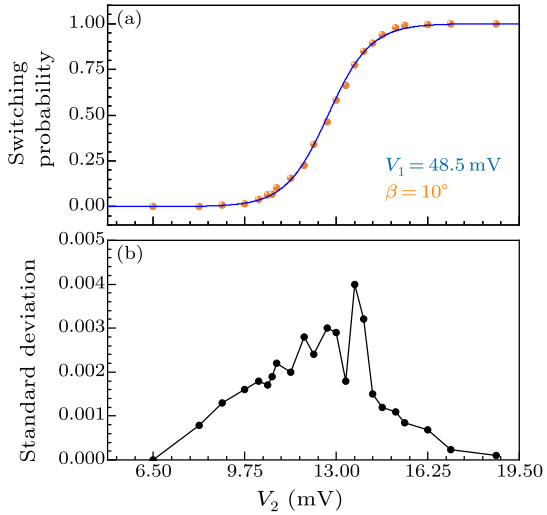

Fig. 2. (a) Switching probability under various input voltages (V_2) at $V_1 = 48.5$ mV and $\beta = 10°$ (orange dots). The data is fitted by a sigmoid-like function (blue line). (b) Standard deviation of these switching probability dots.

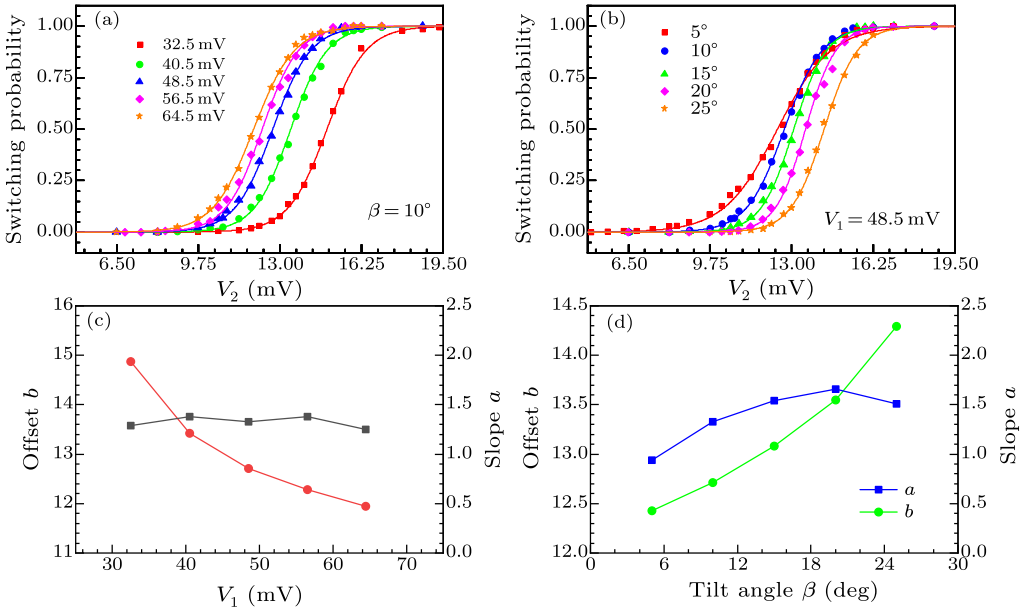

Fig. 3. Switching probability (dots) versus input voltage V_2 (a) at different control voltage V_1 and (b) at different tilt angle β, fitted by the Sigmoid-like function (line). (c) The slope a and offset b of the Sigmoid-like function versus control voltage V_1. (d) The slope a and offset b of the Sigmoid-like functions versus tilt angle β.

We found that the control voltage V_1 and angle β have greater impacts on the 180° switching probability. Therefore, we further studied the relationship between these two parameters and the activation function. In Fig. 3(a), the control voltage is swept from 32.5 mV to 64.5 mV under the angle $\beta = 10°$. Figure 3(c) shows the specific effect of the control voltage on the activation function. When the tilt angle $\beta = 10°$, the offset b of the activation function decreases with increasing voltage V_1, while the slope a of the activation function is not sensitive to changes in voltage V_1. In Fig. 3(b), the angle β is swept from 5° to 25° under the $V_1 = 48.5$ mV. The offset b of the activation function is increased along with the tilt angle β, as shown

in Fig. 3(d). There is a biological mechanism for tuning the input-output functions (S-type) of biological neuron,[31] which is conducive for neurons to implement distinct computational functions and accommodate a variety of new information introduced into the environmental change. Therefore, we can use the voltage V_1 and tilt angle β to mimic tuning mechanism of biological neuron, which can improve adaptability of the voltage-driven spintronic neuron. These results indicate that we can use the voltage-driven spintronics devices to perform as the adaptive neuron, which can further improve the neural network.

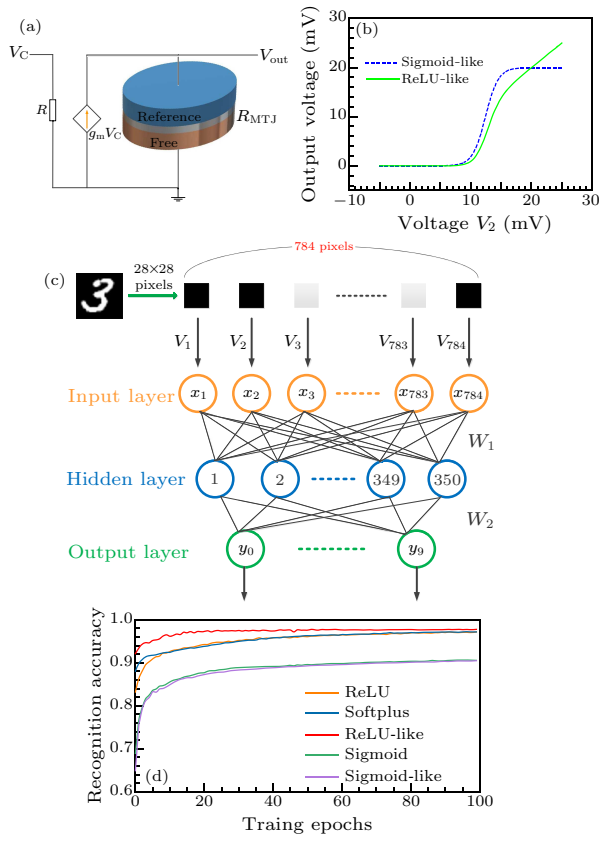

Fig. 4. (a) Schematic of voltage-driven tunable spin neuron. Voltage-driven magnetization switching is applied to the spin neuron, as shown in Fig. 1(a). The g_m is transconductance of voltage controlled current source. (b) Two activation function curves of voltage-driven spin neurons based on hardware characteristics. (c) A schematic illustration of the ANN constructed to recognize the handwriting digit in the MNIST database. (d) The recognition rate obtained from the neural network constructed with the voltage-driven spin neurons, compared to a conventional neural network with ReLU, Sigmoid and Softplus activation functions.

Rotating magnetization of a free layer from the initial state antiparallel to magnetization of the reference layer, the MTJ resistance can reach the maximum value, so that the change of magnetic moment is mapped to the change of resistance. To further improve the adaptability of spin neurons, we propose the reading circuit that can tune the activation function type by changing the reading voltage, as shown in Fig. 4(a). This structure consisting of a voltage controlled current source (VCCS) and the voltage-induced strain manipulation MTJ that can be fabricated in experiment.[32] When the control voltage V_C of VCCS is a constant, the controlled current source is a constant current source $I = g_m V_C$, then the maximum voltage across the MTJ is $V_{MTJmax} = I \times R_{MTJmax}$, where R_{MTJmax} is the maximum MTJ resistance. At this time, this structure (equivalent to DC source + MTJ) has the same function as the reading circuit (DC source + MTJ) of traditional current spin neuron proposed in Refs. [6,20]. To propagate information between neurons, we can use MTJs to convert the resistance change into voltage change. Therefore, the relationship between the output voltage and the input voltage V_2 satisfies the S-type activation function in Figs. 3(a) and 3(b). Assuming $g_m R_{MTJmax} = 1$ at room temperature, therefore $V_{MTJmax} = V_C$. When $V_{MTJmax} = V_C = 20$ mV, activation function is $V_{out} = 20/\{1+\exp[-1.33(V_2-12.71)]\}$, as shown by the blue dotted line in Fig. 4(b). The S-type activation function is commonly used in artificial neural networks, but its disadvantage is that the gradient disappears. Therefore, the ReLU function and Softplus function (smooth version of ReLU function) with better performance are introduced as activation functions of artificial neuron in neural network.[33] We further discuss how we can control the type of activation function of this neuron. When $V_{MTJmax} = V_C = V_2$, the relationship between V_{out} and V_2 changes to $V_{out} = V_2/\{1+\exp[-1.33(V_2-12.71)]\}$, as shown by the green line in Fig. 4(b). It is the same as the ReLU-like function proposed by Google in 2017, which tends to work better than the ReLU function on neural network.[34]

To verify the specific performance of spin neuron we designed, we construct an ANN based on spin neurons to recognize the handwritten digits in the MNIST database, as shown in Fig. 4(c). The artificial synapse structure that stores weights and biases can be implemented by memristors and spin synapses in Refs. [35,36]. In our simulation, a three-layer neural network model was built using the Tensorflow of Python software. All neurons in the neural network are set with the same activation function. We did not simulate the dynamic behaviors of all spin neurons, but we controlled the voltage range to ensure that the activation behaviors of all spin neurons are consistent with the fitting function in Fig. 4(b). Since traditional neurons are replaced by spin neurons we designed, the input voltages of device were used as the input value of neural network and activation function of spin neuron was used as the activation function of the neural network. The output voltage of device can be applied to the electrode pair (yellow) of the next neurons, which achieves the signal transmission. All the images in the dataset have 28×28 pixels, so the input layer has 784 neurons corresponding to 784 pixels. In practice, for a neural network based on a stochastic switching magnetic tunnel junction (spin neurons), since the data of 784 pixels is a gray value in the range of 0–255,

we need to linearly change the gray value to input voltages. We further study whether spin neurons can work as well as three kinds of traditional neurons. Standard activation functions of traditional neurons include Sigmoid function, Softplus function and ReLU function. Changing the neurons will change the activation function of the neural network, so we input five activation functions respectively into the same simulation framework (three layers neural network). Cross-entropy loss function and traditional gradient descent optimization techniques are used in the simulation. The back-propagation algorithm is used to train the network for 100 training epochs under the learning rate of 0.01. After the above simulation, we obtained the recognition rate curves of five kinds of neurons, as shown in Fig. 4(d): ReLU neuron (orange line), Softplus neuron (blue line), Sigmoid neuron (green line), Spin neuron with ReLU-like (red line) and Spin neuron with Sigmoid-like (purple line). The recognition accuracy of ANN based on spin neurons with sigmoid-like activation function is 90.45%, which can work as the artificial neuron with sigmoid function (green line). The recognition accuracy of spintronic ANN with ReLU-like activation function is 97.75%, which is consistent with those of the ReLU and Softplus function, as shown in Fig. 4(d). These results indicate that the energy-efficient spin neurons we designed can replace traditional neurons in theory. It must be aware that the input voltage range should within the limits range of the hardware and within the nonlinear region of the activation function at the same time. Moreover, the learning rate of the neural network can affect the final recognition accuracy, so our recognition results are obtained at a specific learning rate (0.01).

In summary, we have presented a voltage-driven spin neuron model that can be used for energy-efficient neuromorphic computing. The S-type activation function of this spin neuron can be controlled by a control voltage and tilt angle, which shows the ability to mimic the tuning mechanism of biological neurons. We can change activation function type from Sigmoid-like to ReLU-like by controlling voltage of reading circuit, which extends the application range of this adaptive spin neurons. The ANN constructed by voltage-driven spin neurons is used to recognize the handwritten digits in the MNIST database, whose accuracy is 90.45% (Sigmoid-like function) and 97.75% (ReLU-like function) under the learning rate of 0.01. This work provides an effective alternative to the design of lower energy consumption well-adapted spintronic neuron for neuromorphic computing.

References

[1] Zidan M A, Strachan J P, Lu W D 2018 *Nat. Electron.* **1** 22
[2] Lee J, Lu W D 2018 *Adv. Mater.* **30** 1702770
[3] Hassabis D, Kumaran D, Summerfield C, Botvinick M 2017 *Neuron* **95** 245
[4] Kurenkov A, Duttagupta S, Zhang C *et al.* 2019 *Adv. Mater.* **31** 1900636
[5] Camsari Y K, Sutton B M, Supriyo D 2019 *Appl. Phys. Rev.* **6** 011305
[6] Sengupta A, Roy K 2017 *Appl. Phys. Rev.* **4** 041105
[7] Sharad M, Fan D, Roy K 2013 *J. Appl. Phys.* **114** 234906
[8] Liu J H, Yang X K, Zhang M L *et al.* 2019 *IEEE Electron Device Lett.* **40** 220
[9] Torrejon J, Riou M, Araujo F A *et al.* 2017 *Nature* **547** 428
[10] Arai H, Imamura H 2017 *IEEE Trans. Magn.* **53** 1
[11] Cai J L, Fang B, Zhang L K *et al.* 2019 *Phys. Rev. Appl.* **11** 034015
[12] Kei K, Young C, Jong-Ung B *et al.* 2018 *J. Phys. D* **51** 504002
[13] Mizrahi A, Hirtzlin T, Fukushima A *et al.* 2018 *Nat. Commun.* **9** 1533
[14] Sengupta A, Parsa M, Han B, Roy K 2016 *IEEE Trans. Electron Devices* **63** 2963
[15] Agudov N V, Krichigin A V, Valenti D *et al.* 2010 *Phys. Rev. E* **81** 051123
[16] Suh D I, Bae G Y, Oh H S *et al.* 2015 *J. Appl. Phys.* **117** 17D714
[17] Cai J L, Fang B, Wang C *et al.* 2017 *Appl. Phys. Lett.* **111** 182410
[18] Cai J L, Zhang L K, Fang B *et al.* 2019 *Appl. Phys. Lett.* **114** 192402
[19] Vacca M, Graziano M, Crescenzo D L *et al.* 2014 *IEEE Trans. Nanotechnol.* **13** 963
[20] Biswas A K, Atulasimha J, Bandyopadhyay S 2015 *Nanotechnology* **26** 285201
[21] Nasrin S, Drobitch J L, Bandyopadhyay S *et al.* 2019 *IEEE Electron Device Lett.* **40** 345
[22] Biswas A K, Ahmad H, Atulasimha J *et al.* 2017 *Nano Lett.* **17** 3478
[23] Jin T L, Hao L, Cao J W *et al.* 2014 *Appl. Phys. Express* **7** 043002
[24] Fidler J and Schrefl T 2000 *J. Phys. D* **33** R135
[25] Liu J H, Yang X K, Cui H Q *et al.* 2019 *J. Magn. Magn. Mater.* **474** 161
[26] Imre A, Csaba G, Ji L, Orlov A *et al.* 2006 *Science* **311** 205
[27] Beleggia M, Graef M D, Millev Y T *et al.* 2005 *J. Phys. D* **38** 3333
[28] Fashami M S, Roy K, Atulasimha J *et al.* 2011 *Nanotechnology* **22** 155201
[29] Fashami M S, Al-Rashid M, Sun W Y *et al.* 2016 *Nanotechnology* **27** 43LT01
[30] D'Souza N, Fashami M S, Bandyopadhyay S, Atulasimha J 2016 *Nano Lett.* **16** 1069
[31] Lovett-Barron M, Turi G F, Kaifosh P *et al.* 2012 *Nat. Neurosci.* **15** 423
[32] Chen A T, Zhao Y L, Wen Y *et al.* 2019 *Sci. Adv.* **5** eaay5141
[33] Nair V and Hinton G E 2010 *Proceedings of the 27th International Conference on Machine Learning (ICML-10)* (21–24 June 2010, Haifa, Israel) p 807
[34] Ramachandran P, Zoph B and Le Q V 2018 arXiv:1710.05941v2 [cs.NE]
[35] Locatelli N, Cros V, Grollier J 2014 *Nat. Mater.* **13** 11
[36] Borders W A, Akima H, Fukami S *et al.* 2017 *Appl. Phys. Express* **10** 013007