Letter

Microwave-Modulated Spintronic Stochastic Neuron Based on a Magnetic Tunnel Junction

Yanxiang Luo,[1,2,*] Huayao Tu,[1,2] Like Zhang,[3,†] Shangkun Li,[1,2] Rongxin Li,[2] Jialin Cai,[2] Baoshun Zhang,[1,2] Bin Fang,[1,2,‡] and Zhongming Zeng[1,2,4,§]

[1] *School of Nano Technology and Nano Bionics, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China*

[2] *Nanofabrication Facility, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou, Jiangsu 215123, People's Republic of China*

[3] *School of Electronic and Information Engineering, Wuxi University, Wuxi, 214105, People's Republic of China*

[4] *Nanchang (SINANONC) Nano-Devices and Technologies Division, Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Nanchang 330200, People's Republic of China*

(Received 17 January 2023; revised 14 April 2023; accepted 6 June 2023; published 14 July 2023)

The stochastic switching behavior of magnetic tunnel junctions (MTJs) has been proven to be suitable for low-power neuromorphic computing. In this work, we demonstrate that the microwave-modulated stochastic switching behavior in a stochastic MTJ can be used to mimic sigmoid neurons in an artificial neural network to recognize handwritten digits in the Mixed National Institute of Standards and Technology database. The learning speed could be enhanced by the microwave and the flexibility of the neurons could be modulated by the power of rf signal. By modulating the neurons, the training speed of the system could be quicker than a conventional software-based sigmoid neural network, and the recognition accuracy reaches 87% within 30 epochs (while 80 epochs for the latter). Our work provides inspiration for alternative spintronic neuromorphic computing systems.

DOI: 10.1103/PhysRevApplied.20.L011002

Introduction.—An artificial neural network proved to be an accurate way for image classification, scene recognition, and natural language processing. The accuracy can be improved by a deep neural network, but with the expense of long training time. Many hardware-assisted solutions have been proposed to accelerate training, such as distributed computing [1], graphics processing unit (GPU) and field-programmable gate array (FPGA) based hardware accelerator [2–4]. However, these methods require high power consumption or complex hardware designs. The spintronic device, which is based on the spin nature of the electrons, has many functional advantages, such as low levels of power consumption, high stability, high switching speed, and small size, making these devices promising for serving as artificial neurons [5,6] and synapses [7–9]. Recently, the modulation of neurons, which as one kind of parameter optimization, have been used as an efficient way to enhance the training speed and flexibility of a neural network [10–15], but the hardware-based neuron-modulation accelerator is still lacking.

A magnetic tunnel junction (MTJ) is a trilayer "sandwich" consisting of a free layer (FL) separated from a reference layer (RL) by an ultrathin insulating tunneling oxide. When the magnetic moment of the FL is oriented antiparallel (or parallel) to the moment of the RL, the MTJ device has a high (or low) resistance. When the energy barrier caused by the tunneling oxide is close to the thermal disturbance at room temperature, the FL magnetic moment can stochastically switch between parallel and antiparallel configurations without a writing current—a phenomenon that holds great potential for achieving more power-efficient neuromorphic computing [16–22]. For example, recent theoretical works showed that a probabilistic deep-spiking neural system based on the stochastic MTJ, achieved high recognition accuracies (96.3% and 93.8%) in a large number of complex data sets by an artificial neural network (ANN) to spiking neural net (SNN) conversion scheme [20]. In recent years, microwave signals, dc voltage, and heat have been proposed as possible tools for controlling magnetization switching [23–29] and stochastic switching [17,19,22,30–41]. In particular, magnetization switching with the assistance of microwave signals has been proposed as a promising solution in magnetic recording [31,32]. This method requires significantly less energy to switch magnetization when the microwave signal creates a microwave field (H_{rf}) to reduce the coercive field

*yxluo2020@sinano.ac.cn
†lkzhang2021wxu@163.com
‡bfang2013@sinano.ac.cn
§zmzeng2012@sinano.ac.cn

of the FL, which provides an efficiency way to control the magnetic switching behavior of stochastic MTJs.

In this work, we report a microwave-modulated stochastic artificial neuron that is based on the MTJ. The stochastic switching probability as a function of the external static magnetic field is well fitted by a sigmoid function and could be modulated by the power and frequency of the applied microwave signal. This could be interpreted by selective microwave absorption of the FL and be used to mimic the activation function of artificial neurons. The handwritten digit-recognition task is performed by a three-layer ANN based on the microwave-modulated MTJ neurons, with the recognition accuracy reaching 87%. This work provides inspiration for the design of a power-efficient spintronic neuromorphic computing system.

Methods.—Experiments. The layer structure used as the bottom contact is interfaced with the active multiple layers to fabricate the MTJ device and is composed of Ir-Mn(8)/Co-Fe(2.5)/Ru(0.85)/Co-Fe-B(2.5)/MgO(1)/Co-Fe-B(2.5), while the top contact is created with a Ta(5)/Cu(10)/Ru(7) stack (thickness in nanometers). These films are deposited on a 200-nm thermally oxidized silicon substrate by an ultrahigh vacuum magnetron sputtering system and are annealed at 300 °C for 2 h in a 1-T magnetic field.

The deposited films are processed into GSG coplanar waveguide. An electron-beam lithography system is used to pattern elliptical nanopillars to define the MTJ junction followed by Ar ion milling and SiO_2 insulation. The device in this study has nominal dimensions of 400×200 nm^2. The easy axis in the device is defined to be along the long axis of the elliptical nanopillar. After the lift-off procedure, Ti/Au electrodes are evaporated for measurements. The photograph of the final MTJ and GSG probe are shown in Fig. 1(b).

The setup for tunnel magnetoresistance, STFMR (spin-transfer ferromagnetic resonance) and stochastic switching as shown in Fig. 1(a). The applied current and voltage are recorded to obtain the resistance of MTJ using the equation: $R = V/I$. In the STFMR test, the microwave signal is injected on the MTJ through the bias tee, and the rectified voltages are measured by lock-in amplifier. The switching measurements are carried out with dc current source and nanovolt meters under or in the absence of microwave source. In order to simulate the ANN based on the proposed stochastic switching, a simulation framework written by Python is developed to evaluate the performance.

Results and Discussion.—First, we measure the tunnel magnetoresistance (TMR) of the device. The TMR of the device is defined as $\mathrm{TMR} = (R_{AP} - R_P)/R_P$, where R_{AP} (R_P) is the resistance when the magnetic moments of the FL and RL are antiparallel (parallel) to each other. Figure 1(c) displays the magnetoresistance as a function of the in-plane magnetic field with a small bias current of 10 μA. When the external magnetic field increases from −400 to 400 Oe along the easy axis, the resistance changes sharply from 257 to 120 Ω, and the TMR ratio reaches

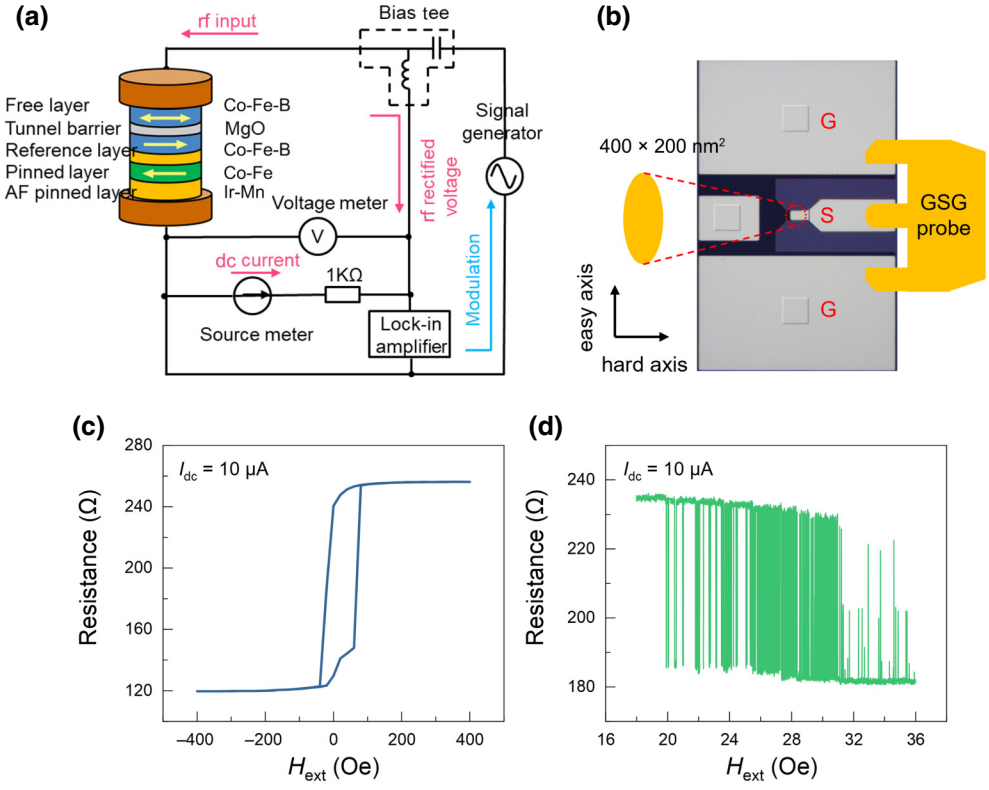

FIG. 1. Schematic of circuit (a) and MTJ (b) used for STFMR ($I_{dc} = 0$) and stochastic switching measurement ($I_{dc} = 10$ μA). Magnetoresistance as a function of the external field applied along the easy axis (c). Magnetoresistance as a function of the external field applied along the hard axis (d).

114%, indicating clear in-plane magnetization of the FL. In addition, the magnetoresistance demonstrates a slight shift toward the positive-field direction due to the stray field induced by the synthetic antiferromagnetic (SAF) layer [42,43]. Figure 1(d) shows the resistance of the MTJ as a function of the magnetic field along the hard axis, indicating stochastic switching behavior. When we sweep the magnetic field with $\Delta H = 1$ Oe, telegraphic switching appears between 18 and 32 Oe, and the two resistances are 184 and 238 Ω. The corresponding tilt angles θ between the magnetic moments of the FL and RL are estimated to be 107° and 141° according to the magnetoresistance relationship: $R^{-1}(\theta) = \left[(1+\cos\theta)R_{\mathrm{P}}^{-1} + (1-\cos\theta)R_{\mathrm{AP}}^{-1}\right]$.

Next, to further study the magnetization dynamics of the device, STFMR measurements are carried out. Figure 2(a) shows the STFMR spectra under different magnetic fields applied along the in-plane hard axis (28 to 32 Oe with the incident rf power $P_{\mathrm{rf}} = -15$ dBm). The STFMR curves show one primary peak under different magnetic fields related to microwave absorption. The STFMR spectra can be computed by fitting well with the symmetric and antisymmetric Lorentzian approximation [44]:

$$V_{\mathrm{dc}} = V_A \frac{4(f_r - f)\Delta f}{4(f_r - f)^2 + \Delta f^2} + V_S \frac{\Delta f^2}{4(f_r - f)^2 + \Delta f^2} + V_C \qquad (1)$$

where f_r is the resonance frequency, Δf is the full width at half maximum, and V_S and V_A are the amplitudes of the symmetric and antisymmetric Lorentzian parts. According to our previous work, the dominating mode (with $f_r = 1890$ MHz) comes from the quasiuniform mode at lower frequency [45].

The microwave signal with a frequency equal to resonance frequencies can compensate for the energy dissipation and lead to forced precession of the magnetic moments as shown in Fig. 2(b). When the height of the energy barrier moves into the thermal regime, the magnetic moment of FL can be switched by thermal fluctuation alone, resulting the telegraphic switching behavior. Thus, with the microwave having different power, the energy barrier can be small enough and the stochastic switching behavior can be modulated by the rf power.

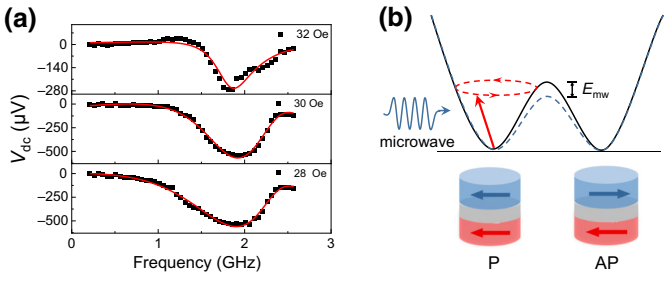

FIG. 2. STFMR spectra at different fields (H_{ext} from 28 to 32 Oe, $P_{\mathrm{rf}} = -15$ dBm) applied along the in-plane hard axis (a). Schematic of microwave-modulated stochastic switching barrier between parallel and antiparallel magnetic configurations in the device (b).

When the magnetic field ranges from 16 to 36 Oe, the telegraphic switching appears. The stochastic switching behavior under three magnetic fields (23, 25, and 28 Oe) in the time domain is shown in Fig. 3(a). Figure 3(b) shows the switching probability as a function of the external magnetic field without or with microwave signal (frequency equal to 1890 MHz and the power changes from −40 to −15 dBm). The switching probability (P) is defined as $P = \tau_{\mathrm{P}}/\tau_{\mathrm{P}} + \tau_{\mathrm{AP}}$, where τ_{AP} and τ_{P} represent the dwell time for high-to-low and low-to-high states, respectively. Previous studies have demonstrated that the switching probability in an MTJ can be expressed as an exponential function [16]:

$$P = \frac{1}{1+\exp[-4(E/kT)(H - H_s/H_{\mathrm{eff}})]} = \frac{1}{1+\exp[-\alpha(H-\beta)]}, \qquad (2)$$

where E is the energy barrier and H_{eff} is the effective field consisting of the magnetic anisotropy field (H_{ani}), exchange field (H_{ex}), shape anisotropy field (H_{sh}), and external magnetic field (H_{ext}). When the external magnetic field increases from 16 to 36 Oe, the switching probability gradually increases from 0 to 1. The data points fit the sigmoid function ($f(x) = 1/(1 + e^{-\alpha(x-\beta)})$), which is widely used for nonlinear neurons in ANNs. The curve with the rf signal has a larger slope than the microwave-free condition. This can be explained by selective microwave absorption under different magnetic fields according to Fig. 2(a). Microwaves with the frequency equal to the resonance frequency can be absorbed under magnetic fields. Then the microwave energy compensates for the energy barrier and enhances the switching probability under lower fields. In addition, there is an offset when the rf signal is applied on the device compared to the microwave-free condition, resulting from the H_{rf} of the rf signal [31].

Note that the coefficients of the neuron (α and β) in Eq. (2) play a role during the learning and training processes. The α is related to the energy barrier E, and the larger α means a higher weight-updating rate and faster obtainment of the optimal solution in an early stage of the learning process, a smaller α would be more effective in further enhancing the recognition accuracy [10–13,46,47]. In addition, the coefficient β ($\beta = H_s$, equal to the magnetic field where the switching probability is 1/2) corresponds to the bias value of the activation functions, which can be used to enhance the flexibility of the neural network. Neurons with a larger coefficient β have a lower activation threshold [14,15]. The microwave, which enables the energy to be provided to compensate the energy barrier and reduce the coercive field, provides an efficient way to modulate the coefficients of neurons.

To further provide insight into the influence of the microwave signal on the learning and training process, we obtain the values of α (0.69 for microwave-free condition and 1.5 for $f_{\mathrm{in}} = 1890$ MHz) and β by fitting the data

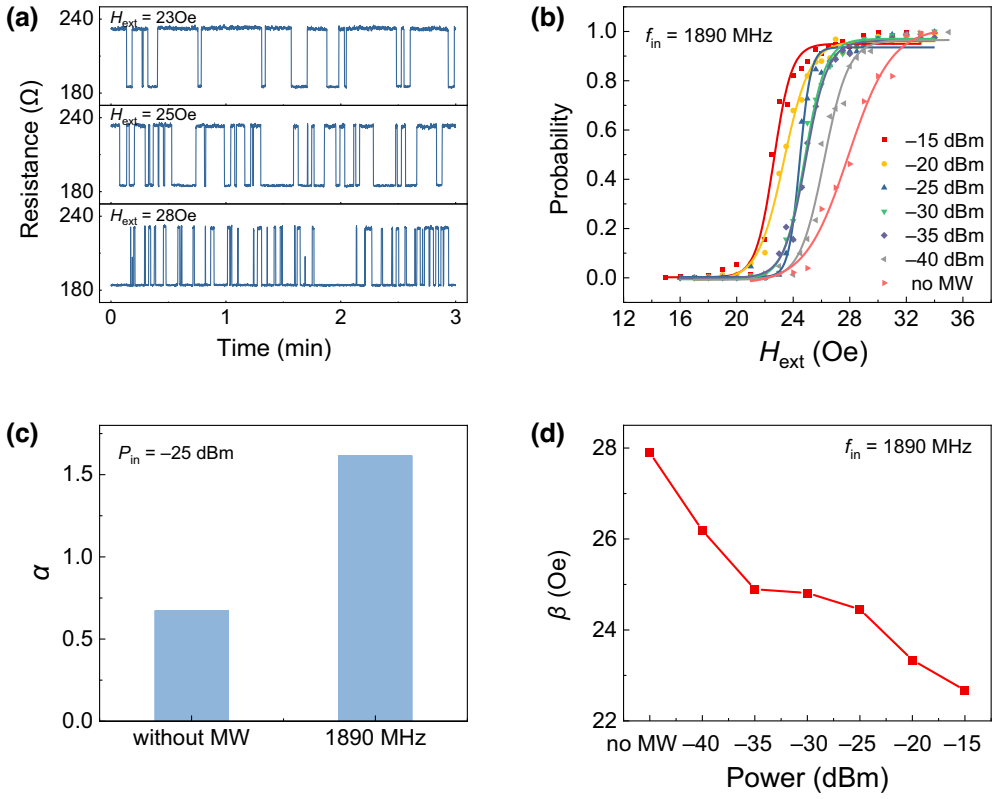

FIG. 3. Stochastic switching probability versus magnetic field under different microwave powers (power range from −15 to −40 dBm with f_{in} = 1890 MHz) (a) and time-domain stochastic switching behavior of the device under different magnetic fields (H_{ext} = 23, 25, and 28 Oe) along the hard axis (b). Parameters α of the sigmoid neuron with and without microwave (f_{in} = 1890 MHz and P_{rf} = −25 dBm) (c), and the β versus different rf power (d).

with Eq. (1), as shown in Fig. 3(c). The α ($\propto \Delta E/kTH_{eff}$) changes twofold because of selective microwave absorption as we mention before. And the learning speed will be enhanced by the microwave signal. Besides, there displays an offset along the field axis when the microwave power increases as shown in Fig. 3(b), indicating the higher stochastic switching probability behavior in higher rf power. It is because the magnetic moment of the FL obtained more energy from the microwave, resulting in an increase of the switching probability under the same magnetic field. The corresponding β are shown in Fig. 3(d). As the microwave power increases, the β decreases from 27.8 to 22.6 Oe. It is because the microwave signal provides a local magnetic field (H_{rf}, $\propto P_{rf}$) [48–51] that eliminates a part of H_s, resulting in a prominent offset along the field axis. The range of external magnetic field is from 16 to 34 Oe, with a regulation rate reaches nearly 28.8% [(27.8− 22.6 Oe)/(34− 16 Oe)].

Finally, an ANN consisting of three layers is constructed based on the stochastic neuron concept to perform the handwritten digit recognition task. This neural network includes an input layer, a hidden layer, and an output layer, as shown in Fig. 4(a). Handwritten digit images are transformed into 28 × 8 pixel values according to their pixels and then fed into 784 sigmoid neurons. The input values are processed by the 400 neurons and are output to the ten terminal neurons. The gradient-descent algorithm is used to enhance the identification performance of the network, and the recognition accuracy reached 87%. In the first 30 epochs, a neural network with two α values (1.5 and 0.69) is adjusted to enhance the recognition speed. Also, as shown in the Supplemental Material [52], neurons with higher α values have higher weight-updating speed. As shown in Fig. 4(b), the MTJ neural network has similar accuracy but is simultaneously quicker in the early stage of the task than a conventional sigmoid neural network would be.

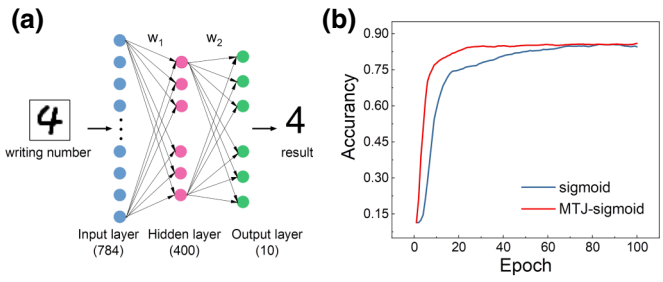

FIG. 4. Schematic diagram of the neural network (a) and recognition accuracy curves obtained from the MTJ and a conventional neural network (b).

Conclusions.—In conclusion, we investigate a microwave-modulated stochastic spintronic neuron based on an MTJ. The switching probability versus the external field is well fitted by a sigmoid function, suggesting that this device could be used as an artificial neuron in a neural network. Furthermore, the stochastic switching behavior of the MTJ could be modulated by the microwave and its power.

Such a modulation mechanism comes from the selective microwave absorption and the influence of rf power and H_{rf} on the coefficients of the neuron (α and β). Finally, a neural network based on this microwave-modulated neuron is constructed that achieved 87% recognition accuracy in handwritten digit-recognition tasks. Our work offers an alternative solution for spintronic neuromorphic computing.

Acknowledgments.—This work is supported by the National Natural Science Foundation of China (Grant No. 11974379), Natural Science Foundation of the Jiangsu Higher Education Institutions of China (Grant No. 22KJB140017) and Wuxi University Research Start-up Fund for Introduced Talents (Grant No. 2022r006), K.C. Wong Education Foundation (Grant No. GJTD-2019-14), Jiangxi Province "Double Thousand Plan" (S2019CQKJ2638), China Postdoctoral Science Foundation Funded Project (Grant No. 2019M661967 and No. 2019TQ0223), National Science Foundation of China (Grant No.12204357). This study is partially supported by the CAS Young Talent program. The authors are grateful for the technical support for Nano-X from the Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences (SINANO).

[1] J. Dean, G. S. Corrado, R. Monga, K. Chen, M. Devin, Q. V. Le, M. Z. Mao, M. A. Ranzato, A. Senior, P. Tucker, K. Yang, and A. Y. Ng, in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, Lake Tahoe, Nevada* (2012), pp. 1223.

[2] K. S. Oh and K. Jung, GPU implementation of neural networks, Pattern Recognit. **37**, 1311 (2004).

[3] S. Biookaghazadeh, P. K. Ravi, and M. Zhao, Toward multi-FPGA acceleration of the neural networks, ACM J. Emerging Technol. Comput. Syst. **17**, 23 (2021).

[4] R. Wu, X. M. Guo, J. Du and J, and B. Li, Accelerating neural network inference on FPGA-based platforms - a survey, Electronics **10**, 25 (2021).

[5] J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, and J. Grollier, Neuromorphic computing with nanoscale spintronic oscillators, Nature **547**, 428 (2017).

[6] M. Romera, P. Talatchian, S. Tsunegi, F. Abreu Araujo, V. Cros, P. Bortolotti, J. Trastoy, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. Ernoult, D. Vodenicarevic, T. Hirtzlin, N. Locatelli, D. Querlioz, and J. Grollier, Vowel recognition with four coupled spin-torque nano-oscillators, Nature **563**, 230 (2018).

[7] K. Yue, Y. Liu, R. K. Lake, and A. C. Parker, A brain-plausible neuromorphic on-the-fly learning system implemented with magnetic domain wall analog memristors, Sci. Adv. **5**, eaau8170 (2019).

[8] J. L. Cai, B. Fang, C. Wang, and Z. M. Zeng, Multilevel storage device based on domain-wall motion in a magnetic tunnel junction, Appl. Phys. Lett **111**, 182410 (2017).

[9] Y. Cao, A. W. Rushforth, Y. Sheng, H. Z. Zheng, and K. Y. Wang, Tuning a binary ferromagnet into a multistate synapse with spin-orbit-torque-induced plasticity, Adv. Funct. Mater. **29**, 1808104 (2019).

[10] N. Kaiyrbekov, O. Krestinskaya, A. P. James, and IEEE, in *2nd International Conference on Computing and Network Communications (CoCoNet), Nazarbayev Univ, Astana, KAZAKHSTAN* (2018), pp. 216.

[11] T. Leibovich-Raveh, D. J. Lewis, S. Al-Rubaiey Kadhim, and D. Ansari, A new method for calculating individual subitizing ranges, J. Numer. Cogn. **4**, 429 (2018).

[12] Y. T. Hsieh, K. Anjum, S. J. Huang, I. Kulkarni, D. Pompili, and IEEE, in *IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), Electr Network* (2021), pp. 59.

[13] P. Hartono and S. Hashimoto, Learning from imperfect data, Appl. Soft. Comput. **7**, 353 (2007).

[14] X. L. Liang and J. Xu, Biased ReLU neural networks, Neurocomputing **423**, 71 (2021).

[15] S. Snyders and C. W. Omlin, in *Proceedings of the Eighth International Conference on Neural Information Processing (ICONI), Shanghai, Peoples R China* (2001), pp. 143.

[16] J. L. Cai, B. Fang, L. K. Zhang, W. X. Lv, B. S. Zhang, T. J. Zhou, G. Finocchio, and Z. M. Zeng, Voltage-Controlled Spintronic Stochastic Neuron Based on a Magnetic Tunnel Junction, Phys. Rev. Appl. **11**, 7 (2019).

[17] G. Srinivasan, A. Sengupta, and K. Roy, Magnetic tunnel junction based long-term short-term stochastic synapse for a spiking neural network with on-chip STDP learning, Sci. Rep. **6**, 13 (2016).

[18] A. Mizrahi, T. Hirtzlin, A. Fukushima, H. Kubota, S. Yuasa, J. Grollier, and D. Querlioz, Neural-like computing with populations of superparamagnetic basis functions, Nat. Commun. **9**, 1533 (2018).

[19] K. Z. Yang and A. Sengupta, Stochastic magnetoelectric neuron for temporal information encoding, Appl. Phys. Lett. **116**, 043701 (2020).

[20] A. Sengupta, M. Parsa, B. Han, and K. Roy, Probabilistic deep spiking neural systems enabled by magnetic tunnel junction, IEEE Trans. Electron Devices **63**, 2963 (2016).

[21] K. Y. Camsari, S. Salahuddin, and S. Datta, Implementing p-bits with embedded MTJ, IEEE Electron Device Lett. **38**, 1767 (2017).

[22] C. M. Liyanagedera, A. Sengupta, A. Jaiswal, and K. Roy, Stochastic Spiking Neural Networks Enabled by Magnetic Tunnel Junctions: From Nontelegraphic to Telegraphic Switching Regimes, Phys. Rev. Appl. **8**, 064017 (2017).

[23] H. Farkhani, T. Bohnert, M. Tarequzzaman, J. D. Costa, A. Jenkins, R. Ferreira, J. K. Madsen, and F. Moradi, LAO-NCS: Laser assisted spin torque nano oscillator-based neuromorphic computing system, Front. Neurosci. **13**, 1429 (2019).

[24] C. Mu, W. Wang, H. Xia, B. Zhang, Q. Liu, and J. Wang, Fast magnetization switching by linear vertical microwave-assisted spin-transfer torque, J. Nanosci. Nanotechnol. **12**, 7460 (2012).

[25] J. L. Drobitch, M. A. Abeed, and S. Bandyopadhyay, Precessional switching of a perpendicular anisotropy

magneto-tunneling junction without a magnetic field, Jpn. J. Appl. Phys. **56**, 100309 (2017).

[26] T. Moriyama, R. Cao, J. Q. Xiao, J. Lu, X. R. Wang, Q. Wen, and H. W. Zhang, Microwave-assisted magnetization switching of $Ni_{80}Fe_{20}$ in magnetic tunnel junctions, Appl. Phys. Lett. **90**, 152503 (2007).

[27] P. Martín Pimentel, B. Leven, B. Hillebrands, and H. Grimm, Kerr microscopy studies of microwave assisted switching, J. Appl. Phys. **102**, 063913 (2007).

[28] T. Moriyama, R. Cao, J. Q. Xiao, J. Lu, X. R. Wang, Q. Wen, and H. W. Zhang, Magnetization reversal by microwave in magnetic tunnel junctions, J. Appl. Phys. **103**, 07a906 (2008).

[29] H. Suto, T. Nagasawa, K. Kudo, K. Mizushima, and R. Sato, Microwave-assisted switching of a single perpendicular magnetic tunnel junction nanodot, Appl. Phys. Express **8**, 023001 (2015).

[30] M. Endo, S. Kanai, S. Ikeda, F. Matsukura, and H. Ohno, Electric-field effects on thickness dependent magnetic anisotropy of sputtered $MgO/Co_{40}Fe_{40}B_{20}/Ta$ structures, Appl. Phys. Lett. **96**, 212503 (2010).

[31] Y. Nozaki, K. Tateishi, S. Taharazako, M. Ohta, S. Yoshimura, and K. Matsuyama, Microwave-assisted magnetization reversal in 0.36 μm wide Permalloy wires, Appl. Phys. Lett. **91**, 122505 (2007).

[32] E. A. Montoya, S. Perna, Y. J. Chen, J. A. Katine, M. d'Aquino, C. Serpico, and I. N. Krivorotov, Magnetization reversal driven by low dimensional chaos in a nanoscale ferromagnet, Nat. Commun. **10**, 543 (2019).

[33] A. Welbourne, A. L. R. Levy, M. O. A. Ellis, H. Chen, M. J. Thompson, E. Vasilaki, D. A. Allwood, and T. J. Hayward, Voltage-controlled superparamagnetic ensembles for low-power reservoir computing, Appl. Phys. Lett. **118**, 202402 (2021).

[34] Z. Liao, K. Ma, S. Tang, M. S. Sarker, H. Yamahara, and H. Tabata, Phase locking of ultra-low power consumption stochastic magnetic bits induced by colored noise, Chaos, Solitons Fractals **151**, 111262 (2021).

[35] J. Deng, V. P. K. Miriyala, Z. Zhu, X. Fong, and G. Liang, Voltage-controlled spintronic stochastic neuron for restricted Boltzmann machine with weight sparsity, IEEE Electron Device Lett. **41**, 1102 (2020).

[36] X. Kou, W. Wang, X. Fan, L. R. Shah, R. Tao, and J. Q. Xiao, High temperature annealing induced superparamagnetism in CoFeB/MgO/CoFeB tunneling junctions, J. Appl. Phys. **108**, 083901 (2010).

[37] B. R. Zink, Y. Lv, and J.-P. Wang, Telegraphic switching signals by magnet tunnel junctions for neural spiking signals with high information capacity, J. Appl. Phys. **124**, 152121 (2018).

[38] N. Locatelli, A. Mizrahi, A. Accioly, R. Matsumoto, A. Fukushima, H. Kubota, S. Yuasa, V. Cros, L. G. Pereira, D. Querlioz, J. V. Kim, and J. Grollier, Noise-Enhanced Synchronization of Stochastic Magnetic Oscillators, Phys. Rev. Appl. **2**, 034009 (2014).

[39] W. Rippard, R. Heindl, M. Pufall, S. Russek, and A. Kos, Thermal relaxation rates of magnetic nanoparticles in the presence of magnetic fields and spin-transfer effects, Phys. Rev. B **84**, 064439 (2011).

[40] P. Talatchian, M. W. Daniels, A. Madhavan, M. R. Pufall, E. Jué, W. H. Rippard, J. J. McClelland, and M. D. Stiles, Mutual control of stochastic switching for two electrically coupled superparamagnetic tunnel junctions, Phys. Rev. B **104**, 054427 (2021).

[41] K. Hayakawa, S. Kanai, T. Funatsu, J. Igarashi, B. Jinnai, W. A. Borders, H. Ohno, and S. Fukami, Nanosecond Random Telegraph Noise in In-Plane Magnetic Tunnel Junctions, Phys. Rev. Lett. **126**, 117202 (2021).

[42] G. Han, M. Tran, C. H. Sim, J. C. Wang, K. Eason, S. T. Lim, and A. Huang, Control of offset field and pinning stability in perpendicular magnetic tunnelling junctions with synthetic antiferromagnetic coupling multilayer, J. Appl. Phys. **117**, 17b515 (2015).

[43] S. Jenkins, A. Meo, L. E. Elliott, S. K. Piotrowski, M. Bapna, R. W. Chantrell, S. A. Majetich, and R. F. L. Evans, Magnetic stray fields in nanoscale magnetic tunnel junctions, J. Phys. D: Appl. Phys. **53**, 044001 (2020).

[44] H. Kubota, A. Fukushima, K. Yakushiji, T. Nagahama, S. Yuasa, K. Ando, H. Maehara, Y. Nagamine, K. Tsunekawa, D. D. Djayaprawira, N. Watanabe, and Y. Suzuki, Quantitative measurement of voltage dependence of spin-transfer torque in MgO-based magnetic tunnel junctions, Nat. Phys. **4**, 37 (2007).

[45] L. K. Zhang, J. L. Cai, B. Fang, B. S. Zhang, L. F. Bian, M. Carpentieri, G. Finocchio, and Z. M. Zeng, Dual-band microwave detector based on magnetic tunnel junctions, Appl. Phys. Lett. **117**, 072409 (2020).

[46] K. Hu, Z. Z. Zhang, X. R. Niu, Y. Zhang, C. H. Cao, F. Xiao, and X. P. Gao, Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function, Neurocomputing **309**, 179 (2018).

[47] Y. Ho and S. Wookey, The real-world-weight cross-entropy loss function: modeling the costs of mislabeling, IEEE Access **8**, 4806 (2020).

[48] E. Albisetti, D. Petti, M. Madami, S. Tacchi, P. Vavassori, E. Riedo, and R. Bertacco, Nanopatterning spin-textures: A route to reconfigurable magnonics, AIP Adv. **7**, 055601 (2017).

[49] Y. Ding, T. J. Klemmer, and T. M. Crawford, A coplanar waveguide permeameter for studying high-frequency properties of soft magnetic materials, J. Appl. Phys. **96**, 2969 (2004).

[50] C. Swoboda, N. Kuhlmann, M. Martens, A. Vogel, and G. Meier, Dynamic coupling of magnetic resonance modes in pairs of mesoscopic rectangles, J. Appl. Phys. **114**, 043905 (2013).

[51] P. Gruszecki, M. Kasprzak, A. E. Serebryannikov, M. Krawczyk, and W. Smigaj, Microwave excitation of spin wave beams in thin ferromagnetic films, Sci. Rep. **6**, 22367 (2016).

[52] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevApplied.20.L011002 for derivations and supportive information.