

# Spin–Orbit Torque-Induced Domain Nucleation for Neuromorphic Computing

Jing Zhou, Tieyang Zhao, Xinyu Shu, Liang Liu, Weinan Lin, Shaohai Chen, Shu Shi, Xiaobing Yan, Xiaogang Liu, and Jingsheng Chen\*

Neuromorphic computing has become an increasingly popular approach for artificial intelligence because it can perform cognitive tasks more efficiently than conventional computers. However, it remains challenging to develop dedicated hardware for artificial neural networks. Here, a simple bilayer spintronic device for hardware implementation of neuromorphic computing is demonstrated. In  $L1_1$ -CuPt/CoPt bilayer, current-induced field-free magnetization switching by symmetry-dependent spin–orbit torques shows a unique domain nucleation-dominated magnetization reversal, which is not accessible in conventional bilayers. Gradual domain nucleation creates multiple intermediate magnetization states which form the basis of a sigmoidal neuron. Using the  $L1_1$ -CuPt/CoPt bilayer as a sigmoidal neuron, the training of a deep learning network to recognize written digits, with a high recognition rate (87.5%) comparable to simulation (87.8%) is further demonstrated. This work offers a new scheme of implementing artificial neural networks by magnetic domain nucleation.

are based on the von Neumann architecture,<sup>[8]</sup> where information is continuously exchanged between memory units and the central processing unit. In contrast, under neuromorphic computing, information is processed in a decentralized manner in proximity to the memory. This strategy not only lowers the energy consumption, but also improves the efficiency in performing complex cognitive tasks, such as recognition, reasoning, and interaction. Despite the robust development of the algorithms<sup>[9]</sup> for various artificial neural networks, the progress of neuromorphic computing has been constrained by a lack of dedicated hardware. To this end, many materials and structures have been proposed, such as floating gate devices,<sup>[10,11]</sup> phase change materials,<sup>[12,13]</sup> ferroelectric materials,<sup>[14,15]</sup> and resistive random-access memory.<sup>[16,17]</sup>

## 1. Introduction

Neuromorphic computing is regarded as a form of unconventional computing that emulates a biological brain on different levels of computing hierarchy.<sup>[1–7]</sup> Traditionally, computers

Spintronic devices<sup>[18–22]</sup> stand out from the above competitors, owing to their ultrafast dynamics and virtually unlimited endurance. Meanwhile, spintronic devices possess key features required for neuromorphic computing, such as nonlinearity,<sup>[20]</sup> stochasticity,<sup>[18]</sup> and nonvolatility.<sup>[1]</sup> Among the state-of-the-art research, studies on spintronic neuromorphic computing revolve around the electrical manipulation of the magnetization of a ferromagnet, by either spin-transfer torque<sup>[19,20,23]</sup> or spin–orbit torque (SOT).<sup>[21,24,25]</sup> Microscopically, the magnetization is switched via either the coherent mode or the incoherent modes, where the latter are further broken down as domain wall propagation and domain nucleation.<sup>[26]</sup> The incoherent modes are particularly valuable for neuromorphic computing, since they not only lower the switching threshold current thus improve the energy efficiency, but also provide stable intermediate states that effectively transform to memristive plasticity. Spintronic devices based on domain wall propagation have been utilized to mimic synapses<sup>[25]</sup> and spiking neurons,<sup>[4]</sup> or to demonstrate spike-timing-dependent plasticity.<sup>[1,4]</sup>


Despite the increasing popularity, current approaches of implementing spintronic devices in artificial neural networks experience challenges in two aspects. First, to achieve field-free deterministic switching, the symmetry of SOT has to be broken, usually by engineering interfacial exchange coupling,<sup>[27,28]</sup> structural asymmetry,<sup>[29,30]</sup> magnetic anisotropy,<sup>[31,32]</sup> or composition gradient.<sup>[33]</sup> These approaches rely on delicately tailored fabrication processes or complex device structures or both. Second, domain wall propagation is prone to the pinning

J. Zhou, T. Zhao, X. Shu, L. Liu, W. Lin, S. Chen, S. Shi, J. Chen  
 Department of Materials Science & Engineering  
 National University of Singapore  
 Singapore 117575, Singapore  
 E-mail: msecj@nus.edu.sg

X. Yan  
 Key Laboratory of Brain-like Neuromorphic Devices  
 and Systems of Hebei Province  
 Key Laboratory of Optoelectronic Information  
 Materials of Hebei Province  
 Hebei University  
 Baoding, Hebei 071002, China

X. Liu  
 Department of Chemistry  
 National University of Singapore  
 Singapore 117543, Singapore

J. Chen  
 Suzhou Research Institute  
 National University of Singapore  
 Suzhou 215123, China

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adma.202103672>.

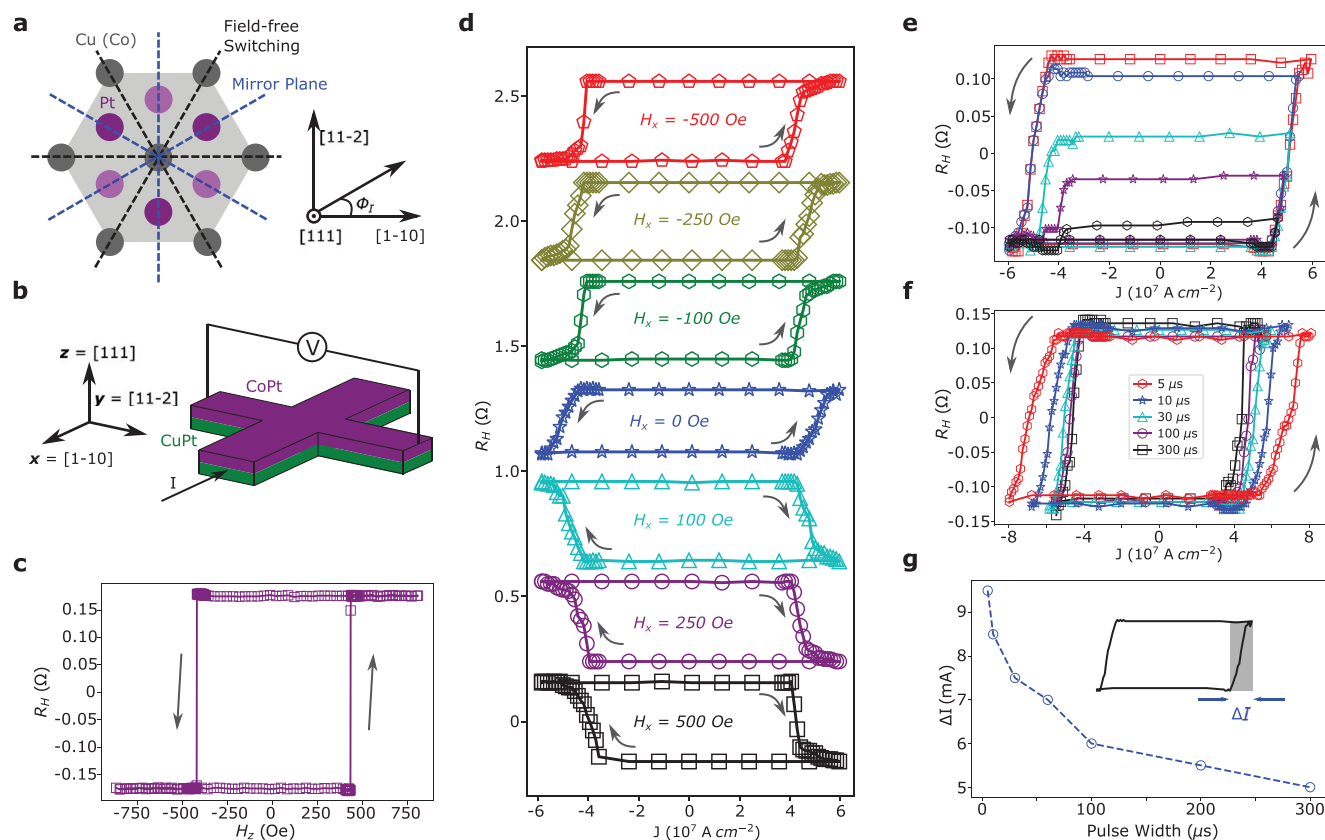
DOI: 10.1002/adma.202103672

effect of defects.<sup>[34]</sup> Although magnetic skyrmions have been proposed to alleviate the problems introduced by pinning,<sup>[35–37]</sup> the electrical manipulation of magnetic skyrmions is technically challenging as well.

We believe that the above challenges can be tackled from a material perspective. Previously, we demonstrated current-induced field-free switching in a simple bilayer structure of L1<sub>1</sub>-CuPt/CoPt, owing to an exotic symmetry-breaking SOT arising from a low-symmetry point group.<sup>[38]</sup> In this work, we show that the mode of field-free switching in the above bilayer is dominated by domain nucleation, which has been rarely employed for neuromorphic computing. Comparing to domain wall propagation, domain nucleation has a much less stringent requirement on defects, since they naturally form the basis of intermediate states. The integration of crystal symmetry and domain nucleation in magnetization switching offers excellent plasticity, allowing the L1<sub>1</sub>-CuPt/CoPt bilayer to be used as a sigmoidal artificial neuron for a deep learning network. To demonstrate our spintronic neuron's potential for on-chip training, we perform live training on the written digit dataset from the Modified National Institute of Standards and Technology (MNIST) database, and achieve a recognition rate comparable to simulation.

## 2. Results and Discussion

Single-crystalline L1<sub>1</sub>-CuPt/CoPt bilayers were deposited on SrTiO<sub>3</sub> (111) substrate at elevated temperatures using the DC magnetron sputtering technique (see the Experimental Section). Referring to **Figure 1a**, the (111) plane of the bilayer has mirror planes with a threefold symmetry. As demonstrated previously,<sup>[38]</sup> no field-free switching was observed when a charge current was applied along the mirror planes (high-symmetry axes), whereas field-free switching of the largest ratio was observed when the charge current was applied in the middle between the mirror planes (low-symmetry axes). The L1<sub>1</sub>-CuPt/CoPt bilayer was patterned into Hall bars to align the current channel with one of the low-symmetry directions ( $\phi_1 = 0^\circ, 60^\circ, 120^\circ$ ) (**Figure 1b**). **Figure 1c** shows the typical anomalous Hall resistance ( $R_H$ ) measured with the magnetic field ( $H_z$ ) applied in the out-of-plane direction. The squareness of the hysteresis loop indicates excellent perpendicular magnetic anisotropy. We applied pulsed DC current (see the Experimental Section) and measured  $R_H$  under different in-plane magnetic field ( $H_x$ ) (**Figure 1d**). When the direction of  $H_x$  was reversed, e.g., from  $-500$  to  $500$  Oe, the loop polarity was also reversed. This is because the damping-like SOT arising from the spin



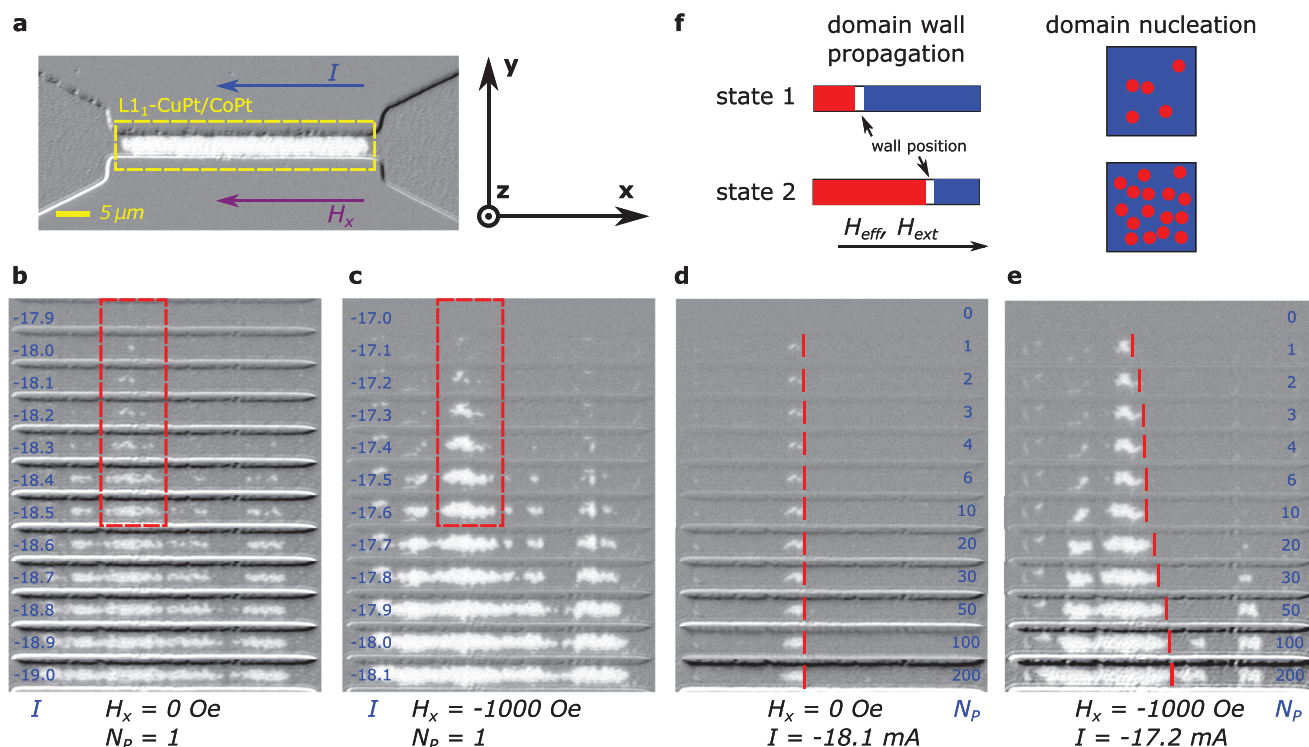
**Figure 1.** Current-induced magnetization switching in L1<sub>1</sub>-CuPt/CoPt. a) Crystal structure and symmetry. No field-free switching for current along high symmetry (mirror plane) direction. b) Schematic of the measurement coordinates. c) Anomalous Hall resistance ( $R_H$ ) versus applied out-of-plane magnetic field ( $H_z$ ). d)  $R_H$  versus current density ( $J$ ) with different in-plane magnetic fields ( $H_x$ ). Arbitrary offsets are added to separate the plots. e) Minor field-free switching loops ( $R_H$  vs  $J$ ) with different maximum  $J$ . f) Maximum field-free switching loops ( $R_H$  vs  $J$ ) with different pulses width of current. g) Tunable current range ( $\Delta I$ ) versus pulse width of current, where  $\Delta I$  covers at least 99% of changes in  $R_H$ . Arrows in (c)–(f) indicate the sweep direction.

Hall effect is odd in magnetic moment. Therefore, an opposite current-induced longitudinal effective field is required when the in-plane component of magnetic moment is flipped. In the absence of an applied magnetic field ( $H_x = 0$  Oe), deterministic bipolar switching was unambiguously observed, with a switching ratio ( $\Delta R_{\text{Ratio}}$ ) of 69.1%. Here,  $\Delta R_{\text{Ratio}}$  is defined as the ratio of current-induced maximum change in Hall resistance ( $\Delta R_{\text{AHE,I}}$ ) to field-induced maximum change in Hall resistance ( $\Delta R_{\text{AHE,H}}$ ).  $\Delta R_{\text{Ratio}}$  can be increased to 96.88% by patterning the CoPt layer into a round pillar, which alleviates the pinning effect of the Hall leads (Section S1, Supporting Information).

Figure 1e–g demonstrates additional features of the field-free switching in  $\text{L1}_1\text{-CuPt/CoPt}$  bilayer, which are crucial for its application in neuromorphic computing. Figure 1e shows the minor switching loops by limiting the maximum current density. The presence of minor loops shows that the intermediate magnetization states are stable, confirming the nonvolatility. In addition, the values of  $R_H$  are roughly the same for all minor loops at large negative pulses ( $\approx -6 \times 10^7 \text{ A cm}^{-2}$ ). This shows that the magnetization state can always be recovered, regardless of the starting point. Figure 1f shows the largest switching loops at different current pulse widths. While  $\Delta R_{\text{AHE,I}}$  decreases slightly with decreased pulse width, the squareness of the loop deteriorates, and the number of distinguishable intermediate states increases. We plot  $\Delta I$ , which is the minimum current range covering at least 99% of  $\Delta R_{\text{AHE,I}}$ , as a function of pulse width. In Figure 1g,  $\Delta I$  increases nonlinearly with decreasing pulse width, which implies remarkable plasticity since the

number of intermediate states can be finely controlled by the pulse width. Notably,  $\Delta I$  can be further increased by decreasing the pulse width to the nanosecond range.<sup>[24]</sup>

We investigated the mode of magnetization switching using the magneto-optical Kerr effect (MOKE) microscopy. Figure 2a shows an example of the MOKE image. We first initialized the magnetization state to  $+M_z$  by applying a large magnetic field in the  $+z$  direction. Then we applied current pulses with different amplitudes  $I$  and an in-plane magnetic field  $H_x$  in the direction of  $I$ . Figure 2b–e shows the changes in magnetization state marked by the yellow rectangle in Figure 2a after successive current pulses. In Figure 2b,c,  $I$  was increased (more negative) by a step of  $-0.1 \text{ mA}$  and the magnetization was captured after applying only one pulse. Comparing the areas marked by the red rectangle, the proportion of  $-M_z$  domains (white color) grows with increasing  $I$  in a dispersed manner when  $H_x = 0$ , which is akin to a domain nucleation process. On the other hand, when  $H_x = -1000$  Oe, the area of  $-M_z$  expands continuously in all directions, implying a domain wall propagation process. In Section S2 in the Supporting Information, we also show the MOKE images for switching from  $-M_z$  to  $+M_z$ . The modes of magnetization switching can be further examined by applying a number ( $N_p$ ) of current pulses with the same amplitude (Figure 2d,e). We chose a relatively small  $I$  for both  $H_x = 0$  Oe and  $H_x = -1000$  Oe, in order to observe the gradual changes. When  $H_x = 0$  Oe (Figure 2d), the right domain wall (marked by the red line) of the initially nucleated  $-M_z$  domain (when  $N_p = 1$ ) remains almost unchanged even after 200 current



**Figure 2.** Modes of incoherent magnetization switching in  $\text{L1}_1\text{-CuPt/CoPt}$ . a) Schematic of the measurement setup. b,c) Changes in magnetization state after one current pulse of different magnitudes, c) with and b) without the applied in-plane magnetic field ( $H_x$ ). d,e) Changes in the magnetization states after multiple current pulses of the same magnitude, e) with and d) without  $H_x$ .  $N_p$  refers to the number of current pulses. f) Schematic of the changes in magnetization state under domain wall propagation and domain nucleation.

pulses, while the left domain wall only moves slightly leftward. In contrast, at  $H_x = -1000$  Oe (Figure 2e), the right domain wall (marked by the red line) moves rightward unambiguously as  $N_p$  increases, and eventually merges with another propagating domain wall when  $N_p = 200$ .

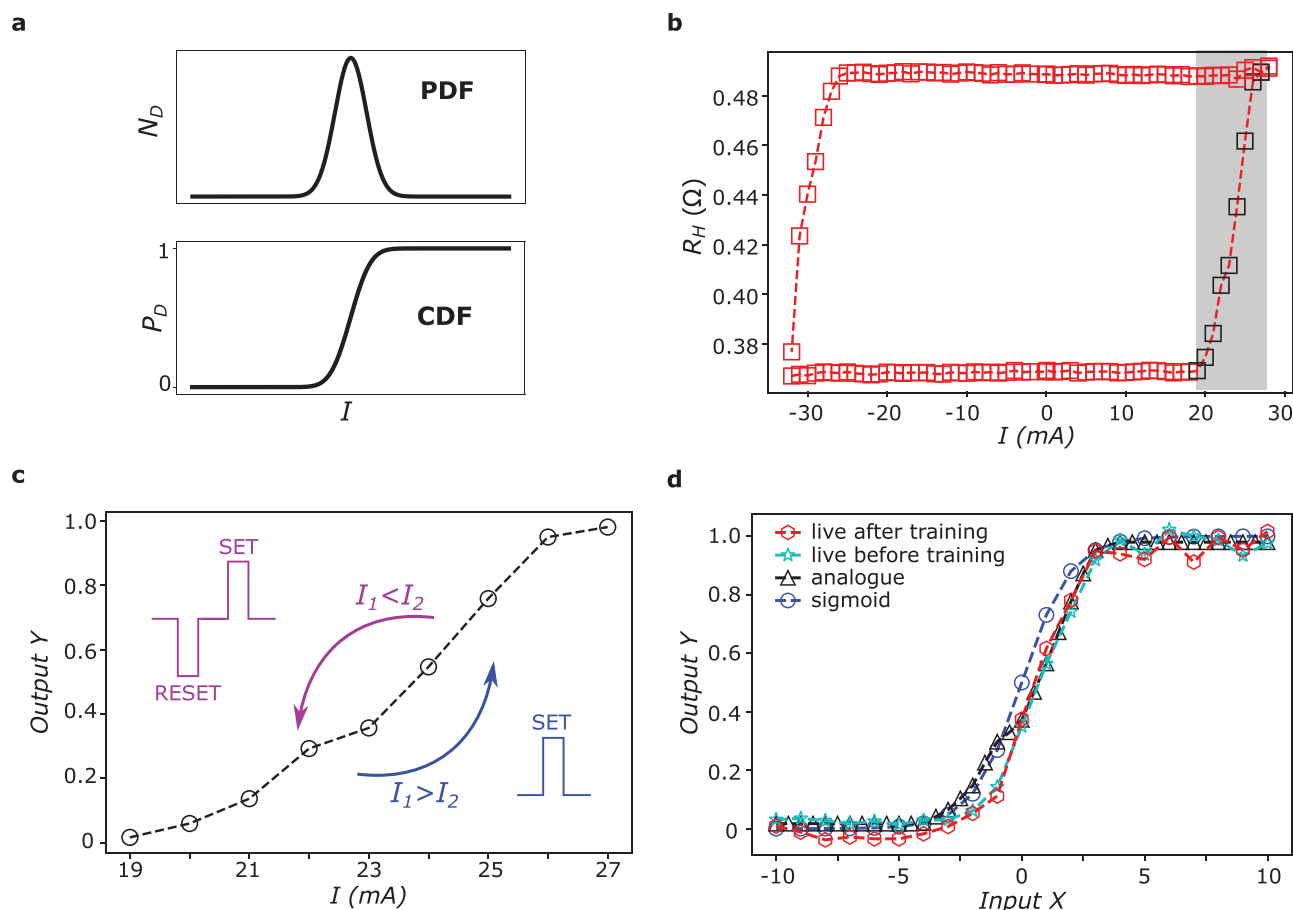
Manipulation of the magnetization switching mode can be understood by the following factors. First, the requirement of an in-plane magnetic field ( $H_x$ ) for domain wall propagation has been well documented.<sup>[39–42]</sup> Application of an  $H_x$  collinear with the current breaks the chirality imposed by Dzyaloshinskii-Moriya interaction (DMI) on the Néel domain wall. As a result, the current-induced out-of-plane effective fields on opposite domain walls have the same sign, enabling domain wall propagation in all directions. Second, domain nucleation can be understood from an energy perspective. The switched magnetization ( $m$ ) per SOT ( $\tau$ ) is proportional to the probability of switching ( $e^{-\frac{E_n}{k_B T}}$ ) and the proportion of unswitched area ( $1-m$ ), i.e.,  $\frac{dm}{d\tau} \propto e^{-\frac{E_n}{k_B T}} (1-m)$ ,<sup>[26]</sup> where  $E_n$ ,  $k_B$ , and  $T$  are the switching energy barrier, Boltzmann constant, and the temperature, respectively.  $E_n$  has a spatial distribution across the sample, which follows the normal distribution statistically. In the absence of a chirality-breaking magnetic field, the proportion of switched magnetization grows when a larger current overcomes higher  $E_n$ . This explains the variation in nucleation with current along the x-direction in Figure 2b. It also explains the switching behavior in Figure 2d, which is localized and almost independent of  $N_p$ , because only the regions with sufficiently low  $E_n$  can be switched for current pulses of a given amplitude. Under domain wall propagation, however, an additional current pulse of the same magnitude triggers additional switching events, since domain wall propagation does not require extra energy, if the domain wall pinning is neglected. This explains the gradual switching in Figure 2e as  $N_p$  increases. Therefore, although the incoherent magnetization switching is based on the combined effects of domain nucleation and domain wall propagation, domain nucleation plays a dominating role in our sample when  $H_x = 0$  Oe, while domain wall propagation is more dominant when  $H_x = -1000$  Oe. In addition, it is worth noting that magnetization switching always begins from the central area along the y-direction in Figure 2b–e. We attribute this to the higher current density at the center, which provides larger SOT. Figure 2f shows the changes in magnetization state under domain wall propagation and domain nucleation. Just like domain wall propagation, domain nucleation in our bilayer supports multiple intermediate states but without applying an external field ( $H_{\text{ext}}$ ) or engineering an effective internal field ( $H_{\text{eff}}$ ). In addition, domain nucleation is less prone to defects, which, in contrast, adversely affects the performance of devices based on domain wall propagation.

In the following sections, we used our L1-CuPt/CoPt bilayer as an artificial neuron to build and train a deep learning network. In a previous study,<sup>[26]</sup> both theoretical calculations and experimental observations show that the number of nucleated domains exhibits a normal distribution with the magnitude of SOT. We illustrate in Figure 3a that this mechanism naturally transforms our bilayer device to a domain nucleation-based sigmoidal neuron. Assuming a normal probability distribution

function (PDF) of the number of nucleated domains ( $N_D$ ) with current ( $I$ ), the cumulative distribution function takes a sigmoidal shape and describes the proportion of switched areas ( $P_D$ ). Figure 3b shows a typical switching loop. The shaded region indicates the selected data points covered by  $\Delta I$ , in a process similar to Figure 1g. Then an analog sigmoid function is constructed by mapping the selected  $R-I$  data points in Figure 3b onto a sigmoid function. The resulting analog sigmoid function is plotted in Figure 3d, which differs only slightly from an ideal sigmoid function, and thus verifies the mechanisms in Figure 3a. Based on the analog sigmoid function, a live sigmoid function was constructed. Each input  $X$  is transformed to a current  $I$  that passes through a Hall bar made of the L1-CuPt/CoPt bilayer, then the measured  $R_H$  is converted to an output  $Y$ . Two points are noteworthy in this process. First, as shown in Figure 3c, if a subsequent current pulse is higher than the previous pulse ( $I_2 > I_1$ ), only a SET pulse is required. If the  $I_2 < I_1$ , however, a RESET pulse is applied before the SET pulse to avoid the hysteresis of minor switching loops (Figure 1e). Second, in order to protect the Hall bar device from overheating, the input current  $I$  is limited within the selected current range ( $\Delta I$ ).

We performed three types of training in this work (see the Experimental Section), as illustrated in Figure 4a. The first type is based on an ideal sigmoid function. In the second type, the switching loop from a single device is measured once to construct the analog sigmoid function, and the rest of the training is performed on a computer. The third type is the focus of this work. Each neuron in the artificial neural network corresponds to a Hall bar device made of L1-CuPt/CoPt bilayer. The  $R_H-I$  data pair was continuously measured when the live sigmoid function was called during training. Figure 3d shows the live sigmoid function from an exemplary device before and after training, where  $1.2 \times 10^5$  to  $2.4 \times 10^5$  current pulses passed through the underlying Hall bar device. The live sigmoid function closely resembles the ideal and the analog sigmoid functions, and changes negligibly after the training, indicating good endurance. In Section S3 in the Supporting Information, we also show the live sigmoid functions from other devices, where all of them survive the training without substantial deterioration, indicating good repeatability across devices. We find that the differences in  $\Delta I$  are small for devices from the same sample. Nevertheless, we calibrate each device individually before training such that the mapping of  $R_H-I$  to  $Y-X$  is sufficiently accurate for all neurons. Figure 4b shows the dataset and network used in our training. We use the MNIST written digit dataset, and the classic 1-hidden-layer feed forward neural network based on gradient descent (see the Experimental Section). Figure 4c shows the circuit for training with live sigmoid function, which we develop based on a previous study.<sup>[25]</sup> We exploit the auxiliary output channels of a lock-in amplifier to control the selector circuit, which allows all circuits in Figure 4c to be controlled synchronously by a single programming language (Python 3.8 in our case) with minimal delays. In Section 4 in the Supporting Information, we present a more detailed hardware connection of the spintronic neurons. The selector circuit activates the connections to each neuron sequentially such that the neurons are accessed individually.





**Figure 3.** Spintronic sigmoidal neuron. a) Schematic illustration of the mechanism behind the domain nucleation-based sigmoidal neuron. Upper: probability distribution function (PDF) of the number of nucleated domains ( $N_D$ ) with current ( $I$ ). Lower: cumulative distribution function of the nucleated domains ( $P_D$ ) with  $I$ . b) A typical switching loop. Data points marked by black color in the shaded region are selected as the basis of the analog and live sigmoid functions. c) Illustration of SET and RESET pulses in each training cycle. d) Exemplary plots of the analog sigmoid function, and a live sigmoid function before and after training.

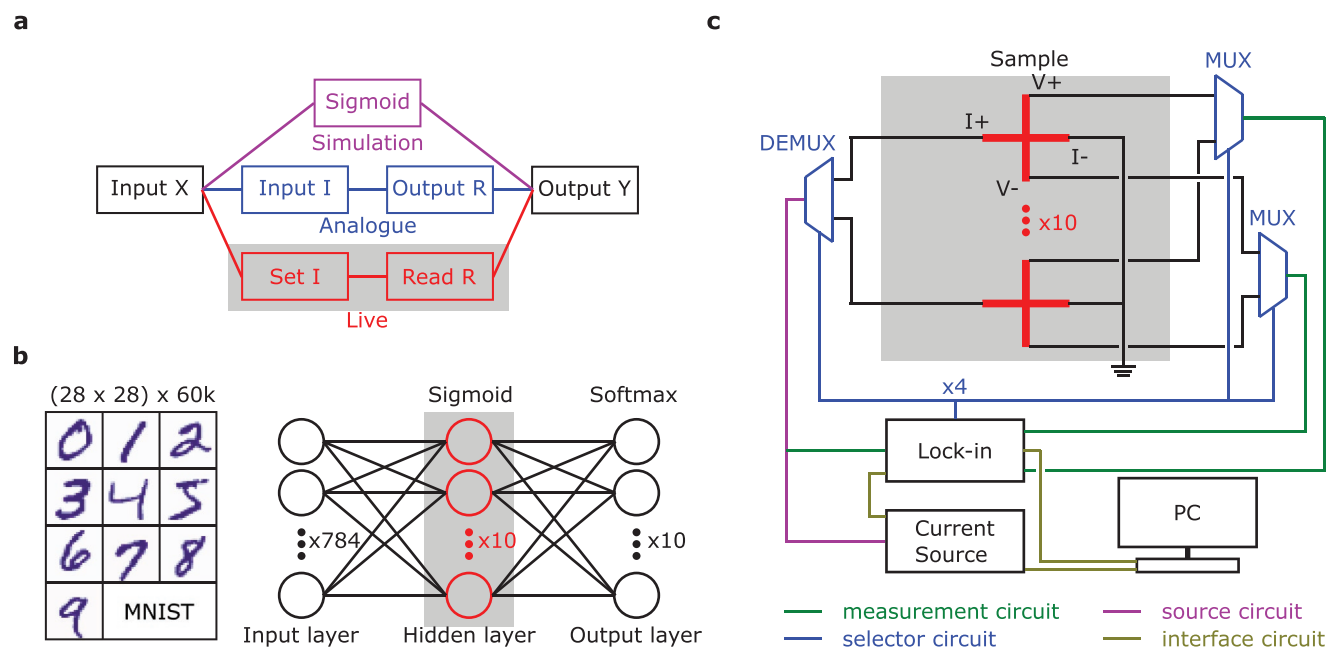
In **Figure 5a**, the recognition rates based on both the analog and live sigmoid functions are close to that of the ideal sigmoid function during the entire training process on the 60 000 training examples. The respective values of loss are shown in **Figure 5b**, all of which decrease continuously toward saturation. These results imply the plausibility of our spintronic artificial sigmoidal neuron. The recognition rate from the analog sigmoid function is even slightly better than that of the ideal sigmoid function. We speculate the reasons behind this difference to be the nonideality of the analog sigmoid function, which effectively introduces additional biases that could increase the recognition rate during the training process. The highest recognition rates achieved from training are 80.9%, 82.2%, and 78.5%, for the ideal, analog, and live sigmoid functions, respectively. We test the trained networks using the 10 000 testing examples, and obtain recognition rates of 87.8%, 89.1%, and 87.5% for the ideal, analog, and live sigmoid functions, respectively.

We further performed a control experiment to validate the effectiveness of our spintronic neuron and to confirm that our spintronic neuron is the critical component, without which training cannot be completed using simulation alone. We

deliberately overloaded the artificial neuron, by decreasing the delay time between successive pulses and using a single device to support ten neurons (**Figure 5c**, inset). The recognition rate (**Figure 5c**) and loss (**Figure 5d**) deviate substantially from those of the ideal sigmoid function after training  $\approx 1500$  examples. The recognition rate (loss) even starts to drop (rise) after training about 3500 examples. We measure the switching loop after 6000 examples. Its offset has shifted and its curvature has changed with substantially reduced  $\Delta R_{\text{AHE},1}$  (**Figure 5d**, inset). These features indicate that the device under investigation has deteriorated, probably due to overheating. These experimental results demonstrate that our artificial neural network relies on the spintronic neuron because hardware failure of the neuron unambiguously results in failure of the network.

### 3. Conclusion

In summary, we have shown that the dominating microscopic mode of field-free magnetization switching in an  $\text{LI}_1\text{-CuPt/CoPt}$  bilayer is domain nucleation rather than domain wall



**Figure 4.** Network and circuit for training. a) Schematic illustration of three types of sigmoid function used for training. b) Data set and network for training. The first samples of each digit are printed as examples. c) Schematic illustration of the circuit and connection for training with the live sigmoid function. Shaded regions in all sub figures highlight the processes that occur on Hall bar devices.

propagation. Owing to the multiple stable intermediate magnetization states, devices made of the bilayer can be utilized as sigmoidal neurons for a deep learning network. Live training based on these spintronic neurons achieves a high recognition rate on the MNIST written digit dataset, comparable to the results of pure simulation based on an ideal sigmoid function. Our work demonstrates the utility of crystal symmetry for hardware implementation of neuromorphic computing. We also highlight that domain nucleation, though less utilized, has huge potential for application in neuromorphic computing because of its advantages over domain wall propagation. We believe the results in our work provide value experimental evidences for evaluating the practical performance of spintronic neurons, and will stimulate more effort in developing all-spin artificial neural networks.

## 4. Experimental Section

**Sample Fabrication:** The  $\text{L1}_1\text{-CuPt}$  (10 nm)/CoPt (4 nm) bilayers were deposited on  $\text{SrTiO}_3$  (111) substrate using the DC magnetron sputtering method with a base pressure of less than  $1 \times 10^{-8}$  Torr. The CuPt layer was co-sputtered at 500 °C using a Cu target and a Pt target to achieve an atomic ratio of Cu:Pt = 1:1. Then the sample was cooled to 300 °C before depositing the CoPt layer, which was also co-sputtered using a Co target and a Pt target to achieve an atomic ratio of Co:Pt = 1:1. The sample was then cooled to room temperature to deposit a 2 nm protective  $\text{Si}_3\text{N}_4$  layer. The bilayer sample was patterned into Hall bars using a combination of photolithography and ion beam etching. The current channel was aligned with the low-symmetry directions (see Figure 1a), with lateral dimensions of  $3 \mu\text{m} \times 40 \mu\text{m}$ . An electrode of Ti (5 nm)/Cu (100 nm) was deposited.

**Electrical Measurement:** A Keithley 6221 source meter was used to generate DC current pulses with pulse widths of 5–300  $\mu\text{s}$ . After 2 s, a Zurich Instrument MFLI lock-in amplifier was used to pass a small AC

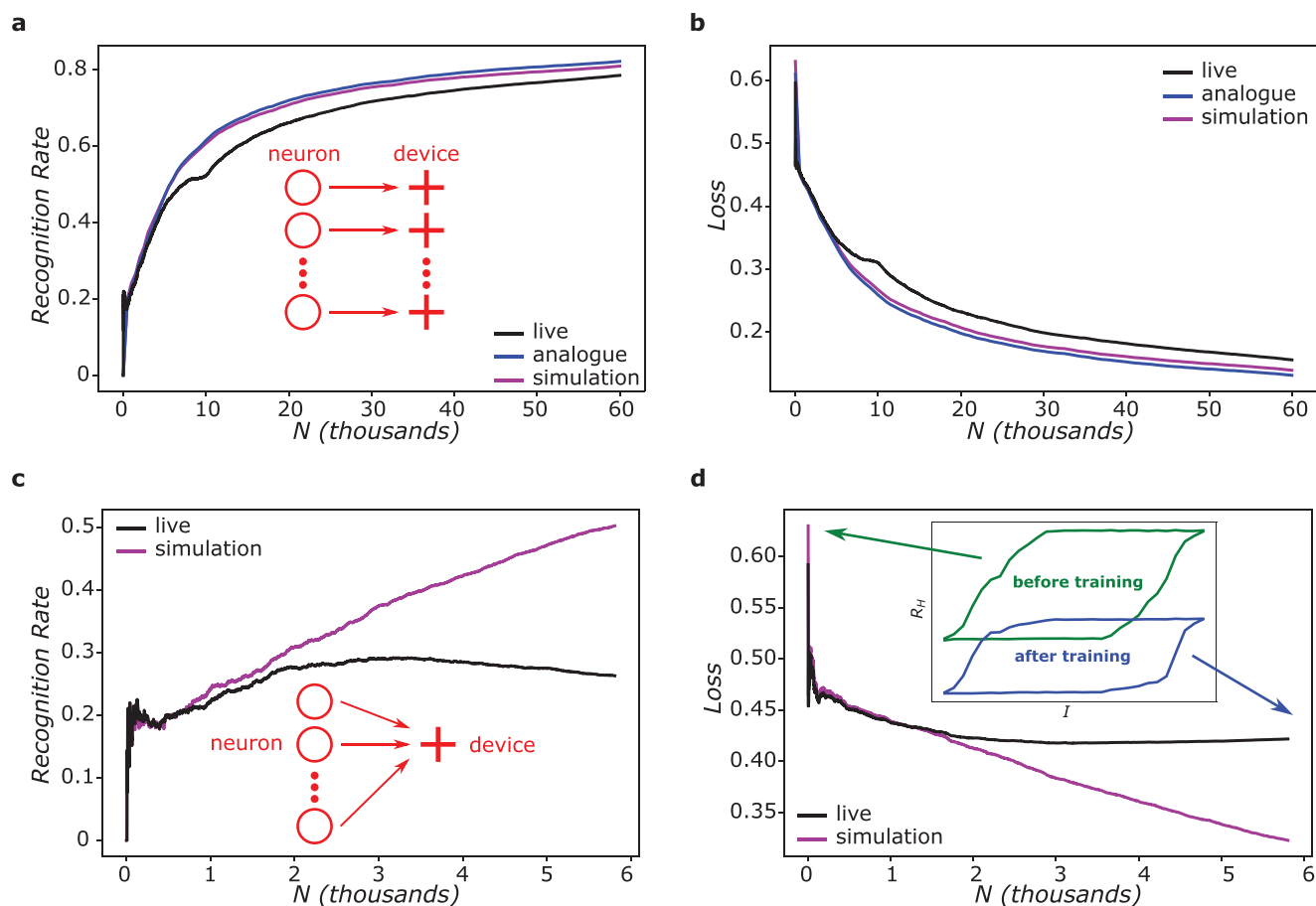
current of 50  $\mu\text{A}$  and 317.3 Hz to the current channel of the Hall bar and measured the modulated  $R_H$ .

**MOKE Measurement:** A magneto-optic Kerr effect microscope manufactured by Evico Magnetics was used to observe the magnetization state. For each series of measurement, the magnetization state was first initialized using a large out-of-plane magnetic field. The background signal was removed to enhance the contrast. Each time after a current pulse was applied, the magnetization was allowed to stabilize for 2 s before taking the image.

**Circuit for Live Training:** The input current for live training had a fixed pulse width of 5  $\mu\text{s}$  and a fixed measurement delay of 2 s, except for the part on overloading (Figure 5c,d), where the measurement delay was reduced to 1.5 s. A bilayer sample with ten Hall bar devices was mounted on a customized PCB. The devices were connected to the PCB by wire bonding. One 1:16 analog demultiplexer (DEMUX) and two 16:1 analog multiplexers (MUX) were used to select one of the ten Hall bar devices, where the extra six channels are vacant. The DEMUX and MUXs were controlled in parallel by four auxiliary output channels (since  $2^4 = 16$ ) of the MFLI lock-in amplifier, which sourced DC voltages of 2 V. The selector circuit was bypassed to overload a single Hall bar device. All instruments for live training were controlled remotely by using Python 3.8.

**Dataset and Algorithms of Training:** The written digit dataset from MNIST was consisted of 60 000 labeled training examples and 10 000 labeled testing samples, all of which had a resolution of  $28 \times 28$  pixels. A classic trilayer feed forward artificial neuron network was used, which was consisted of an input layer, a hidden layer, and an output layer. The training images were converted to 1D arrays of length 784 ( $28 \times 28 = 784$ ). Thus, the input layer had 784 neurons. Both the hidden layer and output

layers had ten neurons. A sigmoid function  $\left(y = \frac{1}{1 + e^{-x}}\right)$  was used as the activation function of the hidden layer, which was replaced by the analog and live sigmoid functions for different types of training. A softmax function  $[S(a) = \frac{e^{a_j}}{\sum_{k=1}^N e^{a_k}}$ , where  $\forall j \in 1..N$ , and  $N = 10$  in this case] was used as the activation function of the output layer. The number of recognized examples ( $N_R$ ) increased by one if the index of



**Figure 5.** Results of training using different sigmoid functions. a) Recognition rates and b) loss for training 60 000 examples using different sigmoid functions. c) Recognition rates and d) loss for training 6000 examples using an overloaded spintronic neuron. Insets in (a) and (c) show the devices supporting neurons. Inset in (d) shows the  $R_H$ - $I$  switching loop before and after training.

the output neuron with the highest probability was the same as the label. The recognition rate  $\left(\frac{N_R}{N_T}\right)$  was computed as the ratio of number of recognized examples ( $N_R$ ) over the total number of trained examples ( $N_T$ ). A simple mean square function was used as the loss function [ $L = \frac{1}{2N_T} \sum (\hat{y}_d - y_d)^2$ , where  $\hat{y}_d$  is the computed output and  $y_d$  is the actual output (label)]. To complete one training cycle, backward propagation and gradient descent were performed, where the detailed procedures are presented in Section S5 in the Supporting Information.

## Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

## Acknowledgements

J.Z. and T.Z. contributed equally to this work. The research was supported by the Singapore Ministry of Education (MOE Tier1 R-284-000-195-114; MOE2018-T2-2-043), A\*STAR (Grant no. A1983c0036, IAF-ICP 11801E0036), and Singapore National Research Foundation under CRP Award (Grant no. NRF-CRP23-2019-0070). J.C. is a member of the Singapore Spintronics Consortium (SG-SPIN).

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Keywords

crystal symmetry, deep learning network, domain nucleation, neuromorphic computing, spin-orbit torques

Received: May 14, 2021  
Revised: July 6, 2021  
Published online: July 24, 2021

- [1] I. Chakraborty, A. Jaiswal, A. K. Saha, S. K. Gupta, K. Roy, *Appl. Phys. Rev.* **2020**, *7*, 021308.
- [2] K. Roy, A. Jaiswal, P. Panda, *Nature* **2019**, *575*, 607.

- [3] J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, M. D. Stiles, *Nat. Electron.* **2020**, 3, 360.
- [4] A. Sengupta, K. Roy, *Appl. Phys. Rev.* **2017**, 4, 041105.
- [5] D. Ielmini, S. Ambrogio, *Nanotechnology* **2020**, 31, 092001.
- [6] G. Finocchio, M. D. Ventra, K. Y. Camsari, K. Everschor-Sitte, P. K. Amiri, Z. Zeng, *J. Magn. Magn. Mater.* **2021**, 521, 167506.
- [7] Z. Wang, H. Wu, G. W. Burr, C. S. Hwang, K. L. Wang, Q. Xia, J. J. Yang, *Nat. Rev. Mater.* **2020**, 5, 173.
- [8] J. V. Neumann, *IEEE Ann. Hist. Comput.* **1993**, 15, 27.
- [9] W. Maass, *Neural Networks* **1997**, 10, 1659.
- [10] V. Kornijcuk, H. Lim, J. Y. Seok, G. Kim, S. K. Kim, I. Kim, B. J. Choi, D. S. Jeong, *Front. Neurosci.* **2016**, 10, 212.
- [11] S. Ramakrishnan, P. E. Hasler, C. Gordon, *IEEE Trans. Biomed. Circuits Syst.* **2011**, 5, 244.
- [12] S. Ambrogio, N. Ciocchini, M. Laudato, V. Milo, A. Pirovano, P. Fantini, D. Ielmini, *Front. Neurosci.* **2016**, 10, 56.
- [13] D. Kuzum, R. G. Jeyasingh, B. Lee, H.-S. P. Wong, *Nano Lett.* **2012**, 12, 2179.
- [14] H. Mulaosmanovic, E. Chicca, M. Bertele, T. Mikolajick, S. Slesazek, *Nanoscale* **2018**, 10, 21755.
- [15] S. Boyn, J. Grollier, G. Lecerf, B. Xu, N. Locatelli, S. Fusil, S. Girod, C. Carrétéro, K. Garcia, S. Xavier, J. Tomas, L. Bellaiche, M. Bibes, A. Barthélémy, S. Saïghi, V. Garcia, *Nat. Commun.* **2017**, 8, 14736.
- [16] C. Sung, H. Hwang, I. K. Yoo, *J. Appl. Phys.* **2018**, 124, 151903.
- [17] Y. Zhang, P. Huang, B. Gao, J. Kang, H. Wu, *J. Phys. D: Appl. Phys.* **2021**, 54, 083002.
- [18] K. Y. Camsari, B. M. Sutton, S. Datta, *Appl. Phys. Rev.* **2019**, 6, 011305.
- [19] A. Sengupta, P. Panda, P. Wijesinghe, Y. Kim, K. Roy, *Sci. Rep.* **2016**, 6, 30039.
- [20] M. Romera, P. Talatchian, S. Tsunegi, F. A. 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, J. Grollier, *Nature* **2018**, 563, 230.
- [21] M. Zahedinejad, A. A. Awad, S. Muralidhar, R. Khymyn, H. Fulara, H. Mazraati, M. Dvornik, J. Åkerman, *Nat. Nanotechnol.* **2020**, 15, 47.
- [22] J. Grollier, D. Querlioz, M. D. Stiles, *Proc. IEEE* **2016**, 104, 2024.
- [23] J. Cai, B. Fang, L. Zhang, W. Lv, B. Zhang, T. Zhou, G. Finocchio, Z. Zeng, *Phys. Rev. Appl.* **2019**, 11, 034015.
- [24] A. Kurenkov, S. DuttaGupta, C. Zhang, S. Fukami, Y. Horio, H. Ohno, *Adv. Mater.* **2019**, 31, 1900636.
- [25] W. A. Borders, H. Akima, S. Fukami, S. Moriya, S. Kurihara, Y. Horio, S. Sato, H. Ohno, *Appl. Phys. Express* **2017**, 10, 013007.
- [26] C. H. Wan, M. E. Stebliy, X. Wang, G. Q. Yu, X. F. Han, A. G. Kolesnikov, M. A. Bazrov, M. E. Letushev, A. V. Ognev, A. S. Samardak, *Appl. Phys. Lett.* **2021**, 118, 032407.
- [27] Y. Oh, S. C. Baek, Y. M. Kim, H. Y. Lee, K. Lee, C. Yang, E. Park, K.-S. Lee, K. Kim, G. Go, J. Jeong, B. Min, H. Lee, K. Lee, B. Park, *Nat. Nanotechnol.* **2016**, 11, 878.
- [28] S. Fukami, C. Zhang, S. DuttaGupta, A. Kurenkov, H. Ohno, *Nat. Mater.* **2016**, 15, 535.
- [29] G. Yu, P. Upadhyaya, Y. Fan, J. G. Alzate, W. Jiang, K. L. Wong, S. Takei, S. A. Bender, L. Chang, Y. Jiang, M. Lang, J. Tang, Y. Wang, Y. Tserkovnyak, P. K. Amiri, K. L. Wang, *Nat. Nanotechnol.* **2014**, 9, 548.
- [30] S. Chen, J. Yu, Q. Xie, X. Zhang, W. Lin, L. Liu, J. Zhou, X. Shu, R. Guo, Z. Zhang, J. Chen, *ACS Appl. Mater. Interfaces* **2019**, 11, 30446.
- [31] L. You, O. Lee, D. Bhowmik, D. Labanowski, J. Hong, J. Bokor, S. Salahuddin, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, 112, 10310.
- [32] L. Liu, Q. Qin, W. Lin, C. Li, Q. Xie, S. He, X. Shu, C. Zhou, Z. Lim, J. Yu, W. Lu, M. Li, X. Yan, S. J. Pennycook, J. Chen, *Nat. Nanotechnol.* **2019**, 14, 939.
- [33] Y. Cao, Y. Sheng, K. W. Edmonds, Y. Ji, H. Zheng, K. Wang, *Adv. Mater.* **2020**, 32, 1907929.
- [34] M.-C. Chen, A. Sengupta, K. Roy, *IEEE Trans. Magn.* **2018**, 54, 8.
- [35] D. Pinna, F. Abreu Araujo, J.-V. Kim, V. Cros, D. Querlioz, P. Bessiere, J. Droulez, J. Grollier, *Phys. Rev. Appl.* **2018**, 9, 064018.
- [36] S. Li, W. Kang, Y. Huang, X. Zhang, Y. Zhou, W. Zhao, *Nanotechnology* **2017**, 28, 31LT01.
- [37] Y. Huang, W. Kang, X. Zhang, Y. Zhou, W. Zhao, *Nanotechnology* **2017**, 28, 08LT02.
- [38] L. Liu, C. Zhou, X. Shu, C. Li, T. Zhao, W. Lin, J. Deng, Q. Xie, S. Chen, J. Zhou, R. Guo, H. Wang, J. Yu, S. Shi, P. Yang, S. Pennycook, A. Manchon, J. Chen, *Nat. Nanotechnol.* **2021**, 16, 277.
- [39] O. J. Lee, L. Q. Liu, C. F. Pai, Y. Li, H. W. Tseng, P. G. Gowtham, J. P. Park, D. C. Ralph, R. A. Buhrman, *Phys. Rev. B* **2014**, 89, 024418.
- [40] G. Yu, P. Upadhyaya, K. L. Wong, W. Jiang, J. G. Alzate, J. Tang, P. K. Amiri, K. L. Wang, *Phys. Rev. B* **2014**, 89, 104421.
- [41] C.-F. Pai, M. Mann, A. J. Tan, G. S. D. Beach, *Phys. Rev. B* **2016**, 93, 144409.
- [42] J. Cao, Y. Chen, T. Jin, W. Gan, Y. Wang, Y. Zheng, H. Lv, S. Cardoso, D. Wei, W. S. Lew, *Sci. Rep.* **2018**, 8, 1355.