A Hopfield neural network in magnetic films with natural learning

Weichao Yu (余伟超),1 Jiang Xiao (萧江),2,3,4,* and Gerrit E. W. Bauer (包格瑞)5,1,6

1Institute for Materials Research, Tohoku University, Sendai 980-8577, Japan
2Department of Physics and State Key Laboratory of Surface Physics, Fudan University, Shanghai 200433, China
3Institute for Nanoelectronics Devices and Quantum Computing, Fudan University, Shanghai 200433, China
4Shanghai Research Center for Quantum Sciences, Shanghai 201315, China
5WPI-AIMR, Tohoku University, Sendai 980-8577, Japan
6Zernike Institute for Advanced Materials, Groningen University, Netherlands

Macroscopic spin ensembles possess brain-like features such as non-linearity, plasticity, stochasticity, self-oscillations, and memory effects, and therefore offer opportunities for neuromorphic computing by spintronics devices. Here we propose a physical realization of artificial neural networks based on magnetic textures, which can update their weights intrinsically via built-in physical feedback utilizing the plasticity and large number of degrees of freedom of the magnetic domain patterns and without resource-demanding external computations. We demonstrate the idea by simulating the operation of a 4-node Hopfield neural network for pattern recognition.

Introduction. Tremendous progress in the last decade has propelled neuromorphic computing to the forefront of information technology. However, brain-inspired algorithms are mainly emulated by conventional von Neumann architectures in which the computing and storage units are physically separated, thereby limiting the power of artificial intelligence algorithms. For example, the power consumption of the Alpha processor (∼1 MW) is 50,000 times higher than that of a human brain (∼20 W). [1, 2] A more sustainable route towards artificial intelligence is an architecture with hard-wired neuromorphic functions. Spintronic systems based on magnets share features of the brain, such as non-linearity, memory, self-oscillations, stochasticity, plasticity, high degrees of freedom, etc. [3] These advantages already led to alternative computing schemes, such as the stochastic [4, 5], in-memory logic [6], as well as neuromorphic [7] computing.

The Artificial Neural Network (ANN) is a widely used model for neuromorphic computing, with artificial neurons and synapses that emulate biological system [8]. Neurons are devices with output signals that spike when the integrated input reaches a certain threshold. While synapses connect the neurons with tunable weights. Both functionalities can be mimicked by spintronic devices. For example, spin-torque oscillators can serve as artificial neurons and recognize spoken digits and vowels [9, 10]. The nonlinear dynamics of skyrmion fabrics can pre-process information for reservoir computing [11–13].

Memristors have a resistance that depends on its history and are widely used as artificial synapses [14–17]. Spintronics offers memristor functionalities by reconfigurable magnetic configurations whose resistance depends on, for example, the positions of magnetic domain walls [18], the number of skyrmions [19–21], or the texture in antiferromagnet/ferromagnet bilayers [22], all of which can be controlled by applied fields or currents.

In most software and hardware realizations of an ANN, the weight-updating process of the synapses is based on external algorithms, such as the back-propagation method in which external computations allocate new weights based on the results of a previous cycle. Unfortunately, this is more expensive in terms of resource and energy consumption than the inferring process itself [23]. In this Letter we propose a platform for neuromorphic computation with superior performance that is based on the plasticity of magnetic textures. We show that an electrically conducting magnetic film may operate as a collection of artificial synapses with weights encoded by the conductances between external electrodes. The network training naturally updates the weights without requiring external computation. As a proof of principle we simulate the on-chip training and inferring (or read-out) of a 4-node Hopfield network implemented on a metallic magnetic thin film with maze spiral domains as shown in Fig. 1(g).

Modeling. The dynamics of the magnetization $\mathbf{M}(r, t)$ of a ferromagnetic film with saturation magnetization $M_s$ is governed by the Landau-Lifshitz-Gilbert (LLG) equation

$$\frac{\partial \mathbf{m}}{\partial t} = -\gamma \mathbf{m} \times \mathbf{H}_{\text{eff}} + \alpha \mathbf{m} \times \frac{\partial \mathbf{m}}{\partial t} + \mathbf{\tau}_d,$$

where $\mathbf{m} = \mathbf{M}/M_s$, $\gamma$ is the gyromagnetic ratio and $\alpha$ is the Gilbert damping constant. The effective magnetic field

$$\mathbf{H}_{\text{eff}} = A \nabla^2 \mathbf{m} + K \mathbf{m} \cdot \mathbf{z} - D \nabla \times \mathbf{m}$$

consists of the exchange interaction, the perpendicular easy-axis anisotropy along $\mathbf{z}$, and a bulk-type Dzyaloshinskii-Moriya interaction (DMI) [24–26], parameterized by $A$, $K$, and $D$, respectively. The current-induced spin-transfer torque [27–29] in Eq. (1)

$$\mathbf{\tau}_d = \frac{\mu_B P}{e M_s} \mathbf{j} \cdot \nabla \mathbf{m}$$

is proportional to the electric current density $\mathbf{j}$, where $\mu_B$ is the Bohr magneton, $P$ the (conductivity) spin polarization, and $-e$ the electron charge. The dipolar interaction is not important for the energetics of sub-micrometer scale textures and may be disregarded for the DMI-stabilized ones considered here.

The electric current density $\mathbf{j}$ is proportional to the local electric field $\mathbf{E}$:

$$\mathbf{j}(r) = \sum_{\mathbf{r}} |\mathbf{m}(\mathbf{r})| \cdot \mathbf{E}(r),$$

granted to personal use only. Further reproduction or distribution is only allowed with permission of the copyright holder.
Figure 1. (a) Snapshot of a maze domain structure at \( t_0 = 0 \) ns (color code for out-of-plane magnetization and arrows for in-plane magnetization) (b) Calculated conductance \( |G| \) in longitudinal (\( \hat{x} \)) and transverse (\( \hat{y} \)) directions. Snapshots of magnetization (c)(e) and current density (d)(f) distributions at \( t_1 = 10 \) ns and \( t_2 = 50 \) ns, respectively. An applied \( V_0 = 0.05 \) V drives a current along the \( x \) direction. (g) Left: Schematical Hopfield network formed by four neurons with values \( u_i \) and weights \( w_{ij} \). Right: Implementation of this network by a magnetic thin film with a texture that stores weights that can be trained by currents. The inputs are the voltages \( V_i \) at the electrodes and the outputs are the currents \( I_i \) into the film. The conductances \( G_{ij} \) are equivalent to the weights in the Hopfield network.

where \( \Sigma[m] \) is the \( 2 \times 2 \) conductivity matrix of a magnetic thin film with the anisotropic magnetoresistance (AMR) [30], i.e., a local resistivity depending on the angle \( \theta \) between the current flow and the local magnetization \( m(r) \) like \( \rho = \rho_\parallel \cos^2 \theta + \rho_\perp \sin^2 \theta \). Inverting this relation, the Cartesian elements \( \Sigma_{ij}[m] = \sigma_\perp + \sigma_\delta m_i m_j \) with \( \sigma_\perp = 1/\rho_\perp \), \( \sigma_\delta = 1/\rho_\parallel - 1/\rho_\perp \) and the AMR ratio \( a = (\rho_\parallel - \rho_\perp) / (\rho_\parallel + \rho_\perp) \) [11–13, 31]. A large enough current-induced magnetization rotation in Eq. (1) rotates the magnetization in turn modulates the current distribution by Eq. (4). We solve the spatiotemporal Eqs. (1, 4) self-consistently under the constraint \( \nabla \cdot j = 0 \) by the COMSOL Multiphysics [32] finite element code.

**Magnetic synapse.** In magnetic films with DMI \( D > 4\sqrt{AK}/\pi \) [26, 33, 34], a spiral maze domain texture emerges [35] as illustrated by Fig. 1(a) for a \( 400 \) nm \( \times \) \( 400 \) nm slab with thickness \( 1 \) \( \mu \)m. Parameters are typical for, e.g., Pt/CoFe/MgO, but a large AMR ratio \( a = 150\% \) as found in \( \text{Sr}_2\text{IrO}_4 \) [36], and \( D = 6.5 \times 10^{-3} \) A. In this system many energetically nearly degenerate textures span a huge configuration space that is accessible by small variations in temperature, field or voltage. Here we focus on the overdamped regime with \( \alpha = 0.3 \), which can be reached by rare earth doping [37]. The low-energy realizations of magnetic textures are then stable in the absence of applied torque and forces, while the effects of weak pinning may be disregarded (see Supplemental Materials (SM) [38]).

Under the action of the spin-transfer torque caused by an electric voltage \( V_0 \) applied across the film, the maze-domain texture evolves with time, simultaneously modulating the conductance via the AMR. Fig. 1(c-f) shows snapshots of the texture with \( V_0 = 0.05 \) V applied in the \( x \) direction, as in Fig. 1(a). The domains tend to align perpendicular to the current flow to minimize the spin-transfer torque, as observed in thin films of lanthanum strontium manganite [39], but the sample boundaries prevent perfect alignment. The reorientation of the texture induced by the current leads to a conductance increase in the longitudinal direction (\( \hat{x} \)) and a conductance decrease in the transverse direction (\( \hat{y} \)), see Fig. 1(b). The magnetic texture acts as memristors that saturates only when high voltages are applied for sufficiently long time.

A self-learning Hopfield network. The magnetic textures equipped with electrodes establish a Hopfield network [40, 41] of fully connected neurons that can recognize patterns. Fig. 1(g) sketches a Hopfield network with 4 neurons, where the synapses between different neurons are realized by the underlying magnetic film with textures. The neuron inputs are the voltages at the electrodes, which take binary values as \( \{V_i = \pm V_0\} \). According to Kirchhoff’s law \( I_i = \Sigma_j G_{ij}[m]V_j \), the current through the electrodes are governed by a (symmetric) conductance matrix \( G_{ij}[m] = G_{ij}^0 + G'_{ij}[m] \) that consists of a texture-independent \( G_0 \) and dependent con-
The voltage sign), there are 8 degenerate patterns for Fig. 2(c) according to center) and mirror symmetry (flipping and -

\[ G_{ij} = -\sum_{j \neq i} G_{ij} > 0 \quad \text{with} \quad G_{ij} < 0 \quad \text{for} \quad i \neq j, \quad (5) \]

where a current is positive when flowing into the electrodes. \( G_0 \) can be measured by saturating the magnetic film by a sufficiently strong magnetic field perpendicular to the film and satisfies the same constraints as \( \hat{G} \) in Eq. (5). The difference \( \hat{C}' = \hat{G} - \hat{G}_0 \) is still bound by \( G_{ij} = -\sum_{j \neq i} G_{ij} \) but the non-diagonal elements \( G_{ij}' \) can have either sign.

We may distinguish the current contributions from the texture-independent conductance \( \hat{G}_0 \) and that from the texture-dependent conductance \( \hat{G}' \): \( I_i = I_i^0 + I_i' \), where

\[ I_i' = \sum_j G_{ij}' V_j. \quad (6) \]

This completes the formulation of Hopfield network with \( V_j = \pm V_0 \) the neuron inputs, \( G_{ij} \) the tunable weights, and \( \text{sign}(I_i') \) the neuron outputs. Analogous to the bipolar Ising spin glass model [43], we define the functional energy

\[ E' = -\sum_i I_i' V_i = -\sum_{i,j} G_{ij}' V_i V_j = 2V_0^2 \sum_{V_i \neq V_j} G_{ij}'. \quad (7) \]

Physically, \(-E'\) is the power consumption of the magnetic thin film (with conductance matrix \( G_{ij} \)) relative to that of the texture-free thin film \( (G_{ij}^0) \): \(-E' = -E - (E_0)\) with \(-E_0 = \sum_i G_{ij}^0 V_i V_j\). With fixed \( G_{ij}^0 \), we know \(-E_0\) in advance for arbitrary inputs \( \{V_i\}\). The energy Eq. (7) therefore measures the additional Joule heating caused by the magnetic texture.

The film can be trained to memorize a pattern encoded by an array of binary values \( V \equiv \{V_i\} \) simply by applying the voltages \( \{V_i\} \) and let magnetization evolve. The texture adjusts itself to the current-induced spin-transfer torque such that the conductances \( (G_{ij}^0) \) and \( G_{ij}' \) between electrode-\( i \) and \( j \) increase when \( V_i \neq V_j \), thereby decreasing the objective function \( E' \) in Eq. (7). The conductances between electrodes with the same voltage cannot be directly trained, but they tend to decrease when other conductances grow.

Fig. 2 shows examples of training a four-neuron network to memorize three patterns, corresponding to the voltages on the four nodes as \( \{+ - + -\}, \{+ + - +\}, \{+ - + +\}, \) respectively. The top panels show the post-training texture and current density distribution. The lower panels of Fig. 2 show the energies Eq. (7) of the trained texture when fed by all possible inputs. The energy is minimized when the trial pattern (up to an global sign change) agree with the memorized pattern. For instance, Fig. 2(d) shows clear minima for the equivalent \( \{+ - + -\} \) and \( \{- + - +\} \) states. Due to the point (rotating the pattern according to center) and mirror symmetry (flipping the voltage sign), there are 8 degenerate patterns for Fig. 2(c) (2 for (a), 4 for (b)), representing 14 out of 16 possible states (that include the trivial \( \{+ + + +\} \) and \( \{- - - -\} \)) spanned by a 4-node Hopfield network.

The memorized pattern can be retrieved by standard inferencing algorithms, for example by feeding the neurons with a random initial pattern of binary voltages \( \{V_i(0)\} \) with small amplitudes that do not perturb the texture. The voltages can be then updated either asynchronously or synchronously [41] with \( V_i(t + 1) = \text{sign}(I_i'(t)) V_0 \). The self-consistent state with \( V_i = \text{sign}(I_i') V_0 \) corresponds to the minimum of the energy function Eq. (7) and the memorized pattern [38].

Natural learning. In other types of hardware-implemented synapses, such as cross-bar grid memristors [44, 45], the weight-updating requires learning algorithms (such as the back-propagation method) [46] that have to be executed externally. In contrast, the weight-updating in magnetic textures is natural and intrinsic. The positive feedback mechanism between the training current and the texture response does not require any external interference. The reinforcement of the interconnection of two “neurons” when activated by different voltages is analogous to Hebb’s learning rule in neuroscience [47] stating that simultaneous activation of neurons leads to increased synaptic strength.

By dividing the continuous training process into discrete temporal slices, i.e. regarding the voltages as a train of pulses, the conductance (weight) matrix evolves as

\[ G_{ij}^{n+1} = G_{ij}^n + \Delta_{ij}^n[\{V_i\}] \quad (8) \]

Here \( \Delta_{ij}^n \) updates at step-\( n \) the weight for the connection between node \( i \) and \( j \), depending mainly on the voltage difference \( V_i - V_j \). The conductance matrix will not reach overall saturation since the magnetic strips twist and move but cannot easily created or destroyed, especially in the presence of topological defects stabilized by the DMI. The strengthening of certain connections therefore weakens others. This competition of the weight modulations is another typical feature of organic synapses that is replicated by the magnetic textures [41].

Discussion. The functionality and efficiency of the proposed neural network relies on the magnitude of the AMR. The AMR ratio \( a \) is \( \sim 8\% \) in \( \text{Ni}_{80}\text{Fe}_{20} \) thin films [48], \(-10\%/6\%\) in single-crystalline \( \text{Co}_x\text{Fe}_{1-x} \) alloys depending on growth direction [49], and \( \sim 80\% \) in bilayered \( \text{La}_{1.2}\text{Sr}_{1.8}\text{Mn}_{2}\text{O}_7 \) single crystals [50]. In antiferromagnetic \( \text{Sr}_2\text{IrO}_4 \) it can reach \( a \sim -160\% \) [36]. A larger DMI than used here (\( >0.002 \text{ J/m}^2 \)) reduces the pitch of the spirals and improves the efficiency of the spin-transfer torque. The parameters and voltages are chosen optimistically to keep computation times manageable. The general idea works for less optimal materials, but at the costs of higher power dissipation and training times. We show in the SM that the planar and anomalous Hall effects [38, 51] may be disregarded because the in-plane and perpendicular magnetizations vanish on average in the absence of an external field [52].

The bulk-type DMI considered above generates Bloch type domain walls, that the current induced torque tends to realign such that the conductance increases. In materials with a dominant field-like spin-transfer torque [29] or with interfacial-type DMI [26, 53], a negative feedback reduces the conductance by applying currents (see SM [38]). Our network scheme can cooperate as well with such negative-feedback
by swapping the roles of voltage and current in the inferring process and using the resistance instead of conductance as weights in the objective function.

Our proof-of-principle device can only store a 4-pixels pattern. Larger pictures can be stored by increasing the number of nodes, and multiple patterns can be memorized when the energy function has multiple local minima [40]. Scaling up the network increases the storing capacity in the form of more complex and multiple input patterns. Performance can be optimized also by the node positions or employing three dimensional textures.

**Conclusion.** We proposed a neural network formed by electric contacts to conducting magnets with a complex magnetization texture. Because of the plasticity of the magnetic textures, the weights of the synapses can be automatically updated during the training process via a positive feedback mechanism between the current-induced spin-transfer torque and the electrical conductance. We numerically simulate training of and retrieval from a 4-node Hopfield network on a spiral magnetic maze stabilized by DMI. The learning is natural, based on physical laws without human intervention. The concept can be generalized to other materials with ”plasticity”, such as reconfigurable ferroelectrics with conducting domain walls [54, 55]. Our work paves the way to realize hardware-based neuromorphic computing with natural learning.

**Acknowledgements.** We are grateful to Shunsuke Fukami, Hangwen Guo, Zhe Yuan and Ke Xia for fruitful discussions. This work was supported by the JSPS Kakenhi (Grant Nos. 20K14369, 19H006450). J.X. was supported by National Science Foundation of China (Grant No. 11722430) and Shanghai Municipal Science and Technology Major Project (Grant No. 2019SHZDZX01). W.Y. acknowledges the support from the State Key Laboratory of Surface Physics.

---

* xiaojiang@fudan.edu.cn

[1] Jacques Mattheij, “Another Way Of Looking At Lee Sedol vs AlphaGo,” https://jacquesmattheij.com/another-way-of-looking-at-lee-sedol-vs-alphago/, online; accessed 17 March 2016.

[2] E. Goi, Q. Zhang, X. Chen, H. Luan, and M. Gu, PhotoniX 1, 3 (2020).

[3] J. Grollier, D. Querlioz, and M. D. Stiles, Proceedings of the IEEE 104, 2024 (2016).

[4] W. A. Borders, A. Z. Pervaiz, S. Fukami, K. Y. Camsari, H. Ohno, and S. Datta, Nature 573, 390 (2019), number: 7774 Publisher: Nature Publishing Group.

[5] M. W. Daniels, A. Madhavan, P. Talatchian, A. Mizrahi, and M. D. Stiles, Physical Review Applied 13, 034016 (2020), publisher: American Physical Society.

[6] W. Yu, J. Lan, and J. Xiao, Physical Review Applied 13, 024055 (2020).

[7] J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, Nature Electronics , 1 (2020), publisher: Nature Publishing Group.

[8] B. Kröse and P. Smagt, An introduction to Neural Networks
