Neural-Network Computation Using Spin-Wave-Coupled Spin-Torque Oscillators

Hiroko Arai[1,2,*] and Hiroshi Imamura[1]

[1]*Spintronics Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki 305-8568, Japan*

[2]*JST PRESTO, Kawaguchi, Saitama 332-0012, Japan*

A neural-network computation scheme is proposed based on a perceptron model having processing units that consist of spin-wave-coupled spin-torque oscillators. This is an oscillatory neural network, where the relative phase of the oscillators is controlled by tuning the Dzylaoshinshkii-Moriya interaction and applying an oscillating magnetic field. Each processing unit receives an input signal that is an oscillating magnetic field and transmits alternating current. The alternating current is a function of the relative phase and generates an oscillating magnetic (Oersted) field around the wire through which the current flows. The generated Oersted field then becomes an input signal to the next processing unit. By solving the Landau-Lifshitz-Gilbert equation, we obtain an activation function of the processing unit. Finally, an artificial neural network is constructed using the obtained activation function to recognize the handwritten digits in the Modified National Institute of Standards and Technology database.

I. INTRODUCTION

An artificial neural network (ANN) is a neuromorphic computational construct designed to mimic the neural network in the brain [1]. Recently, much attention has been focused on information processing using ANNs, such as for object detection and recognition [2,3], and in other fields of business application [4]. Because of the rapid progress in the machine power of computers, in terms of operation speed, memory capacity, etc., software applications using ANNs can be used in practical applications. However, the energy consumption of computers is much larger than that of the brain.

Software applications are typically developed on the central processing unit (CPU) or the graphics processing unit (GPU) of computer machines, which are so-called von Neumann architectures. In order to further improve the performance of the ANN, including its energy consumption, the building of a non-von Neumann architecture is intriguing and very attractive [5]. Neuromorphic architecture has adopted the TrueNorth chip, which is fabricated using complementary metal-oxide semiconductor (CMOS) technology [2]. Not only the conventional devices in CMOS technology [2,6] but also the newer functional devices are candidates for elements of the ANN, such as the atomic switch [7], the memristor [8–11], the coupled-phase oscillator [12,13], and spintronics devices [14–25].

Spintronics devices are easily fabricated on CMOS devices because the technology for manufacturing spintronics devices is similar to that for manufacturing CMOS devices, which is a great advantage in developing an ANN operating with a CPU. Also, they can typically be fabricated in nanosize scale, which is another advantage in reducing the spatial dimensions of the ANN system as a whole. Among various spintronics devices, the spin-torque oscillator (STO) is one of the suitable candidates for the ANN. This is because neurons in the brain show periodic spiking behavior and the STO has been found to essentially have the same oscillation property as a biological neuron [26]. The STO provides additional advantages. For example, the GHz-band precession of the STO provides high-speed operation. Also, the thermal tolerance and radiation resistance will be beneficial for easing the limits of the operational conditions in various environment. So far, various neuromorphic applications using the STO operating principle have been reported, such as associative memories [18–20], pattern matching [22–24], and spoken-digit recognition [25]. Most of these studies have been performed by focusing on the synchronization property, i.e., whether or not the STOs are locked. In other words, the artificial neuron is implicitly assumed to possess the activation function of a step function. However, artificial neurons in ANN software applications allow other activation functions, such as the sigmoid function and the rectified linear unit (ReLU) function [3]. Because the performance of the ANN can be improved by using these nonlinear functions, it is therefore interesting to construct

---

*arai-h@aist.go.jp

024040-1

an artificial neuron that incorporates a STO having the activation function other than the step function.

In this paper, a neural-network computation scheme using a spin-wave-coupled spin-torque oscillator is proposed. Especially, the basic concept and working principles are brought into focus. The activation function of the processing unit is obtained by solving the Landau-Lifshitz-Gilbert equation. Then, an ANN using the obtained activation function is constructed to recognize the handwritten digits in the MNIST database [27].
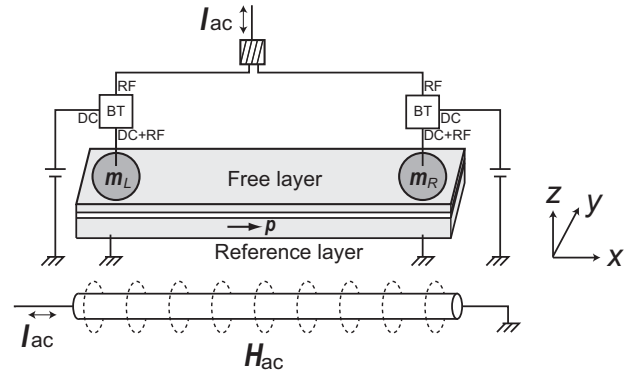

FIG. 1. A schematic illustration of the processing unit consisting of spin-wave-coupled STOs. The magnetization directions of the free layers of the left and the right STOs are represented by the unit vectors m_L and m_R, respectively. The direct current is applied to the STOs through each point contact. The oscillating magnetic field, H_{ac}, generated by the alternating current is also applied. The alternating current, I_{ac}, which is the sum of the ac from the bias tees (BTs) at each point contact, generates the oscillating magnetic (Oersted) field, H_{Oe}, around the wire through which the current flows, as indicated by the dotted circles. The generated Oersted field becomes an input signal to the next processing unit. The Cartesian coordinates are also shown. The direction of the magnetization of the reference layer, p, is fixed to the positive x direction.

II. RESULTS AND DISCUSSION

A. The single unit

Figure 1 shows a schematic illustration of the processing unit consisting of spin-wave-coupled STOs, which is a laterally long magnetic resistant device with double point contacts. The magnetization directions of the free layers of the left and the right STOs are represented by the unit vectors m_L and m_R, respectively. The direct current is applied to the STOs through each point contact to induce the oscillation dynamics of magnetizations underneath the point contacts, the dynamics of which are coupled with each other by spin waves in the free layer. The coupling is assumed to comprise the Heisenberg interaction (HI) and the Dzyaloshinskii-Moriya interaction (DMI). On the one hand, the HI prefers the collinear magnetic structure and, on the other, the DMI prefers the noncollinear magnetic structure. By tuning the ratio of the strength of the DMI to that of the HI, one can control the relative phase of the oscillations of m_L and m_R.

By summing up the alternating current from the bias tees (BTs) at each point contact using a combiner, we obtain the total alternating current, I_{ac}, which is a function of the relative phase. This comes from the fact that the oscillating component of the resistance of each STO is inversely proportional to the inner product $m_{L(R)} \cdot p$, where p is the direction of magnetization in the reference layer. The amplitude of I_{ac} for in-phase oscillation is larger than that for out-of-phase oscillation. The relative phase can also be controlled by applying an oscillating magnetic field, which exerts a torque on m_L and m_R, resulting in the in-phase oscillation. If the spatial separation of the STOs and the strength of the DMI are designed such that the out-of-phase oscillation is induced at $H_{ac} = 0$, the output signal, I_{ac}, increases with an increase in the input signal, H_{ac}. The output signal of the processing unit is the oscillating magnetic (Oersted) field generated around the wire through which I_{ac} flows. The function representing the relation between the input and output signals of the processing unit is called the activation function and it plays an important role in ANNs.

The activation function of spin-wave-coupled STOs is obtained by solving the Landau-Lifshitz-Gilbert equation

having the Sloczewski spin-torque term under an oscillating magnetic field:

$$\frac{d\boldsymbol{m}}{dt} = -\gamma(\boldsymbol{m} \times \boldsymbol{H}_{\text{eff}}) + v\boldsymbol{m} \times (\boldsymbol{m} \times \boldsymbol{p}) + \alpha\left(\boldsymbol{m} \times \frac{d\boldsymbol{m}}{dt}\right),$$
(1)

where γ is the gyromagnetic ratio and α is the Gilbert damping constant. Here, $\boldsymbol{m} = \boldsymbol{m}(\boldsymbol{r}, t)$ represents the magnetization unit vector at position $\boldsymbol{r}$ and at time t. The arguments $\boldsymbol{r}$ and t are omitted for convenience. The effective field $\boldsymbol{H}_{\text{eff}}$ is given by the sum of the HI field, the DMI field, and the oscillating external magnetic field. The HI field is given by

$$\boldsymbol{H}_{\text{HI}} = \frac{2A}{\mu_0 M_s}\nabla^2\boldsymbol{m},$$
(2)

where A is the stiffness constant, μ_0 is the magnetic permeability of a vacuum, and M_s is the saturation magnetization. The DMI field is given by

$$\boldsymbol{H}_{\text{DMI}} = \frac{2D}{\mu_0 M_s}\left(\frac{\partial m_z}{\partial x}, \frac{\partial m_z}{\partial y}, -\frac{\partial m_x}{\partial x} - \frac{\partial m_y}{\partial y}\right),$$
(3)

where D is the DMI constant, which ranges from 0.05 to 1.0 mJ/m^2, depending on the materials [28]. Throughout this paper, we assume that the free layer is designed in such a way that the magnetic anisotropy field cancels with the

demagnetization field. The oscillating external magnetic field is given by

$$\boldsymbol{H}_{\text{ac}} = H_{\text{ac}} \sin(2\pi f t) \boldsymbol{e}_y, \tag{4}$$

where $\boldsymbol{e}_y$ is the unit vector pointing in the positive y direction and f is the oscillation frequency. The second term on the right-hand side in Eq. (1) represents the spin-torque term, where $\boldsymbol{p}$ is fixed to the positive x direction, $\boldsymbol{p} = (1, 0, 0)$. The coefficient v is defined as follows:

$$v = \frac{g\mu_B JP}{2|e|M_s d(1 + P^2 \boldsymbol{m} \cdot \boldsymbol{p})}, \tag{5}$$

where g is the Landé g factor of the electron spin, μ_B is the Bohr magneton, J is current density through a region of the point contact, P is the spin polarization, e is the electron charge, and d is the thickness of the free layer.

The shape of the free layer is assumed as 2 nm wide, 40 nm long, and 1 nm thick. The free layer is modeled by a one-dimensional magnetization chain and is divided into cells with a size of $2 \times 2 \times 1$ (nm^3). The simulations are performed by using the MUMAX^3 code [29]. The STOs are represented by the cells at the left and right ends, $\boldsymbol{m}_L$ and $\boldsymbol{m}_R$, where the direct bias current is applied. The positive current is defined as the electrons flowing from the free layer to the reference layer.

The material parameters for the free layer are assumed as follows: $M_s = 8 \times 10^5$ (A/m), $P = 0.5$, $\alpha = 0.01$, $A = 7 \times 10^{-12}$ (J/m), and $D = 8 \times 10^{-4}$ (J/m^2). The density of the direct bias current is $J = 6 \times 10^{11}$ (A/m^2), and the initial state is $\boldsymbol{m} = (0, 0, 1)$ for all calculations.

B. The activation function

The processing unit is carefully designed to show an antiphase oscillation at $H_{\text{ac}} = 0$, as mentioned earlier. The specific result is shown in Fig. 2(a), where the blue and red curves represent the x component of $\boldsymbol{m}_L$ and $\boldsymbol{m}_R$ as a function of time, respectively. The application of an oscillating magnetic field drastically changes the relative phase of $\boldsymbol{m}_L$ and $\boldsymbol{m}_R$. Figure 2(b) shows the results under the oscillating magnetic field, $H_{\text{ac}} = 300$ (Oe) with $f = 4$ (GHz), which is the oscillation frequency at $H_{\text{ac}} = 0$. The initial antiphase oscillation shown in Fig. 2(c) becomes nearly inphase after a few nanoseconds and the in-phase oscillation is stabilized after just 25 ns, as shown in Fig. 2(d).

The oscillating and nonoscillating components of the current from the STOs are first separated by the bias tee. Then the oscillating components are summed up by a combiner (hatched box) to produce I_{ac}, as shown in Fig. 1. The alternating current I_{ac}, and therefore the magnetic field H_{ac} generated by I_{ac}, are proportional to the sum of the oscillating components of $\boldsymbol{m}_L$ and $\boldsymbol{m}_R$, which is denoted by m_x^{ac} and is shown in Fig. 2(e).
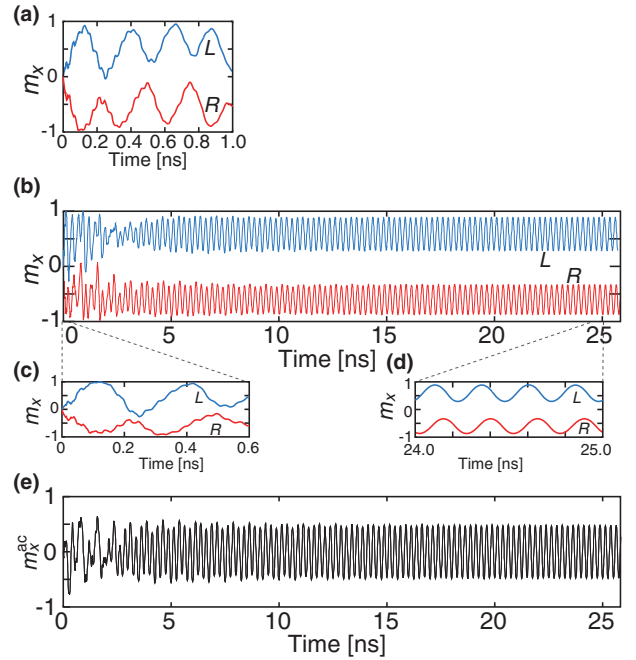
FIG. 2. (a) The time evolution of the x component of $\boldsymbol{m}_L$ and $\boldsymbol{m}_R$ at $H_{\text{ac}} = 0$ is plotted by the blue and red curves, respectively. (b) The same plot at $H_{\text{ac}} = 300$ (Oe). (c) An enlarged view of (b) for $t = 0$–0.6 (nsec). (d) An enlarged view of (b) for $t = 24$–25 (nsec). (e) The time evolution of m_x^{ac} for $H_{\text{ac}} = 300$ (Oe).

We introduce the activation function F, defined as

$$H_{\text{ac}}^{(\text{out})} = cF[H_{\text{ac}}^{(\text{in})}], \tag{6}$$

where $H_{\text{ac}}^{(\text{in})}$ is the external oscillating magnetic field applied to the processing unit and $H_{\text{ac}}^{(\text{out})}$ is the oscillating magnetic field generated by the processing unit. The coefficient c is a constant, the value of which is determined so that the activation function ranges from 0 to 1. The F value at each $H_{\text{ac}}^{(\text{in})}$ is obtained by analyzing m_x^{ac} as shown in Fig. 3, where the amplitude of m_x^{ac} is averaged for $t = 40$–50 (nsec). It increases with an increase in $H_{\text{ac}}^{(\text{in})}$ below 600 Oe because the precession cone angle increases with an increase in $H_{\text{ac}}^{(\text{in})}$. It shows a hump at around 300 Oe because the mode of the spin wave changes. It also shows a jump at $H_{\text{ac}}^{(\text{in})} = 600$ (Oe), above which $\boldsymbol{m}_R$ precesses with $m_z < 0$.

C. The artificial neural network

Since the activation function obtained, shown in Fig. 3, indicates a shifted sigmoidlike function, one can expect that the spin-wave-coupled STOs work as nodes of an ANN. We construct an ANN using the obtained activation function to recognize the handwritten digits in the MNIST database [27]. As shown in Fig. 4(a), the ANN consists of three layers: the input layer, the hidden layer, and the output layer. The input layer has 784 nodes, receiving input
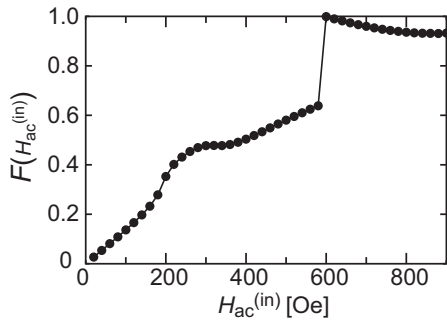
FIG. 3. The value of the activation function F obtained at each H_{ac}.

signals from normalized grayscaled pictures with 28×28 pixels. The hidden layer has 50 nodes. The output layer has ten nodes, corresponding to the digits from 0 to 9.

We briefly note the hardware implementation of the weights. As mentioned above, the input and output signals are the oscillating magnetic field, so that the weights should modify the amplitude and the phase if the weight is negative. These operations can be performed by using a device such as an analog multiplier. In addition, the product-sum operations between the output signals from nodes and weights are performed as a superposition of the waves. However, in this study, we simply perform the product-sum operations by summing the values as mentioned below.

Introducing the variable $x_{orig}^{(i)}$, ranging from 0 to 1, to represent the grayscale of the ith pixel, the input signal of the ith processing unit is set as $H_{ac,i}^{(in)} = 900 x_{orig}^{(i)}$. From Eq. (6), the input-output relation of the ith processing unit is given by

$$H_{ac,i}^{(out)} = cF[H_{ac,i}^{(in)}], \tag{7}$$

where the coefficient c is set as 90 Oe. The output signal from the STO in the input layer is multiplied by the value of weight W_1. The input signal to the node j in the hidden layer is given by

$$H_{ac,j}^{(in)} = \sum_{i=1}^{784} H_{ac,i}^{(out)} W_{1,ij}, \tag{8}$$

where $W_{1,ij}$ is the weight.

The output signal from the jth node in the hidden layer is given by

$$H_{ac,j}^{(out)} = cF[H_{ac,j}^{(in)}]. \tag{9}$$

Likewise, the input signal to the kth node in the output layer is given by

$$H_{ac,k}^{(in)} = \sum_{j=1}^{50} H_{ac,j}^{(out)} W_{2,jk}, \tag{10}$$

where W_2 is the weight between the hidden and output layers. Finally, the output signal from the kth node in the output layer is obtained as

$$H_{ac,k}^{(out)} = cF[H_{ac,k}^{(in)}]. \tag{11}$$

The node with the largest value indicates the digit predicted by the ANN. A learning process is required to update the value of the weights to obtain an accurate prediction from the ANN and this is performed by the back propagation method. For simplicity, in back propagation, we replace the activation function with the analytic sigmoid function as $F(x) = 1/\{1 + \exp[-a(x - x_0)]\}$, with a gain $a = 0.02$ and a shift $x_0 = 600$, which has a similar shape to that shown in Fig. 3. The cross-entropy error function is used for estimation of the error between the prediction by the ANN and the correct answer given by the MNIST database. Sixty thousand images are used for training (learning) and a further 10 000 images are used for testing. Figure 4(b) shows the accuracy of the recognition as a function of epochs. The conventional ANN consists of nodes with an activation function of the typical sigmoid function given by $F(x) = 1/\{1 + \exp(x)\}$. The conventional ANN produces 95% accuracy on average from the same structure as shown in Fig. 4(a). The accuracy of the STO-based ANN is 75%, which is less than that of the conventional ANN. We note that the lower accuracy results from the shape of the activation function. One may consider that the reason for the lower accuracy is disharmony of the shape between the activation functions for
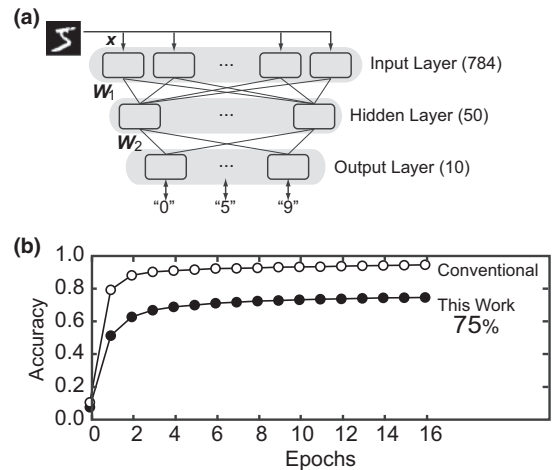

FIG. 4. (a) A schematic illustration of the ANN that we construct to recognize the handwriting digit in the MNIST database. Figures in brackets indicate the numbers of processing units. The output signal from each layer is multiplied by the weight, $W_{1 or 2}$, and then the total amount of the signals is fed to the next layer. (b) The transition in accuracy for the recognition obtained from a conventional neural network (open circles) and by this work (filled circles).

the forward and backward calculations during the training process. However, even when the shifted sigmoid function is used for both the forward and backward processes, the recognition accuracy remains around 75%. In order to increase the accuracy, one should rather tune the hyperparameters, such as the number of nodes in the hidden layer, the number of hidden layers, and the learning rate.

III. CONCLUSIONS

In summary, a neural-network computation is proposed based on a perceptron model having processing units that consist of spin-wave-coupled spin-torque oscillators. The activation function of the processing unit is obtained by solving the Landau-Lifshitz-Gilbert equation, incorporating the Slonczewski term. The artificial neural network is constructed using the activation function obtained from the study. The recognition of the handwritten digits in the MNIST database is demonstrated, with an accuracy of 75%.

ACKNOWLEDGMENTS

[1] D. Hassabis, D. Kumaran, C. Summerfield, and M. Botvinick, Neuroscience-inspired artificial intelligence, Neuron **95**, 245 (2017).

[2] P. A. Merolla, J. V. Arthur, R. Alvarez-Icaza, A. S. Cassidy, J. Sawada, F. Akopyan, B. L. Jackson, N. Imam, C. Guo, Y. Nakamura, B. Brezzo, I. Vo, S. K. Esser, R. Appuswamy, B. Taba, A. Amir, M. D. Flickner, W. P. Risk, R. Manohar, and D. S. Modha, A million spiking-neuron integrated circuit with a scalable communication network and interface, Science **345**, 668 (2014).

[3] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, Nature **521**, 436 (2015).

[4] M. Tkac and R. Verner, Artificial neural networks in business: Two decades of research, Appl. Soft Comp. **38**, 788 (2016).

[5] C. D. Schuman, T. E. Potok, R. M. Patton, J. D. Birdwell, M. E. Dean, G. S. Rose, and J. S. Plank, A survey of neuromorphic computing and neural networks in hardware, arXiv:1705.06963.

[6] P. Maffezzoni, B. Bahr, Z. Zhang, and L. Daniel, Oscillator array models for associative memory and pattern recognition, IEEE Trans. Circ. Sys. I: Regular Papers **62**, 1591 (2015).

[7] T. Hasegawa, K. Terabe, T. Tsuruoka, and M. Aono, Atomic switch: Atom/ion movement controlled devices for beyond von-Neumann computers, Adv. Mat. **24**, 252 (2012).

[8] S.-H. Jo, T. Chang, I. Ebong, B. B. Bhadviya, P. Mazumder, and W. Lu, Nanoscale memristor device as synapse in neuromorphic systems, Nano Lett. **10**, 1297 (2010).

[9] T. Chang, S.-H. Jo, and W. Lu, Short-term memory to long-term memory transition in a nanoscale memristor, ACS Nano **5**, 7669 (2011).

[10] A. Chanthbouala, V. Garcia, R. O. Cherifi, K. Bouzehouane, S. Fusil, X. Moya, S. Xavier, H. Yamada, C. Deranlot, N. D. Mathur, M. Bibes, A. Barthélémy, and J. Grollier, A ferroelectric memristor, Nat. Mat. **11**, 860 (2012).

[11] M. Prezioso, F. Merrikh-Bayat, B. D. Hoskins, G. C. Adam, K. K. Likharev, and D. B. Strukov, Training and operation of an integrated neuromorphic network based on metal-oxide memristors, Nature **521**, 61 (2015).

[12] P. Kaluza, Computation with phase oscillators: An oscillatory perceptron model, Neurocomputing **118**, 127 (2013).

[13] M. J. Cotter, Y. Fang, S. P. Levitan, D. M. Chiarulli, and V. Narayanan, in 2014 IEEE Computer Society Annual Symposium on VLSI (IEEE, 2014), p. 130.

[14] P. Krzysteczko, J. Münchenberger, M. Schäfers, G. Reiss, and A. Thomas, The memristive magnetic tunnel junction as a nanoscopic synapse-neuron system, Adv. Mat. **24**, 762 (2012).

[15] A. Sengupta, S. H. Choday, Y. Kim, and K. Roy, Spin orbit torque based electronic neuron, Appl. Phys. Lett. **106**, 143701 (2015).

[16] 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**, 29545 (2016).

[17] 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).

[18] T. Shibata, R. Zhang, S. Levitan, D. Nikonov, and G. Bourianoff, in 13th International Workshop on Cellular Nanoscale Networks and their Applications (IEEE, 2012), p. 1.

[19] S. P. Levitan, Y. Fang, D. H. Dash, T. Shibata, D. E. Nikonov, and G. I. Bourianoff, in 13th International Workshop on Cellular Nanoscale Networks and their Applications (IEEE, 2012), p. 1.

[20] D. Fan, S. Maji, K. Yogendra, M. Sharad, and K. Roy, Injection-locked spin Hall-induced coupled-oscillators for energy efficient associative computing, IEEE Trans. Nanotech. **14**, 1083 (2015).

[21] W. A. Borders, H. Akima, S. Fukami, S. Moriya, S. Kurihara, Y. Horio, S. Sato, and H. Ohno, Analogue spin-orbit torque device for artificial-neural-network-based associative memory operation, Appl. Phys. Express **10**, 013007 (2016).

[22] K. Yogendra, D. Fan, and K. Roy, Coupled spin torque nano oscillators for low power neural computation, IEEE Trans. Magn. **51**, 4003909 (2015).

[23] D. Nikonov, G. Csaba, W. Porod, T. Shibata, D. Voils, D. Hammerstrom, I. A. Young, and G. I. Bourianoff, Coupled-oscillator associative memory array operation for pattern recognition, IEEE J. Explor. Solid-State Comp. Dev. Circ. **1**, 85 (2015).

[24] K. Kudo and T. Morie, Self-feedback electrically coupled spin-Hall oscillator array for pattern-matching operation, Appl. Phys. Express **10**, 043001 (2017).

[25] 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).

[26] H. Arai and H. Imamura, Stochastic phase synchronization of perpendicularly magnetized spin-torque oscillators with the second-order uniaxial anisotropy, IEEE Trans. Magn. **53**, 6100705 (2017).

[27] Y. Lecun, C. Cortes, and C. J. C. Burges, The MNIST database of handwritten digits, http://yann.lecun.com/exdb/mnist/.

[28] K. Nawaoka, S. Miwa, Y. Shiota, N. Mizuochi, and Y. Suzuki, Voltage induction of interfacial Dzyaloshinskii–Moriya interaction in Au/Fe/MgO artificial multilayer, Appl. Phys. Express **8**, 063004 (2015).

[29] A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sanchez, and B. V. Waeyenberge, The design and verification of MUMAX^3, AIP Adv. **4**, 107133 (2014).