Dynamical Neural Network Based on Spin Transfer Nano-Oscillators

Davi R. Rodrigues ⓘ, Eleonora Raimondo ⓘ, *Graduate Student Member, IEEE*, Vito Puliafito ⓘ, *Member, IEEE*, Rayan Moukhadder ⓘ, Bruno Azzerboni ⓘ, *Senior Member, IEEE*, Abbass Hamadeh ⓘ, Philipp Pirro ⓘ, *Member, IEEE*, Mario Carpentieri ⓘ, *Senior Member, IEEE*, and Giovanni Finocchio ⓘ, *Senior Member, IEEE*

Abstract—Spintronic technology promises to significantly increase the efficiency and scalability of neural networks by employing optimized task-oriented device components that exhibit intrinsic nonlinearity, temporal nonlocality, scalability, and electrical tunability. In particular, the functional response of spin-transfer torque oscillators can be designed to naturally emulate the building blocks of neural networks, such as short-term memory, hierarchy, and nonlinearity. We propose spin-transfer nano-oscillators as a dynamic neuron that can be used in a neural network coupled with a fully connected layer to perform classification tasks. In this concept, successive nodes of the neural network correspond to successive time steps, so that the nonlinearity and memory of the system can be naturally exploited. The tunability of the device allows to project initial configurations in well-defined regions of the phase space where classification is easily performed. Furthermore, training is performed using optimal control theory. We emphasize that the devices benefit from more realistic models compared to simpler analytical models and is robust against device-to-device variations. We tested the performance of the network on two types of datasets and obtained 99% accuracy. Although these systems are computationally expensive, their hardware implementation is simple and inexpensive.

Index Terms—Dynamical neurons, spin-transfer-nano oscillators, spintronics, time nonlocality.

Davi R. Rodrigues, Vito Puliafito, Rayan Moukhadder, and Mario Carpentieri are with the Department of Electrical and Information Engineering, Politecnico di Bari, 70126 Bari, Italy (e-mail: davi.rodrigues@poliba.it; vito.puliafito@poliba.it; rayanroro321997@gmail.com; mario.carpentieri@poliba.it).

Eleonora Raimondo and Giovanni Finocchio are with the Department of Mathematical and Computer Sciences, Physical Sciences and Earth Sciences, University of Messina, I-98166 Messina, Italy (e-mail: eleonora.raimondo@unime.it; gfinocchio@unime.it).

Bruno Azzerboni is with the Department of Engineering, University of Messina, I-98166 Messina, Italy (e-mail: azzerbonib@unime.it).

Abbass Hamadeh and Philipp Pirro are with the Technical University of Kaiserslautern and Landesforschungszentrum OPTIMAS, 67663 Kaiserslautern, Germany (e-mail: hamadeh@rhrk.uni-kl.de; ppirro@rhrk.uni-kl.de).

Digital Object Identifier 10.1109/TNANO.2023.3330535

I. INTRODUCTION

THE ubiquitous use of neural networks on modern technology demands the development of faster, scalable and more efficient means to realize them. A proposed solution is the use of optimized task-specific devices [1], [2]. Moreover, it has been shown that exploiting the intrinsic functional response of the devices, which can be tuned to model the required behaviors of neural networks, such as non-linearity and short-term memory, allows for extremely scalable systems [3], [4]. The analog implementation of neural networks is a growing field with a growing number of proposals [4], [5], [6], [7], [8], [9], [10], [11]. A challenge in the field is finding systems that are scalable, can emulate the properties of a neural network and allow for an inexpensive training of the network [4], [7], [9].

The inherent non-volatility, non-linear behavior, CMOS-compatibility and scalability of spintronic devices put them on the forefront of analog neural networks [12], [13], [14]. For this reason, recently the field of Neuromorphic Spintronics has been set [14], [15]. While the main efforts have been devoted to developing reservoirs for reservoir computing, the idea to use spintronic devices as dynamical neurons, where the control parameters (current or field) are time domain functions, has not been explored.

Here we propose a dynamical neural network where the information is encoded in the power emitted by Spin Transfer Nano-Oscillators (STNOs) [16], [17]. STNOs are nano-sized, low power and highly tunable devices with well-stablished fabrication processes and well-characterized behavior, thus they provide a reliable platform for the realization of analog neural networks [10], [18], [19], [20], [21], [22], [23]. On top of them, those can be realized with magnetic tunnel junctions (MTJs), which are the basic element of the spin-transfer-torque MRAM and are already commercially available [24], [25], [26], [27]. This application makes the research on improving MTJs very active with the development of better deposition processes and nanofabrication techniques which can be also used to make the future generation of STNO with less device-to-device variation. STNOs have been already used for magnetic sensors [21], [28], [29], magnetic field-to-digital converters [30], Ising machines [31], [32], reservoir computing [10], [22], and other neuromorphic computing paradigms [33], [34], [35].

Here, we consider the power emitted by an STNO corresponds to a nonmonotonic function of the injected current which is
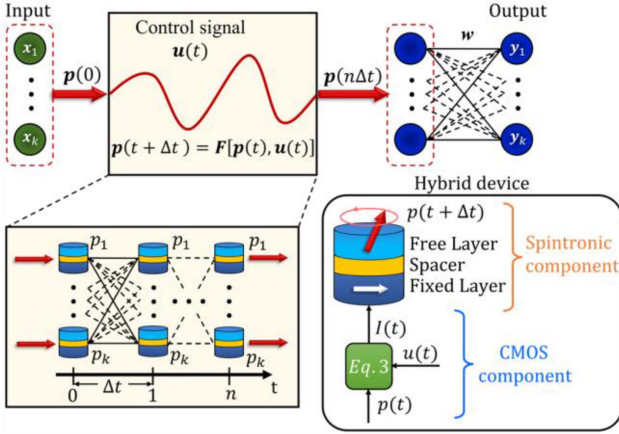
Fig. 1. Sketch of the structure of the dynamical neural network. The dataset is encoded as input power $p(0)$ of the dynamic system. The state at the end time $T_{end} = n\Delta t$, $p(n\Delta t)$, is the input of a fully connected layer for classification tasks. The control signal $u(t)$ and the weights w are the trainable parameters. (Bottom) Schematic of a STNO-based virtual neural network. (Bottom-right) Basic hybrid unit that makes up the Neural Network: a CMOS component read the power and obtain and inject the calculated current back into the STNO. The resistance of the STNO depends on the relative orientation between the magnetization of the free and fixed layer. In state-of-the art STNOs, the spacer is a thin (< 1.2 nm) MgO tunnel barrier.

identified directly from the experimental data of [18] and [36]. We wish to stress that qualitatuive similar data have been measured in a range of different oscillator configurations. We propose a neural network architecture where each node corresponds to a hybrid device that combines CMOS circuitry for electrical reading, current calculation, and injection, along with an STNO serving as the analog component responsible for generating the non-linear and time nonlocal functional response. This unique hybrid configuration capitalizes on the advantages of CMOS technology, enabling precise electrical operations, while harnessing the STNO's distinctive capabilities to deliver the desired non-linearity and inherent memory. In the proposed dynamical neural network, the STNO's output power evolves dynamically in response to the calculated input at each time step, which corresponds to a linear combination of the power output from the previous step. This dynamic evolution allows information to propagate through the network, utilizing the stationary power output of the STNOs at each time step as a means of conveying information. For an overview of the proposed dynamical neural network, see Fig. 1. The training of the network relies on the physical properties of the network and modifies the input matrix at each instance. This can be done by functional optimal control theory [9] coupled with the backpropagation scheme.

We propose this dynamical neural network for simple classification tasks to show the feasibility of the approach driving for a future experimental demonstration. The initial data set is projected in a phase space that evolves in time such that at a final step each class can be easily recognizable by a fully connected layer. We highlight that this concept resembles the concept of physical reservoir computing [10], [22], [37]. However, there are key distinctions: (i) In our approach, different nodes are represented by successive states of the device, enabling us to fully exploit the system's nonlinearity and memory. (ii) The dynamical system itself is tunable, with control parameters

that vary over time. (iii) A time-evolving control parameter is utilized to modulate the functional response, optimizing task performance. (iv) Training can be accomplished using optimal control theory. These characteristics give rise to a dynamical neuron, which incorporates an activation function exhibiting nonlocality in time. By time nonlocality, we refer to the intrinsic memory of the system, wherein the properties of the initial state strongly constrain the region of phase space for the final state. [38]. Dynamical neurons offer significant advantages over existing reservoir computer proposals, providing greater versatility and scalability without significantly increasing computational costs. Notably, the current tunability [3], [17], [22], [39], [40] of STNOs supports their implementation as dynamical neurons.

II. Dynamical Neural Networks Based on STNOs

A. Functional Response of STNOs

While we consider experimental data of STNOs published in [18] for the calculations performed here, where the device consists of two ferromagnetic layers of Permalloy (Py) separated by an nonmagnetic layer of Copper (Cu), see Fig. 2(a) [18], the results presented are fully generalizable to any other STNO with a monotonic relationship describing power vs current as we discuss ahead in the text. For the device in [18], the lower, thicker, Py layer has a fixed magnetization direction, while in the upper Py layer the magnetization is able to move in the presence of external torques. The spins of an injected electrical current get polarized by the fixed layer and interact with the magnetization of the free layer through the spin transfer torque [41], [42]. The unitary magnetization of the free layer, $\mathbf{m}$, in the presence of the spin torque behaves according to the LLG equation [16], [19], [43]

$$\frac{d\mathbf{m}}{dt} = \gamma\mathbf{m} \times \mathbf{H}_{eff} + \mathbf{T}_G + \mathbf{T}_S \qquad (1)$$

where γ is the gyromagnetic ration, $\mathbf{H}_{eff}$ is the effective magnetic field, $\mathbf{T}_G$ is the dissipative torque, and

$$\mathbf{T}_S = \frac{\varepsilon g\mu_B I}{2eM_S LS}\mathbf{m} \times (\mathbf{m} \times \hat{\mathbf{p}}), \qquad (2)$$

is the Slonczewski-Berger torque, where ε is the dimensionless spin-polarization efficiency, g is the Landé factor, μ_B is the Bohr magneton, I is the applied current, e is the electron charge, M_S is the saturation magnetization of the free layer, L is the thickness of the free layer, S is the area of the magnetic free layer, and $\hat{\mathbf{p}}$ is the direction of the magnetization in the fixed layer. If the electrical current I is above a certain threshold I_c it induces an auto-oscillation of the free layer magnetization [19], [20]. In the auto-oscillation regime, the magnetization rotates with a fixed component along a direction defined by the effective magnetic field and the vector $\hat{\mathbf{p}}$ with an emitted power proportional to this component. This behavior has been experimentally observed [18] and analytically described [19]. According to the analytical model developed on [19] and based on (1) and (2), the stationary power p of an STNO with an injected current I is given by

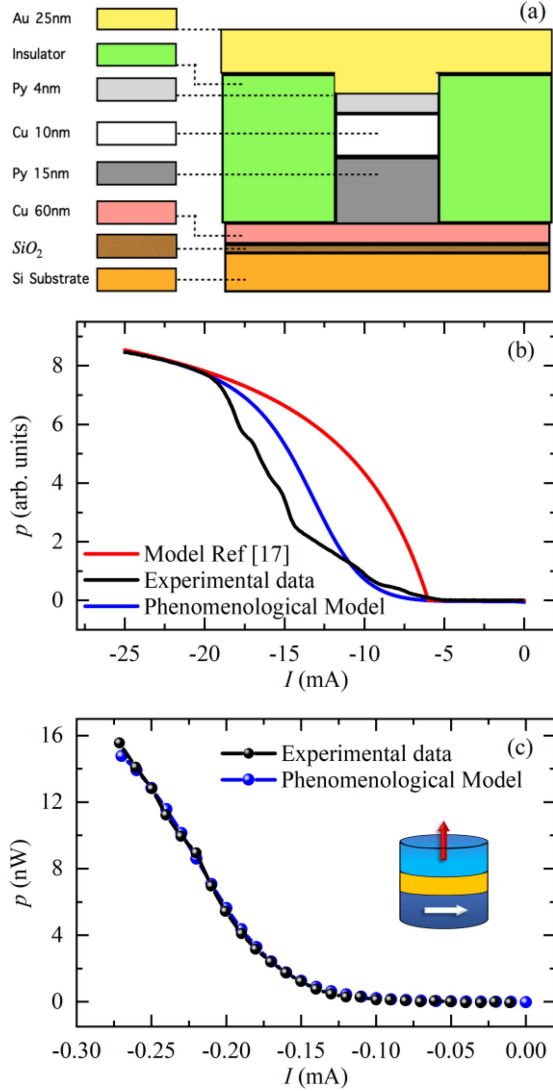$$p(I) = \frac{(I - I_c)}{(I + QI_c)}, \qquad (3)$$

Fig. 2. (a) The STNO structure of experiment performed in [18]. (b) A comparison of the power vs current relationship, experimental data corresponding to the data in [18] measured for an applied perpendicular magnetic field of 10kOe (black line); phenomenological models described by (4) (blue line) and the model described by the (3) and developed in [17] (red line), both with a critical current of −6 mA and Q equal to 0.1. (c) A comparison of the experimentally obtained power vs current from [36] (black line) and the fitted model from (4) (blue line). The critical current is −10 μA and Q equal to 0.3. The inset shows the STNO configuration studied in [36].

where Q is a measure of the non-linearity of the system. Real devices, however, present an exponential tail near the critical current that is not captured by (3) [18], see Fig. 2(b). As seen in experiments, the analytical model captures well the power curve away from the critical current, but fails to predict the behavior near I_c. The observed tail in the vicinity of I_c is due to specific experimental conditions and material properties of the STNO. Therefore, we propose a phenomenological model given by

$$p(I) = \frac{(1 + \tanh(aI - b))}{2} \frac{(I - I_c)}{(I + QI_c)}, \qquad (4)$$

which is based on the model of [19] with a prefactor function where a and b are fitting parameters identified directly by the experimental data and that includes the behavior near the critical

current. The prefactor function can have different expressions according to the power vs current characteristic of the STNO used as dynamical neuron and includes a higher non-linearity which better describes the proposed device behavior. We have observed that this prefactors works for many experimental data on STNO published in literature [35], [44], [45], [46].

The power emitted by the STNO can be obtained by analyzing the component of the magnetization of the free layer in the direction of the magnetization of the fixed layer, see next section for more details. The power output can be measured electrically [21], [28], [47].

B. Dynamical Neuron

The information in the dynamical neural network is encoded in the stationary power output of the k STNOs at each instance in time, $\mathbf{p}(t) = \{p_1(t), \ldots, p_k(t)\}$, which evolve iteratively over time according to

$$p_i(t + \Delta t) = p(I_i(\mathbf{p}(t))). \qquad (5)$$

Here the input current at the i-th STNO, I_i, is a linear function of the power outputs of all the STNOs in the previous time step,

$$I_i(\mathbf{p}(t)) = \left(\sum_{j=1}^{k} A_{ij}(t) p_j(t) + B_i(t)\right)^2. \qquad (6)$$

The matrices $\mathbf{A}(t)$ and vectors $\mathbf{B}(t)$ correspond to controllable external parameters. The input data is projected into the initial configuration of the system by assigning a correlated power output for a set of STNOs. For example, given an image with continuous values for each pixel, we assign an STNO to each pixel, and the value of the pixel corresponds to a normalized value of the power output of the respective STNO. The system then evolves as a functional response to the external parameters $\mathbf{A}(t)$ and $\mathbf{B}(t)$, such that at a final time T_{end}, the configuration of the system allows for an easier classification of the input data with a simple artificial neural network. The training of the dynamical neural network follows the optimal control theory developed in [9].

In general, the basic idea of the optimal control theory is to minimize a certain measure given a certain set of constraints [48], [49]. Specifically in the learning algorithm developed in [9], the authors demonstrated a method to minimize the loss function, which measures the accuracy of a neural network, imposing as constraints the equations of motion of the dynamical system. This physically inspired method leverages the impact of tunable external perturbations, in the case of this manuscript given by $\mathbf{A}(t)$ and $\mathbf{B}(t)$, to control the dynamics of the physical system such that at a final time step one obtains a configuration that minimizes the loss function. It relies on finding the optimal set of external perturbations that drives the system to an optimal and easily classifiable configuration.

This learning algorithm, therefore, requires prior knowledge of the power output curve $\mathbf{p}(I)$ of the STNOs in the network, to find the optimal set of $\mathbf{A}(t)$ and $\mathbf{B}(t)$. In realistic conditions, the specific power curve for each STNO, which depends on the specific composition and structure of the devices, can be generated before the training process.
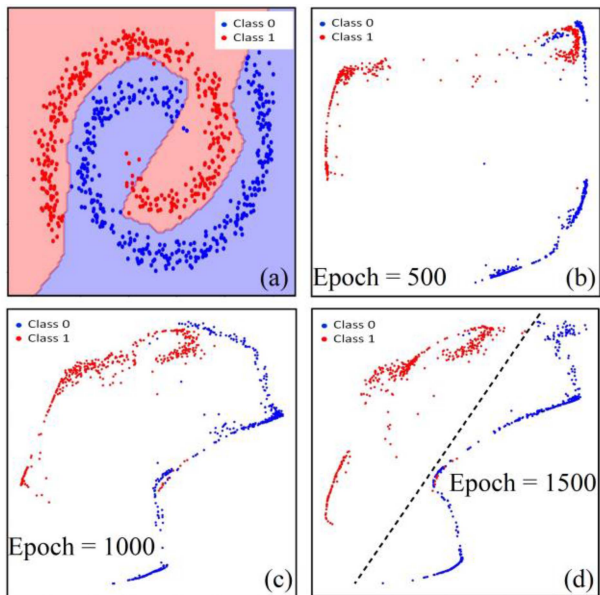
Fig. 3. (a) Spiral dataset for binary classification with classification in two regions identified by the colors blue and red. (b)–(d) final state $p(T_{end})$ for three different epochs during the training. (d) shows the distribution of the last training epoch, where a linear separation of the data is obtained.

Fig. 4. Average accuracy as a function of the standard deviation of the parameter a. The inset on the bottom right shows the probability distribution of the possible a values and the meaning of the standard deviation σ. Notice that a standard deviation $\sigma = 1$, means that a can take values between 0 and 8.

Fig. 5. Accuracy and Loss versus number of epochs for training and testing of the networks.

C. Scalability

State-of-art STNOs further increase the scalability of the dynamical neural network by requiring lower currents and no applied external magnetic field [36], [46]. In Fig. 2(c), we considered the experimentally realized STNO from [36], where the authors reported STNOs based on a magnetic tunnel junction configuration. The device consist of magnetic stacks of CoFeB separated by a tunnel barrier of MgO, see Fig. 2(c) inset. In the experiments performed, the authors obtained microwave emission with large output power, peak on the order of 10 nW, excited at ultra low currents, $I_c \approx -10\mu A$, and in the absence of external magnetic fields. It was also shown that thinner free layers poduce even higher power outputs at lower currents. The power output curve and the phenomenological model from (4) is plotted on Fig. 2(c) for a comparison. Moreover, we mention that the full experimental results reveal a greater non-linearity of the emmited power curve which can increase the accuracy of the dynamical neural network while decreasing the computation costs in term of energy for larger problems.

D. Results

We verified the performance of the STNO dynamical neural network with two simple classification tasks, see Figs. 3 and 4. First, we consider the binary classification problem given by data spread on a spiral, labeled in two categories as already presented for an optical-based solution [9]. This abstract dataset, made up of 1000 points, is divided into two: 80% of the data is used for the neural network training and 20% for testing. In this case, we mapped the two coordinates of the points on the spiral to the power output of two STNOs. The accuracy at the end of the training reaches 93.88% for the training set and
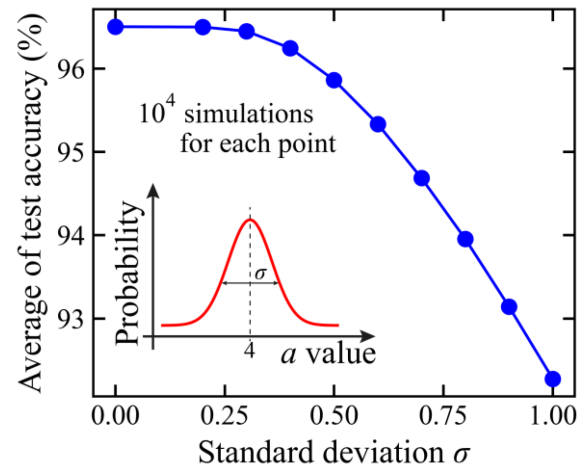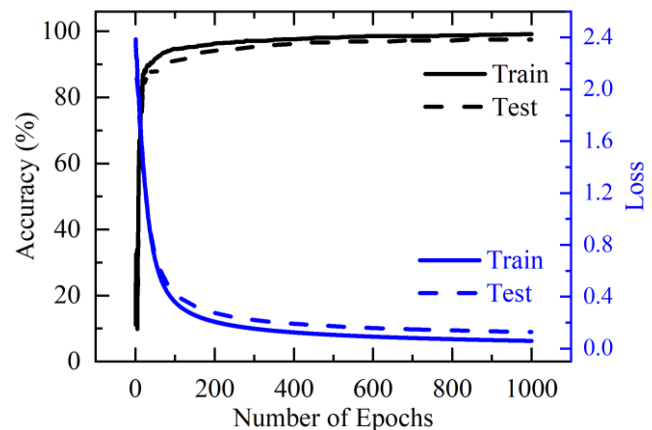
92% for the test set. Fig. 3(a) shows the results of the binary classification. Fig. 3(b)–(d) shows the final state $\mathbf{p}(T_{end})$ for three different epochs during the training of the network. As we can see in Fig. 3(d), at the end of the training the spiral dataset is disentangled and becomes linearly separable.

To assess the robustness of the proposed network against device-to-device variations, a statistical analysis was conducted on the network's accuracy in relation to deviations from the fitting model. A trained network, utilizing the fitting parameters employed for Fig. 3, was subjected to a classification task using devices that deviated from the training fitting model. Each device was characterized by a distinct value of parameter a in (4), drawn from a normal Gaussian distribution. Fig. 4 illustrates the average accuracy of 10000 runs for the classification task, plotted against the imposed standard deviation of the a parameters for each device. Remarkably, the network maintains high accuracy for deviations up to $\sigma = 0.25$. Consequently, it can be concluded that minor disparities between device behavior and the fitting model have minimal impact on device performance.

In the second case, we trained the network to recognize the categories of the "DIGIT" database, consisting of 1797 images of handwritten digits from 0 to 9, and image size of 8×8 pixels. The values for each of the 64 pixels were mapped to the power output of 64 STNOs. Also in this case, we divided the dataset into two: 80% of the data is used for the neural network training and 20% for testing. The accuracy obtained during training was 99.16% while the accuracy of the test was 97.5%. Fig. 5 shows the results in terms of accuracy and loss for the training and the test set, as a function of the number of epochs for the "DIGIT" database recognition.

To perform these tests, we consider a set of identical STNOs with a power output, (4), curve given by $a = 4$, $b = 5$, $I_c = 1$, and $Q = 2$.

III. SUMMARY AND CONCLUSION

The continuous development of spintronic technology and the performance of spintronic devices in terms of power consumption together with their intrinsic nonlinearities are the engine of development of neuromorphic spintronics which aims to design and realize spintronic-based energy-friendly hardware building block of neuromorphic computing. Here, we propose a new functionality for STNO realizing dynamical neurons which are characterized by an activation function which can be related to a differential equation. The main advantages of these devices are scalability, nanoscale size (those are the smallest auto-oscillators known in nature) [16], and potentially with a reduced cost being CMOS compatible. We show that the standard training algorithms based on functional optimization are sufficient to train the dynamical network implemented with a fully connected network of STNOs. We have tested the STNO dynamical neural network with two sets of data, including the "DIGIT" database, for digit recognition. The network achieved an accuracy of 99% for the latter. Those results achieved for a simple dataset are the proof-of-concept of the idea and we believe can motivate experimental efforts on developing a hardware implementation of this network. It is expected that a hardware implementation can be realized.

As an outlook, leveraging the transient dynamics of the spintronic component presents challenges that can be overcome through proper modeling and efficient calculation and modulation of external inputs. Addressing these challenges would greatly increase device efficiency and speed, benefiting from the nonlinear nature and fast characteristic time of magnetization dynamics.

REFERENCES

[1] B. J. MacLennan, "Analog Computation," in *Encyclopedia of Complexity and Systems Science*, Berlin, Germany: Springer, 2015, pp. 1–31.

[2] H. P. Graf and L. D. Jackel, "Analog electronic neural network circuits," *IEEE Circuits Devices Mag.*, vol. 5, no. 4, pp. 44–49, Jul. 1989, doi: 10.1109/101.29902.

[3] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, Apr. 2004, doi: 10.1126/science.1091277.

[4] L. G. Wright et al., "Deep physical neural networks trained with backpropagation," *Nature*, vol. 601, no. 7894, pp. 549–555, Jan. 2022, doi: 10.1038/s41586-021-04223-6.

[5] G. Wetzstein et al., "Inference in artificial intelligence with deep optics and photonics," *Nature*, vol. 588, no. 7836, pp. 39–47, Dec. 2020, doi: 10.1038/s41586-020-2973-6.

[6] D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, "Physics for neuromorphic computing," *Nature Rev. Phys.*, vol. 2, no. 9, pp. 499–510, Sep. 2020, doi: 10.1038/s42254-020-0208-2.

[7] T. W. Hughes, I. A. D. Williamson, M. Minkov, and S. Fan, "Wave physics as an analog recurrent neural network," *Sci. Adv.*, vol. 5, no. 12, Dec. 2019, Art. no. 2512, doi: 10.1126/sciadv.aay6946.

[8] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Comput. Sci. Rev.*, vol. 3, no. 3, pp. 127–149, Aug. 2009, doi: 10.1016/j.cosrev.2009.03.005.

[9] G. Furuhata, T. Niiyama, and S. Sunada, "Physical deep learning based on optimal control of dynamical systems," *Phys. Rev. Appl.*, vol. 15, no. 3, Mar. 2021, Art. no. 034092, doi: 10.1103/PhysRevApplied.15.034092.

[10] J. Torrejon et al., "Neuromorphic computing with nanoscale spintronic oscillators," *Nature*, vol. 547, no. 7664, pp. 428–431, Jul. 2017, doi: 10.1038/nature23011.

[11] X. Lin et al., "All-optical machine learning using diffractive deep neural networks," *Science*, vol. 361, no. 6406, pp. 1004–1008, Sep. 2018, doi: 10.1126/science.aat8084.

[12] G. W. Burr et al., "Neuromorphic computing using non-volatile memory," *Adv. Phys.: X*, vol. 2, no. 1, pp. 89–124, 2017, doi: 10.1080/23746149.2016.1259585.

[13] B. Dieny et al., "Opportunities and challenges for spintronics in the microelectronics industry," *Nature Electron.*, vol. 3, no. 8, pp. 446–459, 2020, doi: 10.1038/s41928-020-0461-5.

[14] J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, "Neuromorphic spintronics," *Nature Electron.*, vol. 3, no. 7, pp. 360–370, 2020, doi: 10.1038/s41928-019-0360-9.

[15] G. Finocchio, M. Di Ventra, K. Y. Camsari, K. Everschor-Sitte, P. Khalili Amiri, and Z. Zeng, "The promise of spintronics for unconventional computing," *J. Magn. Magn. Mater.*, vol. 521, 2021, Art. no. 20125, doi: 10.1016/j.jmmm.2020.167506.

[16] Z. Zeng, G. Finocchio, and H. Jiang, "Spin transfer nano-oscillators," *Nanoscale*, vol. 5, no. 6, 2013, Art. no. 2219, doi: 10.1039/c2nr33407k.

[17] T. Chen et al., "Spin-torque and spin-hall nano-oscillators," *Proc. IEEE*, vol. 104, no. 10, pp. 1919–1945, Oct. 2016, doi: 10.1109/JPROC.2016.2554518.

[18] A. Hamadeh et al., "Autonomous and forced dynamics in a spin-transfer nano-oscillator: Quantitative magnetic-resonance force microscopy," *Phys. Rev. B*, vol. 85, no. 14, Apr. 2012, Art. no. 140408, doi: 10.1103/PhysRevB.85.140408.

[19] A. Slavin and V. Tiberkevich, "Nonlinear auto-oscillator theory of microwave generation by spin-polarized current," *IEEE Trans. Magn.*, vol. 45, no. 4, pp. 1875–1918, Apr. 2009, doi: 10.1109/TMAG.2008.2009935.

[20] J. - V. Kim, V. Tiberkevich, and A. N. Slavin, "Generation linewidth of an auto-oscillator with a nonlinear frequency shift: Spin-torque nano-oscillator," *Phys. Rev. Lett.*, vol. 100, no. 1, Jan. 2008, Art. no. 017207, doi: 10.1103/PhysRevLett.100.017207.

[21] S. Kaka, M. R. Pufall, W. H. Rippard, T. J. Silva, S. E. Russek, and J. A. Katine, "Mutual phase-locking of microwave spin torque nano-oscillators," *Nature*, vol. 437, no. 7057, pp. 389–392, Sep. 2005, doi: 10.1038/nature04035.

[22] M. Romera et al., "Vowel recognition with four coupled spin-torque nano-oscillators," *Nature*, vol. 563, no. 7730, pp. 230–234, 2018, doi: 10.1038/s41586-018-0632-y.

[23] M. Zahedinejad et al., "Memristive control of mutual spin Hall nano-oscillator synchronization for neuromorphic computing," *Nature Mater.*, vol. 21, no. 1, pp. 81–87, Jan. 2022, doi: 10.1038/s41563-021-01153-6.

[24] S. Ikegawa, F. B. Mancoff, J. Janesky, and S. Aggarwal, "Magnetoresistive random access memory: Present and future," *IEEE Trans. Electron Devices*, vol. 67, no. 4, pp. 1407–1419, Apr. 2020, doi: 10.1109/TED.2020.2965403.

[25] X. Wang, P. Yu, and Y. Jiang, "Optimization on dry-etching process for high-density STT-MRAM," *IEEE Trans. Nanotechnol.*, vol. 20, pp. 161–167, 2021, doi: 10.1109/TNANO.2021.3058260.

[26] K. Ali, F. Li, S. Y. H. Lua, and C. H. Heng, "Area efficient high through-put dual heavy metal multi-level cell SOT-MRAM," *IEEE Trans. Nanotechnol.*, vol. 19, pp. 613–619, 2020, doi: 10.1109/TNANO.2020.3012669.

[27] S. Jung et al., "A crossbar array of magnetoresistive memory devices for in-memory computing," *Nature*, vol. 601, no. 7892, pp. 211–216, 2022, doi: 10.1038/s41586-021-04196-6.

[28] P. M. Braganca, B. A. Gurney, B. A. Wilson, J. A. Katine, S. Maat, and J. R. Childress, "Nanoscale magnetic field detection using a spin torque oscillator," *Nanotechnology*, vol. 21, no. 23, May 2010, Art. no. 235202, doi: 10.1088/0957-4484/21/23/235202.

[29] P. J. Metaxas et al., "Sensing magnetic nanoparticles using nano-confined ferromagnetic resonances in a magnonic crystal," *Appl. Phys. Lett.*, vol. 106, no. 23, Jun. 2015, Art. no. 232406, doi: 10.1063/1.4922392.

[30] D. I. Albertsson, J. Akerman, and A. Rusu, "A magnetic field-to-digital converter employing a spin-torque nano-oscillator," *IEEE Trans. Nanotechnol.*, vol. 19, pp. 565–570, 2020, doi: 10.1109/TNANO.2020.3007344.

[31] A. Houshang et al., "Phase-binarized spin hall nano-oscillator arrays: Towards spin hall ising machines," *Phys. Rev. Appl.*, vol. 17, no. 1, Jan. 2022, Art. no. 014003, doi: 10.1103/PhysRevApplied.17.014003.

[32] D. I. Albertsson, M. Zahedinejad, A. Houshang, R. Khymyn, J. Åkerman, and A. Rusu, "Ultrafast ising machines using spin torque nano-oscillators," *Appl. Phys. Lett.*, vol. 118, no. 11, Mar. 2021, Art. no. 112404, doi: 10.1063/5.0041575.

[33] M. Zahedinejad et al., "Two-dimensional mutually synchronized spin hall nano-oscillator arrays for neuromorphic computing," *Nature Nanotechnol.*, vol. 15, no. 1, pp. 47–52, Jan. 2020, doi: 10.1038/s41565-019-0593-9.

[34] D. E. Nikonov et al., "Coupled-oscillator associative memory array operation for pattern recognition," *IEEE J. Explor. Solid-State Comput. Devices Circuits*, vol. 1, pp. 85–93, Dec. 2015, doi: 10.1109/JX-CDC.2015.2504049.

[35] L. Martins et al., "Non-volatile artificial synapse based on a vortex nano-oscillator," *Sci. Rep.*, vol. 11, no. 1, Dec. 2021, Art. no. 16094, doi: 10.1038/s41598-021-95569-4.

[36] Z. Zeng et al., "Ultralow-current-density and bias-field-free spin-transfer nano-oscillator," *Sci. Rep.*, vol. 3, no. 1, Mar. 2013, Art. no. 1426, doi: 10.1038/srep01426.

[37] R.T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, "Neural ordinary differential equations," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 31, 2018. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2018/hash/69386f6bb1dfed68692a24c8686939b9-Abstract.html

[38] M. Di Ventra, *MemComputing*. Oxford, U.K.: Oxford Univ. Press, 2022.

[39] J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, "Information processing capacity of dynamical systems," *Sci. Rep.*, vol. 2, no. 1, pp. 1–7, Jul. 2012, doi: 10.1038/srep00514.

[40] G. Tanaka et al., "Recent advances in physical reservoir computing: A review," *Neural Netw.*, vol. 115, pp. 100–123, 2019, doi: 10.1016/j.neunet.2019.03.005.

[41] M. D. Stiles and A. Zangwill, "Anatomy of spin-transfer torque," *Phys. Rev. B*, vol. 66, no. 1, Jun. 2002, Art. no. 014407, doi: 10.1103/PhysRevB.66.014407.

[42] J. C. Sankey, Y. T. Cui, J. Z. Sun, J. C. Slonczewski, R. A. Buhrman, and D. C. Ralph, "Measurement of the spin-transfer-torque vector in magnetic tunnel junctions," *Nature Phys.*, vol. 4, no. 1, pp. 67–71, 2008, doi: 10.1038/nphys783.

[43] L. Mazza et al., "Computing with injection-locked spintronic diodes," *Phys. Rev. Appl.*, vol. 17, no. 1, Jan. 2022, Art. no. 014045, doi: 10.1103/PhysRevApplied.17.014045.

[44] B. Fang et al., "Giant spin-torque diode sensitivity in the absence of bias magnetic field," *Nature Commun.*, vol. 7, no. 1, Apr. 2016, Art. no. 11259, doi: 10.1038/ncomms11259.

[45] Z. Zeng et al., "High-power coherent microwave emission from magnetic tunnel junction nano-oscillators with perpendicular anisotropy," *ACS Nano*, vol. 6, no. 7, pp. 6115–6121, Jul. 2012, doi: 10.1021/nn301222v.

[46] H. Fulara et al., "Giant voltage-controlled modulation of spin Hall nano-oscillator damping," *Nature Commun.*, vol. 11, no. 1, pp. 1–7, Aug. 2020, doi: 10.1038/s41467-020-17833-x.

[47] M. N. Baibich et al., "Giant magnetoresistance of (001)Fe/(001)Cr magnetic superlattices," *Phys. Rev. Lett.*, vol. 61, no. 21, pp. 2472–2475, Nov. 1988, doi: 10.1103/PhysRevLett.61.2472.

[48] E. Todorov, "Optimal control theory," *Bayesian Brain*, vol. 21, no. C, 2006, pp. 268–298.

[49] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, "Neural ordinary differential equations," 2018, *arXiv:1806.07366*.