

PHYSICAL SCIENCES

Harnessing spatiotemporal transformation in magnetic domains for nonvolatile physical reservoir computing

Jing Zhou^{1*}, Jikang Xu², Lisen Huang¹, Sherry Lee Koon Yap¹, Shaohai Chen¹, Xiaobing Yan^{2*}, Sze Ter Lim^{1*}

Combining physics with computational models is increasingly recognized for enhancing the performance and energy efficiency in neural networks. Physical reservoir computing uses material dynamics of physical substrates for temporal data processing. Despite the ease of training, building an efficient reservoir remains challenging. Here, we explore beyond the conventional delay-based reservoirs by exploiting the spatiotemporal transformation in all-electric spintronic devices. Our nonvolatile spintronic reservoir effectively transforms the history dependence of reservoir states to the path dependence of domains. We configure devices triggered by different pulse widths as neurons, creating a reservoir featured by strong nonlinearity and rich interconnections. Using a small reservoir of merely 14 physical nodes, we achieved a high recognition rate of 0.903 in written digit recognition and a low error rate of 0.076 in Mackey-Glass time series prediction on a proof-of-concept printed circuit board. This work presents a promising route of nonvolatile physical reservoir computing, which is adaptable to the larger memristor family and broader physical neural networks.

INTRODUCTION

Recent advances in neural networks attempt to incorporate physical principles into computational models, owing to the limitation of data-driven approaches on multiscale problems and the isomorphism between mathematical operation and hardware physics (1–6). For instance, the physics-informed neural network enforces physical laws as constraints or prior knowledge during the in silico learning process (2, 5). The deep physical neural network (PNN) explores beyond in silico learning by training physical transformation with backpropagation (3). A closely related yet distinct approach is physical reservoir computing (PRC), which harnesses the rich dynamics of physical process for machine learning without training the physical process itself (1, 4, 6). The algorithm of RC was originally derived from the echo state networks and liquid state machines (7–9). It is an improved computing paradigm for temporal data processing since the conventional recurrent neural network (RNN) suffers from an exploding (and vanishing) gradient problem (9). The architecture of RC is featured by an input-driven reservoir of complex but fixed interconnectivity and a memoryless readout layer. Although RC benefits from clearly defined mathematical preconditions and streamlined training process, translating these advantages into practicable reservoir-building strategies remains a longstanding challenge (10). This is where the physical implementation of RC, or PRC, was introduced. A wealth of physical substrates, including photonic (11–16), mechanical (17–21), electronic (22–26), quantum (27–31), and even biochemical systems (32–36), proves to have the intrinsic memory and nonlinear dynamics required by the black box-like reservoir. These physical substrates result in a hybrid in situ–in silico RC architecture that is more resource efficient for information processing.

Despite the versatility of PRC, the construction of nearly all physical reservoirs relies on some delay-based architecture. This means that the effect of an input to the physical system is not immediate but occurs after a time delay. For instance, using explicitly implemented delay lines in the mechanical (37), optical (38), and electrical (23) circuits, the current output is fed back to the system to modify subsequent inputs. These circuitry-based reservoirs suffer from the less efficient serial operation enforced by the arbitrary delay line and limited interconnectivity. More recent advances achieved delays by leveraging intrinsic relaxation processes of materials, such as magnetic damping of a spin-torque nano-oscillator (11, 26), filament rupture in metal oxide (22, 39), ferroelectric polarization back-switching (25), and reagent consumption in redox reaction (33). These relaxation processes naturally propagate the effects of an input to future time steps. The relaxation-based approaches demonstrate high fidelity to the black box-like nature of reservoir, thanks to the hidden and extensive interconnection associated with nonlinear material dynamics. However, such material- or device-based reservoirs are restricted by the characteristic time constants of the relaxation processes, limiting their effectiveness to signals of similar timescale only. As expected, the delay-based conventional physical reservoirs are essentially volatile, where the reservoir states are lost upon the removal of input. This compromises the application of PRC in many real-life situations where the data are generated continuously in a prolonged period, such as near-sensor or in-sensor computing (1).

In this work, we address these challenges by demonstrating nonvolatile PRC using all-electric spintronic memristors. Instead of the physical sense of time, we interpret the history dependence of reservoir states as the path dependence of ferromagnetic domains. This transforms a volatile reservoir in the temporal domain to a nonvolatile reservoir in the spatial domain, removing the limit on input timescale. The electrical readouts of our reservoir states are finely manipulated by current-induced domain wall motion, where the combination of spin-orbit torque (SOT) and thermal effects induces strong nonlinearity. Stacking multiple dynamic responses of the ferromagnetic domains generates a high-dimensional reservoir

¹Institute of Materials Research and Engineering (IMRE), Agency for Science Technology and Research (A*STAR), 2 Fusionopolis Way, #08-03 Innovis, Singapore 138634, Republic of Singapore. ²Key Laboratory of Brain-like Neuromorphic Devices and Systems of Hebei Province, Key Laboratory of Optoelectronic Information Materials of Hebei Province, Hebei University, Baoding 071002, China.

*Corresponding author. Email: zhou_jing@imre.a-star.edu.sg (J.Z.); yanxiaobing@ime.ac.cn (X.Y.); lim_sze_ter@imre.a-star.edu.sg (S.T.L.)

characterized by strong interconnection, without the need for nonlinear preprocessing masks or time multiplexing. Our reservoir shows competitive performance in both written digit classification and Mackey-Glass time series prediction. The potential of on-chip learning is also demonstrated using a customized printed circuit board (PCB).

RESULTS

Device and reservoir building

All-electric spintronic memristors are used as the hardware building blocks of our reservoir (see Materials and Methods for details). SOT usually requires an external magnetic field for deterministic switching, introducing additional constraints on circuits while complicating the device operation. Our devices internalize this field, allowing them to be manipulated purely by electrical signals: The out-of-plane magnetization is modified by input voltage (V_p) pulses and gauged using Hall resistance (R_H). This field-free feature is indispensable for circuit-level application toward large-scale integration with conventional chip technologies. Our devices are essentially nonvolatile as their magnetization states are retained without any input. More device properties were reported in our previous work (40).

Our reservoir-building strategies tap on the nonlinear path dependence of ferromagnetic domains. Figure 1A illustrates the path dependence in typical R_H - V_p responses. Paths 1 and 2 form the full magnetic hysteresis loop where the middle states and saturation states are labeled as **A** and **B** and **C** and **D**, respectively. Despite being both at the same macroscopic readout of $R_H = 0.5$, states **A** and **B** have different prior states (path 1 for **A** and path 2 for **B**). As a result, state **A** will take path 1, but state **B** will take path 3 toward state **C**. Similarly, states **A** and **B** take paths 4 and 2 toward state **D**, respectively. Paths 3 and 4 are the widely observed minor hysteresis loops in SOT-driven magnetization switching (41–43). Microscopically, the minor loops arise from the different domain structures of states **A** and **B**, which respond differently to input stimuli, forming different future trajectories. This translates conveniently into the basic dynamics in an echo state network: The reservoir state $\mathbf{x}(t)$ at time step t depends on both current input $\mathbf{u}(t)$ and past states $[\mathbf{x}(t-n), n \in \mathbb{R}]$.

The electrical switching of perpendicular magnetization is highly sensitive to the input pulse width (τ_p). Switching triggered by small τ_p (<1 ns) is dominated by SOT, whereas thermal effects, such as Joule heating-induced lowering of switching barrier, dominate for large τ_p (>1 μ s). As a result, switching depends nonlinearly on τ_p , which is usually observed as a disproportionate increase in V_p or the number of pulses (N_p) required for full switching as τ_p decreases (44–46). Moreover, the switching behavior becomes more complex for intermediate pulse widths (1 ns $< \tau_p < 1$ μ s) since the contributions of SOT and thermal effects are comparable. Subsequently, tuning the working τ_p interval—defined as the dynamic range (DR)—of the spintronic device creates multiple virtual nodes (VNs) (37), each producing a different dynamic response to the same input (Fig. 1B). Owing to the strong nonlinearity in τ_p , the input-output characteristics of VN depends on not only the width ($\Delta\tau_p$) but also the exact position of DR. For instance, DR 2 and DR 4 in Fig. 1B, although having the same $\Delta\tau_p$, will produce two distinct VNs. This allows a diverse combination of VNs, or neurons, to be stacked, collectively forming a reservoir of high-dimensional data space (Fig. 1C). Compared with VNs in photonic PRCs, our VNs only share the similar neural functions. However, our VNs are generated in fundamentally different approaches because they do not rely on time-domain multiplexing enabled by delay lines (47). Our reservoir merges two advantages of conventional approaches on top of being nonvolatile. First, we use arrays of spintronic devices (14 in this work) of nominally the same physical characteristics (fig. S1) for the simultaneous generation of VNs, leading to the more preferred parallel computing for analog neural networks (1, 48). The very small device-to-device variations also enrich the reservoir dynamics by further diversifying the neuron responses (note S1). Next, as we will show later, the same overarching τ_p dependence helps to establish extensive and complex interconnections among neurons, similar to the role of a long delay line in optical circuit (14, 38).

Nonlinearity and fading memory

We first demonstrate the path dependence using Hall measurement (Fig. 2A). The current-induced SOT in a longitudinal V_p partially reverses the magnetization, which is monitored by the Hall resistance

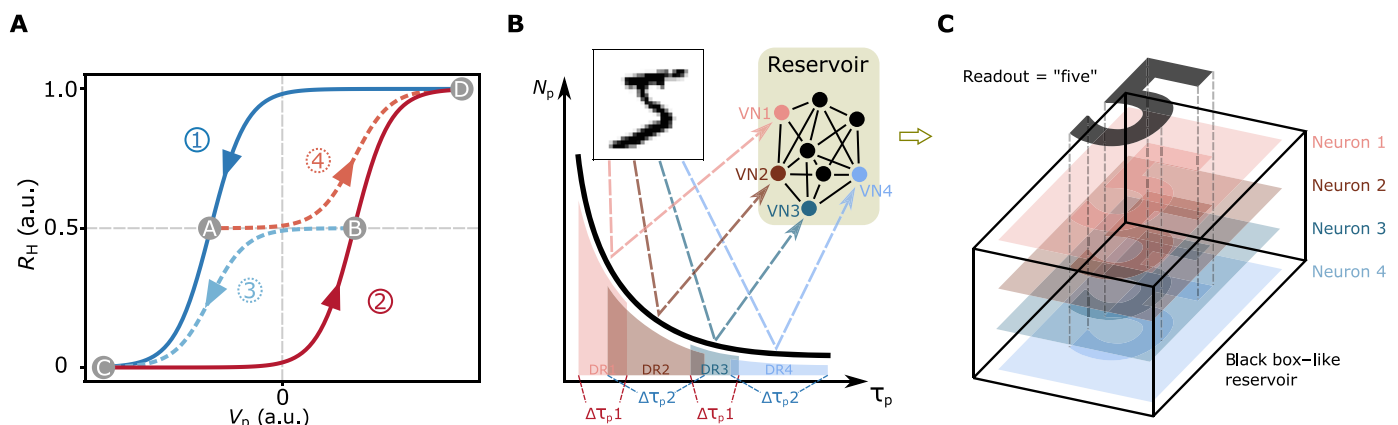


Fig. 1. Reservoir-building strategies. (A) Path dependence in electric switching of magnetization, with voltage pulse (V_p) as the input and Hall resistance (R_H) as the output. Solid (dotted) lines indicate full (partial) switching loops. Filled and hollow circles label magnetization states and switching paths, respectively. (B) Nonlinear dependence of number of voltage pulses (N_p) for full switching on pulse width (τ_p). Inset illustrates the creation of virtual nodes (VNs) by projecting input at different pulse width intervals defined as the dynamics ranges (DRs). $\Delta\tau_p$ refers to the difference in maximum and minimum τ_p . (C) A black box-like high-dimensional reservoir by stacking multiple VNs. a.u., arbitrary units.

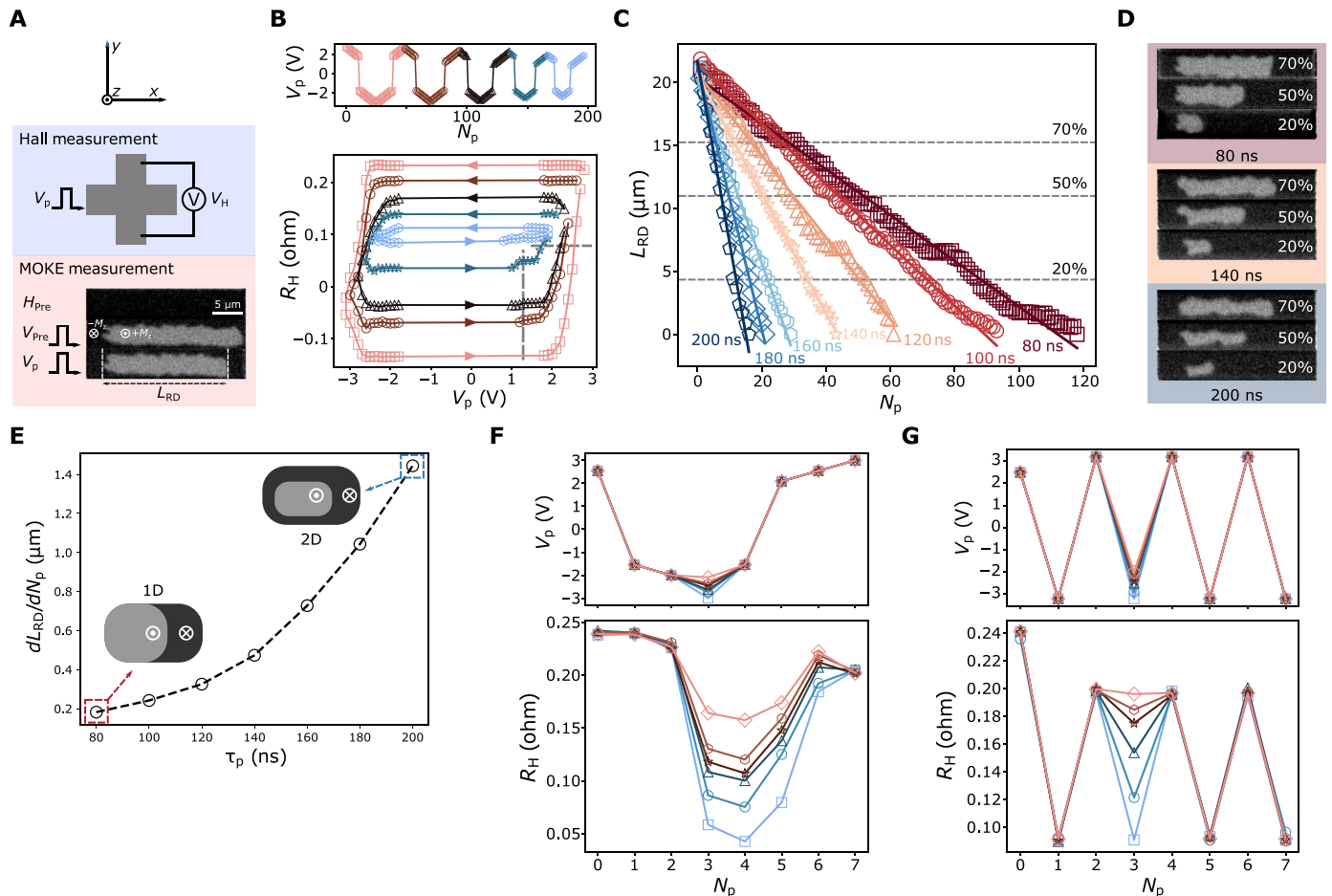


Fig. 2. Switching dynamics in spintronic memristor. (A) Schematic setup of Hall and MOKE measurements. V_p , V_{pre} , and H_{pre} denote the amplitudes of switching pulse, preset pulse, and preset field, respectively. L_{RD} refers to the length of the remaining domain. (B) Variation of Hall resistance (R_H) with decreasing V_p . $\tau_p = 200$ ns. (C) Variation in L_{RD} with pulse number (N_p) for $V_p = -4.1$ V and $\tau_p = 80$ to 200 ns. (D) MOKE images for $\tau_p = 80$, 140, 200 ns when 20, 50, and 70% when L_{RD} remain. (E) Extracted domain wall motion per pulse (dL_{RD}/dN_p) from line fits in (C). Inset shows the one-dimensional (1D) and 2D domain wall motion at low and high τ_p . (F and G) R_H induced by six pulse trains where only V_p for $N_p = 3$ is different. $N_p = 0$ is a preset pulse. The maximum absolute value of voltage pulse is 3 V in (F) and 3.2 V in (G).

$R_H = V_H/I_L$, considering the transverse Hall voltage V_H induced by a longitudinal current I_L . By applying a train of V_p segments of opposite polarities but diminishing amplitudes, we observe minor switching loops that gradually spiral inward, as shown in Fig. 2B. The vertical dotted line highlights that the same V_p triggers different changes in R_H depending on the present state. The horizontal dotted line, on the other hand, implies that the same R_H state requires different V_p to change depending on the history of the state. Collectively, we conclude the simultaneous dependence of R_H on both input V_p and the history of R_H .

Next, using polar magneto-optic Kerr effect (MOKE) microscopy, we demonstrate that SOT-driven domain wall motion (49) is nonlinear in τ_p , which provides the practical tool for inducing nonlinear reservoir dynamics. Referring to Fig. 2A, input V_p switches the positive out-of-plane magnetization ($+M_z$, gray color) to the negative one ($-M_z$, black color), leading to the leftward movement of domain walls near the right terminal of the microstrip. We define a parameter L_{RD} in Fig. 2A to quantify domain wall motion. In Fig. 2C, L_{RD} is roughly linear in the number of pulses N_p for $\tau_p = 80$ to 200 ns, indicating the stable effect of V_p on domain walls. We verify that the

step-like features near $L_{RD} \approx 7 \mu m$ for $\tau_p = 120$ ns and $\tau_p = 80$ ns are unlikely to be caused by local pinning (note S2 and fig. S2). In a closer examination, we find that the domain wall motion appears to be one-dimensional (1D) when $\tau_p = 80$ ns, as shown by the domain snapshots in Fig. 2D. However, as τ_p increases to 200 ns, the $+M_z$ domains shrink in both x and y directions, indicating 2D domain wall motion (movies S1 to S3). This is consistent with the extreme case of V_{pre} , where $\tau_p = 5 \mu s$, such that one pulse is sufficient to switch most areas of the $+M_z$ domains (Fig. 2A). Using the linear fits in Fig. 2C, we summarize in Fig. 2E the extracted domain movement per pulse (dL_{RD}/dN_p), which is unambiguously nonlinear in τ_p . The relation in Fig. 2E implies that if an input is encoded as τ_p of an electric pulse, any linear changes in the input will lead to nonlinear changes in L_{RD} thus nonlinear changes in electric output (i.e., R_H). Furthermore, since R_H is proportional to the area (i.e., 2D) of $\pm M_z$ whereas L_{RD} is 1D, we expect the nonlinearity in $R_H - \tau_p$ to be even stronger than $L_{RD} - \tau_p$. Therefore, modulating τ_p is an effective means of high-dimensional nonlinear transformation of input data.

The mathematical preconditions of a working reservoir (11, 22)—point-wise separation property and fading memory—are readily

achieved in our spintronic neuron. The separation property requires the reservoir states $\mathbf{x}(t)$ to distinguish any differences in the paired input $\mathbf{u}(t)$ and output $\mathbf{y}(t)$ at time t . Fading memory means that the reservoir state $\mathbf{x}(t)$ depends on both present and past inputs $[\mathbf{u}(t-n), n \in \mathbb{R}]$, and this dependence weakens as n increases. To show these, we apply six trains of seven pulses to the same spintronic memristor with the same initial state, where only the third pulse is different in each train. The resulting R_H states are summarized in Fig. 2F. The state trajectories begin to branch out at $N_p = 3$, implying the separation property since the memristor recognizes small differences in the input. Moreover, the differences in R_H diminish for $N_p \geq 4$, showing that the effects of inputs at $N_p = 3$ extend into the future and eventually fade out. In addition, our devices have highly repeatable readout and excellent endurance, which are crucial for the stable operation of the network. In the single-domain limit, the estimated thermal stability is high, consistent with the demonstrated long data retention time (>10 days) (note S3 and figs. S3 to S5). However, caution must be taken to exploit this fading memory. In a control experiment, we increase the maximum magnitude of V_p from 3 V in Fig. 2F to 3.2 V in Fig. 2G and assign alternating pulse polarities. The resulting R_H trajectories only differ at $N_p = 3$, showing no fading memory at all, because values of V_p in Fig. 2G are strong enough to preset the R_H state such that prior history is washed out. Therefore, a prerequisite for establishing fading memory is to have the input energy below the preset level. This echoes the delay-based PRCs (11, 23, 26), where the time delay between successive inputs must be less than the characteristic time constant to keep the fading memory property. However, our approach differs from the delay-based PRCs in two critical aspects. First, our architecture has temporal information encoded within individual inputs, lifting the need for pulse train-like input of certain frequencies. The resulting task performance has a more flexible timescale since the reservoir no longer requires specific rates of data input. Second, our method of encoding temporal information is only possible with nonvolatile memristor, without which the reservoir would require continuous input streams to retain its state and becomes delay based.

Written digit recognition

In this section, with reference to the benchmark task of written digit recognition, we illustrate the network structure and training protocol in our PRC. Referring to Fig. 3A, an input (u_n) of digit “5” from the Modified National Institute of Standards and Technology (MNIST) database (50) is mapped to pulse trains of fixed amplitudes (-3.0 and 2.8 V for negative and positive inputs, respectively), where the pulse width ($50 \text{ ns} < \tau_p < 200 \text{ ns}$) varies linearly with the magnitude of input (see table S1 for examples). Figure 3B shows the reconstructed responses of four different VNs to the same input digits. Apparently, all VNs successfully capture the outlines of the digits with varying details, as if the VNs “perceive” the same input in slightly different manners. Notably, horizontal streaks are observed in VN responses. They arise from the large number of zeros in the pixels of the input images. The corresponding voltage pulse has minimal τ_p that barely alters the magnetization thus R_H . We consider this an advantage of nonvolatile reservoir, which can tolerate a larger number of low-energy input while memorizing the state. In other words, our spintronic reservoir does not have to be continuously excited. Furthermore, unlike prior works (22, 25), we did not binarize the image pixels, making our task more challenging. The R_H trains produced by all VNs are concatenated to form the reservoir

state x_n of one digit, and all input digits’ reservoir states collectively form the final reservoir X .

In Fig. 3A, we also illustrate the software-based reservoir (red arrows) in a classic echo state network (see Materials and Methods for details), which differs from our hardware-based reservoir (blue arrows) in three aspects. First, our reservoir requires no dimension-raising preprocessing mask W_{in} , which is indispensable in prior approaches (14, 23, 26, 30, 37, 38). In some cases of time multiplexing, the preprocessing mask relies on trigonometric functions to introduce nonlinearity while explicitly adding the dimension of time, which inevitably undermines the role of reservoir. Next, our reservoir dynamics is not determined by an arbitrarily designed internal state matrix W , substantially lowering energy consumption since matrix manipulation is computationally expensive. Last, our reservoir has the fading memory and the nonlinearity internalized in the spintronic memristor’s switching dynamics, instead of relying on the leak rate and activation function. Overall, our physical reservoir has all essential characteristics of a software-based reservoir while using much less predetermined hyperparameters and computation steps. The training procedure in our physical RC is similar to prior

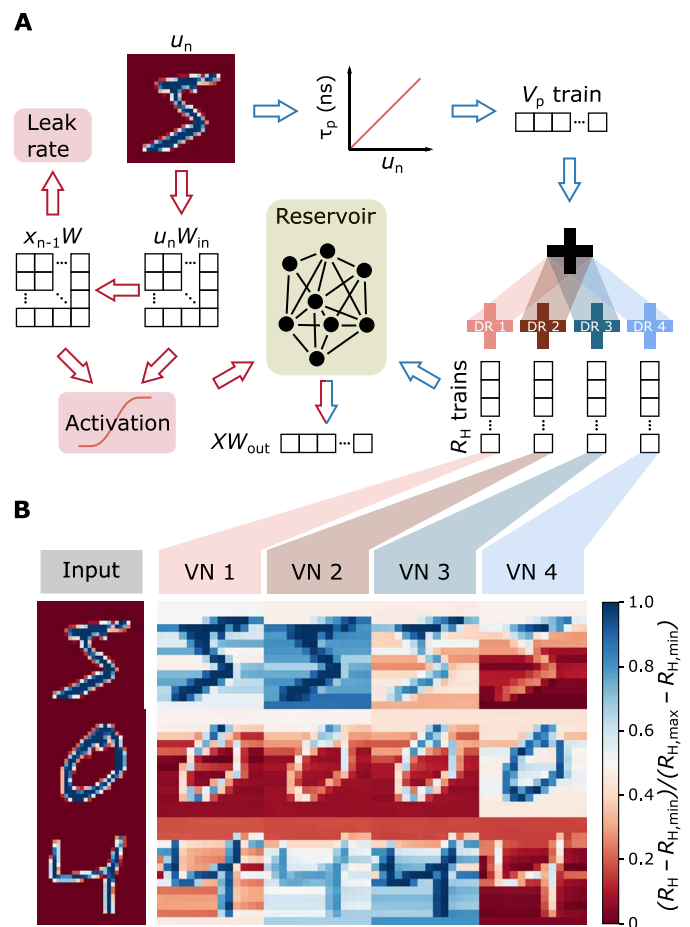


Fig. 3. Network structure of reservoir computing. (A) Comparison of the structures of a classic echo state network (red arrows) and our physical RC (blue arrows). (B) The same spintronic device responds to the same input differently when different DRs are chosen, resulting in different VNs. $R_{H,max}$ and $R_{H,min}$ refer to the maximum and minimum R_H readouts, respectively.

approaches—we use the combination of pseudoinverse and ridge regression to compute the output weight W_{Out} (see Materials and Methods for details).

To deliver practically meaningful task performance, we develop a proof-of-principle PCB (Fig. 4A) while exercising 10-fold cross-validation for training (fig. S6). The PCB elevates the technical sophistication of our approach from measurement of discrete devices to circuit-level integration. It ensures the parallel connection of 14 devices to the field programmable gating array (FPGA) during the course of computation, mimicking the environment of on-chip learning toward embedded applications. Cross-validation ensures the statistical validity—a requirement frequently overlooked in prior works—where up to 20,000 samples are taken as the data pool and grouped randomly into 10 equally populated subsets. Each subset serves as the testing data, while the rest serves as the training data in a single trial, which repeats 10 times to produce an average performance. Figure 4B shows some single-trial confusion matrices for data pools of size N_s . When $N_s = 2000$, the training accuracy is unity ($A_{\text{Train}} = 1.00$), but the testing accuracy is very low ($A_{\text{Test}} = 0.633$). This implies that our reservoir can easily interpret a small dataset, but these samples are too few to generalize for untrained data. Despite the low A_{Test} score, our reservoir is making human-like mistakes, such as confusing “9” with “4” and “8” with “3” (Fig. 4B, middle). Expectedly, increasing N_s decreases A_{Train} and increases A_{Test} , both of which eventually saturate (Fig. 4C).

By progressively increasing the number of neurons, or VNs (N_{VN}), we demonstrate the robustness of our reservoir building

strategy (Fig. 4D). In the control experiment with $N_{\text{VN}} = 0$, the reservoir includes raw inputs only, producing the lowest accuracy rates, therefore implying the crucial role of hardware dynamics. The largest performance enhancement comes from the first neuron, where both A_{Train} and A_{Test} increase from less than 0.7 for $N_{\text{VN}} = 0$ to roughly 0.8 for $N_{\text{VN}} = 1$. Additional neurons further increase the task performance, which appears to saturate at $N_{\text{VN}} = 6$. The highest average accuracy rates are $A_{\text{Train}} = 0.916 \pm 0.001$ and $A_{\text{Test}} = 0.880 \pm 0.001$ with $N_s = 20,000$ and $N_{\text{VN}} = 6$. These results demonstrate that stacking different neurons derived from the same spintronic dynamics effectively enhances reservoir dynamics and task performance. However, the observed saturation and the small number of neurons in reservoir remain intriguing, which we will investigate further in the next section.

Mackey-Glass time series prediction

While written digit classification only showcases our physical RC's broader capabilities in machine learning tasks, time series prediction is a more specific benchmark task for RC (or RNN in general) (1, 6). We use a dataset (51) of the Mackey-Glass equation (52), which describes a chaotic time series widely applied in physical and biological systems. The same linear mapping process in Fig. 3A is used to generate input to devices, where pulses have fixed amplitudes (-2.8 and 2.2 V) and varying pulse widths ($50 \text{ ns} < \tau_p < 200 \text{ ns}$). After flushing the first 100 data points in the series to remove the initial condition, we train 1000 data points to find W_{Out} for autonomously predicting the next 400 data points. We prepared 156 neurons

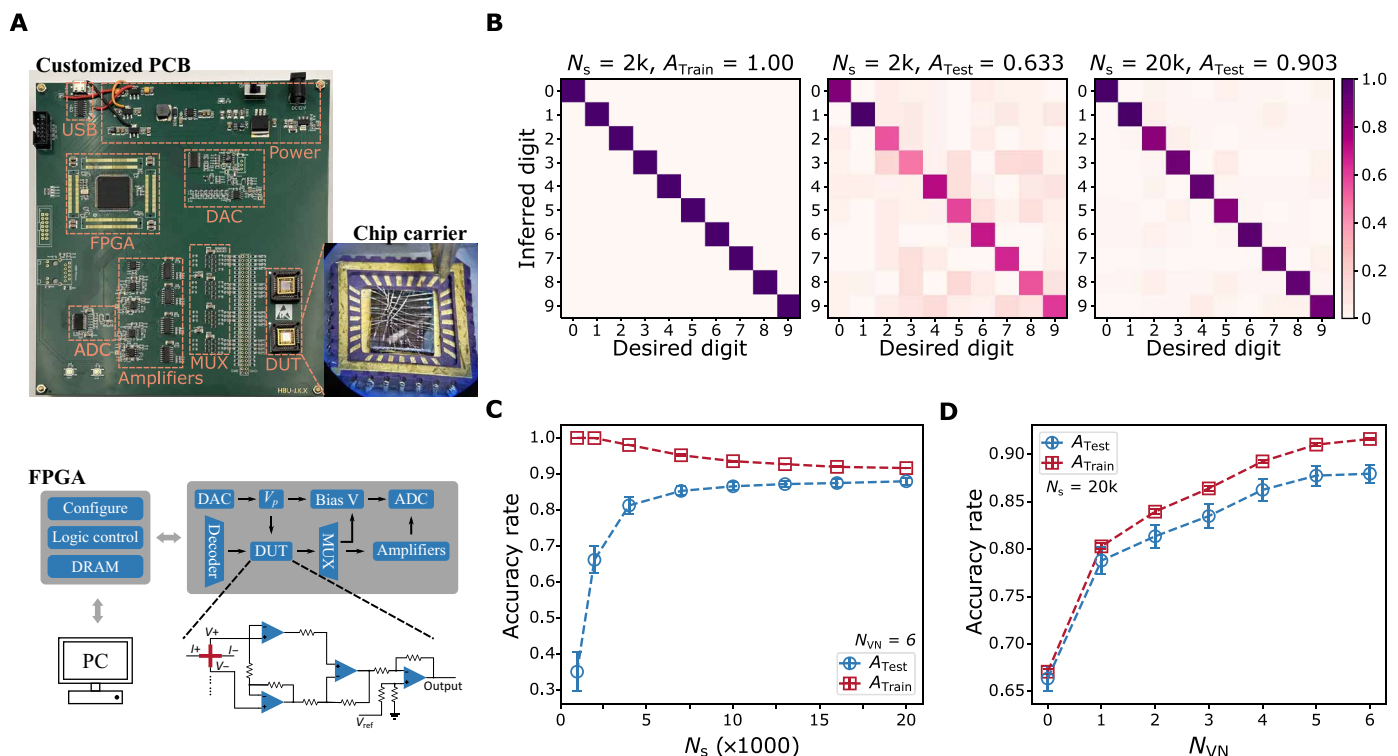


Fig. 4. Written digit recognition with spintronic physical RC. (A) Illustration of hardware circuit. Top: A photo of the customized PCB and a chip carrier. Bottom: Schematic components of the circuit. The spintronic devices and their supporting components are labeled as devices under test (DUT). (B) Confusion matrices for training (left) and testing (middle and right). (C) Dependence of the training (A_{Train}) and testing (A_{Test}) accuracy rates on the number of trained samples (N_s). (D) Dependence of A_{Train} and A_{Test} on the number of neurons (N_{VN}). Error bars in (C) and (D) indicate 1 SD from 10-fold cross-validation. DAC, digital-to-analog converter; MUX, multiplexer.

from 14 spintronic memristors, corresponding to 156 different DRs of 14 $\Delta\tau_p$. Inspired by the strategy of “titration of chaos” (53), we add the neurons one at a time to the reservoir and compute the normalized root mean square error (NRMSE) of testing. A neuron is kept in the reservoir only if it decreases the NRMSE. This process repeats for all 156 neurons to obtain the final reservoir and minimized NRMSE of each trial (Fig. 5A). The 156 neurons are randomized into 100 distinct trial sequences to study the effects of neuron interaction on reservoir dynamics. Figure 5B shows the prediction during a particular trial. As more neurons are added, the reservoir produces better prediction for the untrained data. The final reservoir has 57 neurons, and its prediction almost superimposes with the target, showing robust performance of our physical reservoir on the Mackey-Glass time series.

A quantitative analysis on the 100 trials reveals the strong and complex interconnection among neurons. The reduction of NRMSE with N_{VN} for five representative trials is plotted in Fig. 5C in reddish colors. Regardless of the different trends and final N_{VN} values, the five trials have saturation NRMSEs close to 0.08, which also applies to all other trials (fig. S7). This shows the enhancing effect of more neurons on the reservoir dynamics, leading to better predictions on time series, consistent with the observation in Fig. 4D. Moreover, we find that a specific neuron is included in the final reservoir of some trials but not others. This is strong evidence of the complex interaction among neurons because the same neuron can be either helpful or detrimental depending on the existing neurons in the reservoir. It also explains the saturated performance in Figs. 4D and 5C as the additional neuron does not interact with the existing ones well.

We summarize the occurrence frequency of neurons in the final reservoirs of all trials to further investigate neuron interaction. Referring to Fig. 5D, neurons with wider DRs ($\Delta\tau_p \geq 100$ ns) are picked more frequently by the reservoir than those with narrower DRs ($\Delta\tau_p \leq 20$). We attribute this to the fact that wider DRs correspond to stronger nonlinearity. Unexpectedly, some neurons with

$\Delta\tau_p \leq 20$, where the underlying device responses is approximately linear in τ_p , also have fairly good frequency (~ 20). This is another strong evidence of the hidden interactions among neurons beyond their individual contribution to the reservoir dynamics. Therefore, the results in Fig. 5 (C and D) validate with our hypothesis (Fig. 1B) that the overall τ_p dependence of hardware facilitates the interaction among VNs.

The hidden interaction among neurons of different $\Delta\tau_p$ is also demonstrated using a control experiment. Instead of multiple $\Delta\tau_p$, we set up 50 different neurons of a single $\Delta\tau_p$ ($= 100$ ns) and repeat the reservoir construction process in Fig. 5A. Five representative NRMSE trends are plotted in Fig. 5C in bluish colors. The saturated NRMSE is close to 0.15—roughly 80% more than that of multiple $\Delta\tau_p$ —indicating a performance deterioration with single $\Delta\tau_p$. Moreover, for most trials with single $\Delta\tau_p$, the reduction in NRMSE is slower than those with multiple $\Delta\tau_p$, implying that additional neuron of the same $\Delta\tau_p$ is less efficient in enriching the reservoir dynamics. In addition, we point out that the differences in the trends and saturation of NRMSE between single and multiple $\Delta\tau_p$ are unlikely to stem from initial conditions, as highlighted by the inset of Fig. 5C. For some trials with multiple $\Delta\tau_p$, although the reservoir begins with underperforming neurons (NRMSE > 0.7), its performance rapidly improves after adding the first several neurons. These results suggest that the quantity and diversity of neurons are both critical for constructing a reservoir with rich internal dynamics for time series prediction.

DISCUSSION

In summary, we have demonstrated a spintronic implementation of nonvolatile PRC, where the spintronic hardware can be controlled purely by electrical signals. We establish a strong isomorphism between the path dependence in the electrical manipulation of ferromagnetic domains and the history dependence of reservoir states. Such spatiotemporal transformation removes the constraints of

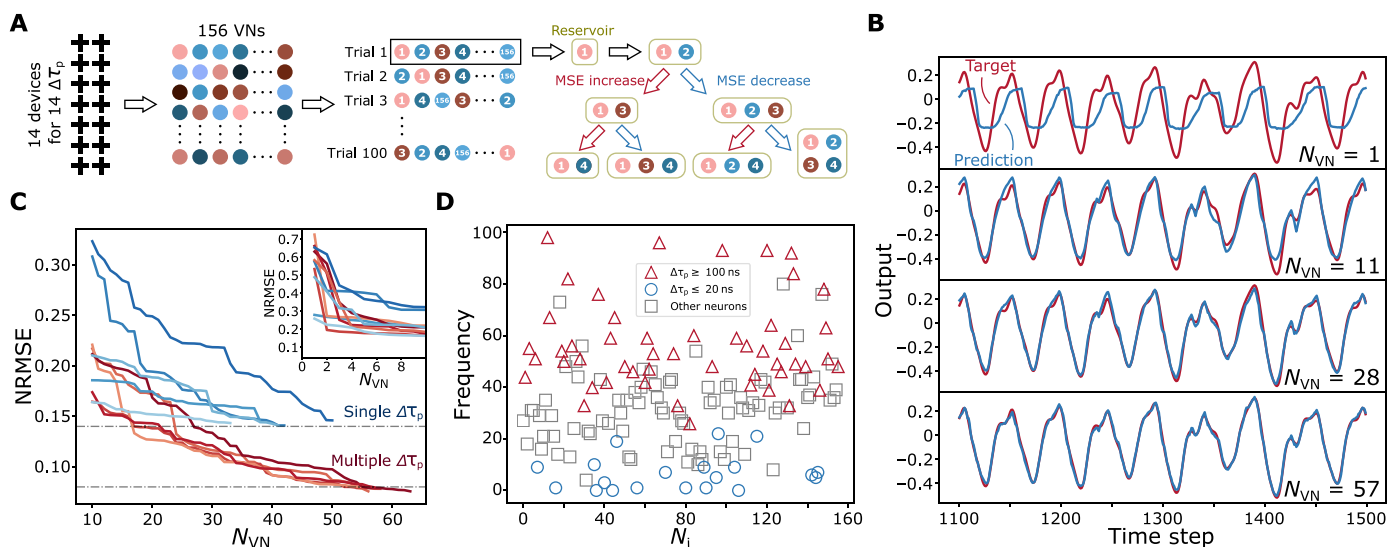


Fig. 5. Autonomous prediction of Mackey-Glass time series with spintronic physical RC. (A) Schematics of experimental design for studying neuron interaction. (B) Improvement of prediction with increasing number of neurons (N_{VN}). (C) Dependence of NRMSE on N_{VN} . Inset shows the initial condition. (D) The selection frequency of each neuron (N_i) in 100 trials.

input timescale on the efficient responses of physical reservoir, therefore substantially boosting the application scope of PRC. In principle, the time delay between successive inputs to our spintronic reservoir can extend beyond several years since nonvolatile spintronic devices are known to have long data retention time (>10 years) (54). With the help of a customized PCB, we use a parallel computing framework by generating a large number of VNs from multiple spintronic devices. We show that the shared τ_p dependence of multidimensional domain wall motion mediates the interaction of VNs, contributing to the rich dynamics of the spintronic reservoir and its excellent performance on benchmark tasks.

Using our spintronic reservoir, the highest single-trial A_{Test} is 0.903 in the written digit classification, and the lowest NRMSE is 0.076 in Mackey-Glass time series prediction, both of which are among the best performances of PRC (table S2). Furthermore, the robust performance of our spintronic RC is enhanced by ultralow reading energy (10.1 fJ) and fairly low writing energy (0.92 nJ). The present writing energy is further reduced to 37 pJ using nanoscale magnetic tunnel junction (MTJ). Our circuit has a low power budget of 16.64 mW and a high processing speed of 100 kHz, both are promising when compared with prior works. The total time per trial for written digit recognition and Mackey-Glass prediction tasks are 18 s and 90 ms, respectively (note S4 and tables S3 to S5). In the digit classification task, the energy consumption of our PRC network is estimated to be 99.6 and 94.8% lower than the software implementation of the state-of-the-art echo state network (7) and feedforward neural network (41), respectively (note S4 and table S6).

While the present work only offers a proof-of-concept solution to nonvolatile PRC, we envisage marked improvement in its functionality by incorporating more sophisticated materials, devices, and circuit components. First, strictly speaking, our architecture is partially parallel. Using a larger hardware array and a one-VN-one-memristor structure, a completely parallel architecture can be readily engineered. Second, replacing the Hall geometry with MTJs will enhance the signal-to-noise ratio, energy efficiency, and scalability (55). Notably, smaller MTJs also retain the nonlinear τ_p dependence (fig. S8) (56) and nonvolatility required for PRC. For an MTJ array on the same current channel (57), the collective resistance readout depends on the switching probability, which is finely controlled by V_p and τ_p . With the help of a MTJ crossbar array as the hardware readout layer (58), an all-spintronic PNN can be built using our approach. Next, the domain wall motion in our device is chiral (40), which provides additional tuning knobs on reservoir. A related technology is the magnetic skyrmion (59, 60), whose chirality and quasiparticle nature are likely to enrich reservoir dynamics. Last, our nonvolatile approach applies to both physical computation and physical transformation. Therefore, it can extend beyond the relatively rigid training paradigm of RC to the more hierarchical learning in deep neural networks or the so-called physics-aware training (3). We also expect our approach to be applicable to other nonvolatile memristors, such as ferroelectrics (61).

MATERIALS AND METHODS

Sample fabrication

The spintronic memristor has a stack of PtMn (15)/CoFeB (3)/W (1.5)/CoFeB (0.9)/MgO (1.8), where numbers in the parentheses are nominal thickness in nanometers. All layers were deposited at room

temperature using a magnetron sputtering technique on a 200-mm thermally oxidized silicon wafer, with a base chamber pressure of 2×10^{-8} torr. The field-free feature was enabled by an in-plane exchange bias field, which was engineered by annealing the device at 250°C and an in-plane field of 2 T for 2 hours followed by field cooling to room temperature. The microdevices were patterned by photolithography and ion beam etching. The Hall bars for electrical measurement has a dimension of 5 μm for both the current channel and Hall leads. The microstrips for MOKE microscopy are 5 μm by 25 μm .

Single-device measurement

A Tektronix AFG3252 function generator was used to apply voltage pulses to Hall bars and microstrips. The Hall voltage was measured using an ac modulation technique by a Zurich MFLI lock-in amplifier with a probing current of 50 μA and a modulation frequency of 317.3 Hz. Polar MOKE images were obtained using a MagVision Kerr system. To ensure the same original state before pulsing, a two-step initialization procedure was used, including a preset out-of-plane field (H_{pre}) and a preset voltage pulse (V_{pre}). For Hall bars, $\mu_0 H_{\text{pre}} = 0.05$ T and $V_{\text{pre}} = 2.5$ V at $\tau_p = 200$ ns. For microstrips, $\mu_0 H_{\text{pre}} = -0.05$ T and $V_{\text{pre}} = 3$ V at $\tau_p = 5$ μs .

Circuit and software

The PCB for hardware network consists of two chip carriers and peripheral circuits. Two dies—each provides seven Hall bar devices—from the same 200-mm wafer were installed on the chip carriers, and electrical connections were made by wire bonding. An FPGA was configured with a digital-to-analog converter for sourcing voltage pulses. Twenty-eight 1:2 analog demultiplexers were used to tune the pulse widths for writing the 14 devices in parallel. The measurement circuit has four 8:1 multiplexers, a two-stage amplification module, and an analog-to-digital converter (ADC). The FPGA read through the ADC Hall voltages induced by a probing current of 50 μA . The PCB was interfaced with a computer using USB. Python (version 3.7.11) was used to communicate with the FPGA for data processing and to run training algorithms of the two benchmark tasks. The FPGA was programmed using Quartus II (version 13.1).

Tasks and training

The training algorithms in this work are adapted from the classic echo state network. The reservoir state at time step t is expressed as

$$x_{t+1} = (1 - \alpha)x_t + \alpha \tanh(W_{\text{in}} u_{t+1} + W x_t) \quad (1)$$

Here, u_t is the input; W_{in} is the weight matrix connecting input and reservoir states; W is the internal weight matrix for connection among reservoir state; α is the leak rate, and the hyperbolic tangent is the activation function. Stacking x_t for all time steps generates the reservoir state matrix X , which relates to the real output y (or predicted output y') as

$$W_{\text{Out}} X = y \quad (2)$$

where W_{Out} is the weight matrix connecting reservoir state and output. To find W_{Out} , we first alter Eq. 2 as $X^T W_{\text{Out}}^T = y^T$, where $()^T$ indicates transpose matrix. Then

$$W_{\text{Out}}^T = (X X^T)^{-1} X y^T \quad (3)$$

where $(XX^T)^{-1}X$ is the Moore-Penrose pseudoinverse of X^T . In practice, we use y from the training dataset and apply `linalg.solve()` function from the numpy library of Python on

$$(XX^T + \text{Reg} \cdot I)W_{\text{Out}}^T = Xy^T \quad (4)$$

to solve for W_{Out} , where $\text{Reg} (=10^{-8})$ and I are regularization coefficients of ridge regression and identity matrix, respectively. The computed W_{Out} and Eq. 2 are used to find out y' for the testing dataset.

The MNIST written digit dataset has 60,000 samples for training and 10,000 samples for testing, all of which are labeled and have a resolution of 28×28 pixels. Since we exercise cross-validation, the difference between training and testing samples in the original dataset is trivial. Therefore, we use the first 20,000 samples in the training dataset. Equation 1 is irrelevant for building the physical reservoir of both tasks. For written digit, X was constructed using the process described in Fig. 3A. Ideally, X is 3D with a size of $N_{\text{VN}} \times N_{\text{time}} \times N_s$, where N_{time} is the number of time step (equals to the number of pixels per image). However, to reduce computation cost and time, we downsized X to 2D by concatenating the N_{VN} and N_{time} axes, and X becomes $(N_{\text{VN}} \times N_{\text{time}}) \times N_s$. For the same reasons, we downgrade the image resolution from 28×28 to 14×14 ($N_{\text{time}} = 196$). Since there are 10 distinct digits, W_{Out} is $10 \times (N_{\text{VN}} \times N_{\text{time}})$. We used a winner-take-all strategy (similar to the softmax function) to identify the predicted digit. Last, A_{Train} and A_{Test} were computed as $\frac{N_R}{N_s}$, where N_R is the number of correctly recognized digits.

The Mackey-Glass dataset has 10,000 data points with a delay of 17, i.e., it is chaotic. The X was constructed as depicted in Fig. 5A. Since $N_s = 1$ for this task, X has a size of $N_{\text{VN}} \times N_{\text{time}}$ for training, and W_{Out} has a size of $1 \times N_{\text{VN}}$. In the autonomous prediction mode, y'_t predicted by W_{Out} is fed back into the series as the input of the next time step u_{t+1} . The NRMSE of prediction is computed using

$$\text{NRMSE} = \sqrt{\frac{\frac{1}{N} \sum_{i=1}^{i=N} (y_i - y'_i)^2}{\frac{1}{N} \sum_{i=1}^{i=N} (y_i - \bar{y})^2}} \quad (5)$$

where \bar{y} is the average of real output in the testing dataset.

Supplementary Materials

The PDF file includes:

Supplementary Notes S1 to S4
Figs. S1 to S8
Tables S1 to S6
Legends for movies S1 to S3
References

Other Supplementary Material for this manuscript includes the following:

Movies S1 to S3

REFERENCES AND NOTES

- X. Liang, J. Tang, Y. Zhong, B. Gao, H. Qian, H. Wu, Physical reservoir computing with emerging electronics. *Nat. Electron.* **7**, 193–206 (2024).
- G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning. *Nat. Rev. Phys.* **3**, 422–440 (2021).
- L. G. Wright, T. Onodera, M. M. Stein, T. Wang, D. T. Schachter, Z. Hu, P. L. McMahon, Deep physical neural networks trained with backpropagation. *Nature* **601**, 549–555 (2022).
- H. Jaeger, B. Noheda, W. G. van der Wiel, Toward a formal theory for computing machines made out of whatever physics offers. *Nat. Commun.* **14**, 4911 (2023).
- M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707 (2019).
- G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, A. Hirose, Recent advances in physical reservoir computing: A review. *Neural Netw.* **115**, 100–123 (2019).
- H. Jaeger, The “echo state” approach to analysing and training recurrent neural networks—with an erratum note. *German National Research Center for Information Technology* **148**, 13 (2001).
- W. Maass, T. Natschlager, H. Markram, Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Comput.* **14**, 2531–2560 (2002).
- M. Lukoševičius, H. Jaeger, Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **3**, 127–149 (2009).
- M. Yan, C. Huang, P. Bienstman, P. Tino, W. Lin, J. Sun, Emerging opportunities and challenges for the future of reservoir computing. *Nat. Commun.* **15**, 2056 (2024).
- K. Nakajima, I. Fischer, *Reservoir Computing Theory, Physical Implementations, and Applications* (Springer Singapore, 2021); <https://doi.org/10.1007/978-981-13-1687-6>.
- K. Vandoorne, W. Dierckx, B. Schrauwen, D. Verstraeten, R. Baets, P. Bienstman, J. V. Campenhout, Toward optical signal processing using photonic reservoir computing. *Opt. Express* **16**, 11182–11192 (2008).
- D. Brunner, M. C. Soriano, G. V. D. Sande, *Photonic Reservoir Computing: Optical Recurrent Neural Networks* (De Gruyter, 2019).
- Y. Paquot, F. Dupont, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, S. Massar, Optoelectronic reservoir computing. *Sci. Rep.* **2**, 287 (2012).
- L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, I. Fischer, Photonic information processing beyond Turing: An optoelectronic implementation of reservoir computing. *Opt. Express* **20**, 3241–3249 (2012).
- F. Dupont, B. Schneider, A. Smerieri, M. Haelterman, S. Massar, All-optical reservoir computing. *Opt. Express* **20**, 22783–22795 (2012).
- K. Tanaka, Y. Tokudome, Y. Minami, S. Honda, T. Nakajima, K. Takei, K. Nakajima, Self-organization of remote reservoirs: Transferring computation to spatially distant locations. *Adv. Intell. Syst.* **4**, 2100166 (2022).
- P. Bhoval, S. Li, Physical reservoir computing with origami and its application to robotic crawling. *Sci. Rep.* **11**, 13002 (2021).
- K. Nakajima, H. Hauser, T. Li, R. Pfeifer, Information processing via physical soft body. *Sci. Rep.* **5**, 10487 (2015).
- K. Caluwaerts, J. Despraz, A. İçen, A. P. Sabelhaus, J. Bruce, B. Schrauwen, V. SunSpiral, Design and control of compliant tensegrity robots through simulation and hardware validation. *J. R. Soc. Interface* **11**, 20140520 (2014).
- K. Nakajima, H. Hauser, R. Kang, E. Guglielmino, D. Caldwell, R. Pfeifer, A soft body as a reservoir: Case studies in a dynamic model of octopus-inspired soft robotic arm. *Front. Comput. Neurosci.* **7**, 91 (2013).
- C. Du, F. Cai, M. A. Zidan, W. Ma, S. H. Lee, W. D. Lu, Reservoir computing using dynamic memristors for temporal information processing. *Nat. Commun.* **8**, 2204 (2017).
- L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, I. Fischer, Information processing using a single dynamical node as complex system. *Nat. Commun.* **2**, 468 (2011).
- Y. Zhong, J. Tang, X. Li, X. Liang, Z. Liu, Y. Li, Y. Xi, P. Yao, Z. Hao, B. Gao, H. Qian, H. Wu, A memristor-based analogue reservoir computing system for real-time and power-efficient signal processing. *Nat. Electron.* **5**, 672–681 (2022).
- Z. Chen, W. Li, Z. Fan, S. Dong, Y. Chen, M. Qin, M. Zeng, X. Lu, G. Zhou, X. Gao, J.-M. Liu, All-ferroelectric implementation of reservoir computing. *Nat. Commun.* **14**, 3585 (2023).
- 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, J. Grollier, Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017).
- Q. H. Tran, K. Nakajima, Learning temporal quantum tomography. *Phys. Rev. Lett.* **127**, 260401 (2021).
- R. Martínez-Peña, G. L. Giorgi, J. Nokkala, M. C. Soriano, R. Zambrini, Dynamical phase transitions in quantum reservoir computing. *Phys. Rev. Lett.* **127**, 100502 (2021).
- S. Ghosh, T. Paterek, T. C. H. Liew, Quantum neuromorphic platform for quantum state preparation. *Phys. Rev. Lett.* **123**, 260404 (2019).
- K. Fujii, K. Nakajima, Harnessing disordered-ensemble quantum dynamics for machine learning. *Phys. Rev. Applied* **8**, 024030 (2017).
- S. Ghosh, K. Nakajima, T. Krisnanda, K. Fujii, T. C. H. Liew, Quantum neuromorphic computing with reservoir computing networks. *Adv. Quantum Technol.* **4**, 2100053 (2021).
- A. Goudarzi, M. R. Lakin, D. Stefanovic, “DNA reservoir computing: A novel molecular computing approach” in *DNA Computing and Molecular Programming* (Springer, 2013), pp. 76–89. https://doi.org/10.1007/978-3-319-01928-4_6.

33. O. Obst, A. Trinchì, S. G. Hardin, M. Chadwick, I. Cole, T. H. Muster, N. Hoschke, D. Ostry, D. Price, K. N. Pham, T. Wark, Nano-scale reservoir computing. *Nano Commun. Netw.* **4**, 189–196 (2013).
34. X. Liu, K. K. Parhi, Reservoir computing using DNA oscillators. *ACS Synth. Biol.* **11**, 780–787 (2022).
35. T. Matsuo, D. Sato, S.-G. Koh, H. Shima, Y. Naitoh, H. Akinaga, T. Itoh, T. Nokami, M. Kobayashi, K. Kinoshita, Dynamic nonlinear behavior of ionic liquid-based reservoir computing devices. *ACS Appl. Mater. Interfaces* **14**, 36890–36901 (2022).
36. B. Jones, D. Stekel, J. Rowe, C. Fernando, Is there a Liquid State Machine in the Bacterium *Escherichia Coli*? in *2007 IEEE Symposium on Artificial Life* (IEEE, 2007), pp. 187–191.
37. G. Dion, S. Mejaouri, J. Sylvestre, Reservoir computing with a single delay-coupled non-linear mechanical oscillator. *J. Appl. Phys.* **124**, 152132 (2018).
38. D. Brunner, M. C. Soriano, C. R. Mirasso, I. Fischer, Parallel photonic information processing at gigabyte per second data rates using transient states. *Nat. Commun.* **4**, 1364 (2013).
39. J. Moon, W. Ma, J. H. Shin, F. Cai, C. Du, S. H. Lee, W. D. Lu, Temporal data classification and forecasting using a memristor-based reservoir computing system. *Nat. Electron.* **2**, 480–487 (2019).
40. J. Zhou, L. Huang, H. J. Chung, J. Huang, T. S. Suraj, D. J. X. Lin, J. Qiu, S. Chen, S. L. K. Yap, Y. T. Toh, S. K. Ng, H. K. Tan, A. Soumyanarayanan, S. T. Lim, Chiral interlayer exchange coupling for asymmetric domain wall propagation in field-free magnetization switching. *ACS Nano* **17**, 9049–9058 (2023).
41. J. Zhou, T. Zhao, X. Shu, L. Liu, W. Lin, S. Chen, S. Shi, X. Yan, X. Liu, J. Chen, Spin-orbit torque-induced domain nucleation for neuromorphic computing. *Adv. Mater.* **33**, 2103672 (2021).
42. S. Fukami, C. Zhang, S. DuttaGupta, A. Kurenkov, H. Ohno, Magnetization switching by spin-orbit torque in an antiferromagnet–ferromagnet bilayer system. *Nat. Mater.* **15**, 535–541 (2016).
43. S. Chen, R. Mishra, H. Chen, H. Yang, X. Qiu, Mimicking synaptic plasticity with a wedged Pt/Co/Pt spin-orbit torque device. *J. Phys. D Appl. Phys.* **55**, 095001 (2021).
44. K. Garello, C. O. Avci, I. M. Miron, M. Baumgartner, A. Ghosh, S. Auffret, O. Boulle, G. Gaudin, P. Gambardella, Ultrafast magnetization switching by spin-orbit torques. *Appl. Phys. Lett.* **105**, 212402 (2014).
45. W.-B. Liao, T.-Y. Chen, Y.-C. Hsiao, C.-F. Pai, Pulse-width and temperature dependence of memristive spin-orbit torque switching. *arXiv:2012.05531* (2020).
46. N. Sato, F. Xue, R. M. White, C. Bi, S. X. Wang, Two-terminal spin-orbit torque magnetoresistive random access memory. *Nat. Electron.* **1**, 508–511 (2018).
47. G. V. der Sande, D. Brunner, M. C. Soriano, Advances in photonic reservoir computing. *Nanophotonics* **6**, 561–576 (2017).
48. J. D. Kendall, S. Kumar, The building blocks of a brain-inspired computer. *Appl. Phys. Rev.* **7**, 011305 (2020).
49. 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, Central role of domain wall depinning for perpendicular magnetization switching driven by spin torque from the spin Hall effect. *Phys. Rev. B* **89**, 024418 (2014).
50. Y. LeCun, C. Cortes, C. J. C. Burges, The MNIST database of handwritten digits (1998); <https://yann.lecun.com/exdb/mnist/>.
51. M. Lukoševičius, Simple Echo State Network implementations (2016); https://mantas.info/code/simple_esn/.
52. L. Glass, M. Mackey, Mackey-Glass equation. *Scholarpedia* **5**, 6908 (2010).
53. T. Devolder, D. Rontani, S. Petit-Watelot, K. Bouzehouane, S. Andrieu, J. Létang, M.-W. Yoo, J.-P. Adam, C. Chappert, S. Giron, V. Cros, M. Sciamanna, J.-V. Kim, Chaos in magnetic nanocontact vortex oscillators. *Phys. Rev. Lett.* **123**, 147701 (2019).
54. V. Krizakova, M. Perumkunnil, S. Couet, P. Gambardella, K. Garello, Spin-orbit torque switching of magnetic tunnel junctions for memory applications. *J. Magn. Magn. Mater.* **562**, 169692 (2022).
55. S. S. P. Parkin, C. Kaiser, A. Panchula, P. M. Rice, B. Hughes, M. Samant, S.-H. Yang, Giant tunnelling magnetoresistance at room temperature with MgO (100) tunnel barriers. *Nat. Mater.* **3**, 862–867 (2004).
56. J. Zhou, L. Huang, S. L. K. Yap, D. J. X. Lin, B. Chen, S. Chen, S. K. Wong, J. Qiu, J. Lourembam, A. Soumyanarayanan, S. T. Lim, Synergizing intrinsic symmetry breaking with spin-orbit torques for field-free perpendicular magnetic tunnel junction. *APL Mater.* **12**, 081105 (2024).
57. A. Kumar, D. J. X. Lin, D. Das, L. Huang, S. L. K. Yap, H. R. Tan, H. K. Tan, R. J. J. Lim, Y. T. Toh, S. Chen, S. T. Lim, X. Fong, P. Ho, Multistate compound magnetic tunnel junction synapses for digital recognition. *ACS Appl. Mater. Interfaces* **16**, 10335–10343 (2024).
58. L. Liu, D. Wang, D. Wang, Y. Sun, H. Lin, X. Gong, Y. Zhang, R. Tang, Z. Mai, Z. Hou, Y. Yang, P. Li, L. Wang, Q. Luo, L. Li, G. Xing, M. Liu, Domain wall magnetic tunnel junction-based artificial synapses and neurons for all-spin neuromorphic hardware. *Nat. Commun.* **15**, 4534 (2024).
59. S. Chen, J. Lourembam, P. Ho, A. K. J. Toh, J. Huang, X. Chen, H. K. Tan, S. L. K. Yap, R. J. J. Lim, H. R. Tan, T. S. Suraj, M. I. Sim, Y. T. Toh, I. Lim, N. C. B. Lim, J. Zhou, H. J. Chung, S. T. Lim, A. Soumyanarayanan, All-electrical skyrmionic magnetic tunnel junction. *Nature* **627**, 522–527 (2024).
60. D. Pinna, F. Abreu Araujo, J. V. Kim, V. Cros, D. Querlioz, P. Bessiere, J. Droulez, J. Grollier, Skyrmion gas manipulation for probabilistic computing. *Phys. Rev. Applied* **9**, 064018 (2018).
61. K. Everschor-Sitte, A. Majumdar, K. Wolk, D. Meier, Topological magnetic and ferroelectric systems for reservoir computing. *Nat. Rev. Phys.* **6**, 455–462 (2024).
62. M. L. Alomar, E. S. Skibinsky-Gitlin, C. F. Frasser, V. Canals, E. Isern, M. Roca, J. L. Rosselló, Efficient parallel implementation of reservoir computing systems. *Neural Comput. Appl.* **32**, 2299–2313 (2020).
63. M. L. Alomar, M. C. Soriano, M. Escalona-Morán, V. Canals, I. Fischer, C. R. Mirasso, J. L. Rosselló, Digital implementation of a single dynamical node reservoir computer. *IEEE Trans. Circuits Syst. II: Express Briefs* **62**, 977–981 (2015).
64. M. L. Alomar, V. Canals, N. Perez-Mora, V. Martínez-Moll, J. L. Rosselló, FPGA-based stochastic echo state networks for time-series forecasting. *Comput. Intell. Neurosci.* **2016**, 3917892 (2016).
65. Z. Dai, F. Xiang, C. He, Z. Wang, W. Zhang, Y. Li, J. Yue, D. Shang, A scalable small-footprint time-space-pipelined architecture for reservoir computing. *IEEE Trans. Circuits Syst. II: Express Briefs* **70**, 3069–3073 (2023).
66. X. Liang, Y. Zhong, J. Tang, Z. Liu, P. Yao, K. Sun, Q. Zhang, B. Gao, H. Heidari, H. Qian, H. Wu, Rotating neurons for all-analog implementation of cyclic reservoir computing. *Nat. Commun.* **13**, 1549 (2022).
67. G. Milano, G. Pedretti, K. Montano, S. Ricci, S. Hashemkhani, L. Boarino, D. Ielmini, C. Ricciardi, In materia reservoir computing with a fully memristive architecture based on self-organizing nanowire networks. *Nat. Mater.* **21**, 195–202 (2022).
68. J. C. Gartside, K. D. Stenning, A. Vanstone, H. H. Holder, D. M. Arroo, T. Dion, F. Caravelli, H. Kurebayashi, W. R. Branford, Reconfigurable training and reservoir computing in an artificial spin-vortex ice via spin-wave fingerprinting. *Nat. Nanotechnol.* **17**, 460–469 (2022).
69. W. Ma, T. Hennen, M. Lueker-Boden, R. Galbraith, J. Goode, W. H. Choi, P. F. Chiu, J. A. J. Rupp, D. J. Wouters, R. Waser, D. Bedau, A Mott Insulator-Based Oscillator Circuit for Reservoir Computing in *2020 IEEE International Symposium on Circuits and Systems (ISCAS)* (IEEE, 2020), pp. 1–5.
70. W. M. Zhong, C. L. Luo, X. G. Tang, X. B. Lu, J. Y. Dai, Dynamic FET-based memristor with relaxor antiferroelectric HfO₂ gate dielectric for fast reservoir computing. *Mater. Today Nano* **23**, 100357 (2023).
71. S. Wang, Y. Li, D. Wang, W. Zhang, X. Chen, D. Dong, S. Wang, X. Zhang, P. Lin, C. Gallicchio, X. Xu, Q. Liu, K.-T. Cheng, Z. Wang, D. Shang, M. Liu, Echo state graph neural networks with analogue random resistive memory arrays. *Nat. Mach. Intell.* **5**, 104–113 (2023).
72. Y. Wang, Q. Wang, S. Shi, X. He, Z. Tang, K. Zhao, X. Chu, Benchmarking the Performance and Energy Efficiency of AI Accelerators for AI Training in *2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID)* (IEEE, 2020), pp. 744–751.

Acknowledgments: We wish to thank W. Zizhe for insightful discussion on reservoir dynamics and training methodology. **Funding:** This study is supported by the career development fund (A*STAR grant no. C222812013), Key Projects Supported by the Regional Innovation and Development Joint Fund (grant no. U23A20365), National Key R&D Plan “Nano Frontier” Key Special Project (grant nos. 2024YFA1208400 and 2021YFA1200502), Cultivation Projects of National Major R&D Project (grant no. 92164109), Disruptive Technology Innovation Project of the National Key R&D Program (grant no. DT01202402075), National Natural Science Foundation of China (grant no. 61874158), Special Project of Strategic Leading Science and Technology of Chinese Academy of Sciences (grant no. XDB44000000-7), Yanzhao Young Scientist Project of Hebei Province (grant no. F2023201076), Support Program for the Top Young Talents of Hebei Province (grant no. 70280011807), Interdisciplinary Research Program of Natural Science of Hebei University (grant no. DXK202101), Natural Science Foundation of Hebei Province (grant nos. F2021201045), Baoding Science and Technology Plan Project (grant nos. 2172P011, 2272P014) Outstanding Young Scientific Research and Innovation Team of Hebei University (grant no. 605020521001), Special Support Funds for National High Level Talents (grant no. 041500120001), Hebei Province High-Level Tal-Ent Funding Project (grant no. B20231003), Science and Technology Project of Hebei Education Department (grant nos. QN2020178 and QN2021026). **Author contributions:** J.Z. conceptualized this work. J.Z., X.Y., and S.T.L. supplied resources and supervised this work. J.Z. performed the electrical and MOKE measurements, composed Python codes for training, and wrote the original manuscript. J.X. designed and fabricated the PCB circuits. L.H. fabricated devices and performed MOKE measurement. S.L.K.Y. deposited films and fabricated devices. S.C. performed electrical measurement and supported algorithm development. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 4 July 2024
Accepted 9 December 2024
Published 10 January 2025
10.1126/sciadv.adr5262