Toward All-Electric Non-volatile Intelligence in Spintronic Reservoir

Jing Zhou[1], Jikang Xu[2], Lisen Huang[1], Sherry Lee Koon Yap[1], Shaohai Chen[1], Xiaobing Yan[2], *Senior Member, IEEE,* and Sze Ter Lim[1]

[1]Institute of Materials Research and Engineering, Agency for Science, Technology and Research (A*STAR), 2 Fusionopolis Way, Innovis #08-03, Singapore 138634, Republic of Singapore, zhou_jing@imre.a-star.edu.sg; lim_sze_ter@imre.a-star.edu.sg

[2]Key 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, yanxiaobing@ime.ac.cn

Integrating physical dynamics with computational models is gaining traction for boosting neural network efficiency. Physical reservoir computing leverages the intrinsic dynamics of materials for temporal processing but faces challenges in constructing efficient reservoirs. Here, we move beyond delay-based designs by harnessing spatiotemporal transformations in all-electric, non-volatile spintronic devices. By triggering devices with varied pulse widths, we emulate neurons and build a compact reservoir with strong non-linearity and dense interconnectivity. Despite using only 14 physical nodes, our system achieves a 0.903 recognition rate on handwritten digits and a 0.076 error rate in Mackey-Glass prediction, validated on a proof-of-concept PCB.

***Index Terms*—Reservoir Computing, Nonvolatile Memory, Hall Effect Devices, Spin Valves.**

I. Introduction

The recurrent neural network (RNN) is a special framework of artificial intelligence (AI) designed for temporal data processing, which has broad applications in physics, biology and finance. RNN suffers from the haunting problem of exploding (or vanishing) gradient during training, to which the conventional approaches, such as long short-term memory and backpropagation through time, remain inefficient. Reservoir computing (RC) emerges as an improved RNN paradigm[1]. It is featured by an input-driven reservoir of high-dimensional data space and a memoryless readout layer. Only the readout layer is trained, making RC particularly suitable for the edge AI, where resources are constrained. Recent advances leverage physical processes to implement RC with analog responses of hardware – known as physical RC (PRC) – including mechanical, photonic and electronic systems[2, 3]. However, conventional PRC approaches rely on delay-based volatile reservoirs, enforcing a pulse-train-like input stream with a narrow bandwidth, limiting AI task performance. Here, for the first time, we provide an approach to construct non-volatile PRC using all-electric spintronic devices, completely separating input timescale from task timescale[4]. Our physical reservoir is realized on a customized printed circuit board (PCB) to mimic on-chip training. It exhibits excellent performance in benchmark classification and prediction tasks while lowering energy consumption by 99.6% compared to software implementation.

II. Methods and Approaches

A. Fading Memoroy

A mathematical precondition of reservoir computing is fading memory, which requires the future readout of reservoir to depend on past inputs and this dependence weakens over time. We demonstrate this property with Hall bars made of perpendicularly magnetized ferromagnetic layer, which can be switched by SOT without external magnetic field[5]. Fig. 1(a) shows SOT-induced Hall resistance (R_H) changes when 6 voltage pulse trains are applied. Note that only the third pulse (N_p = 3) has a different pulse amplitude (V_p) and the remaining V_p are the same for each pulse across the 6 trains. Clearly, the R_H response begins to diverge at N_p = 3 and eventually converge at N_p = 7, thus phenomenologically showing fading memory. We attribute this behaviour to the path dependence of domain evolution in magnetic hysteresis. That is, as long as the magnetization is not fully switched, its future trajectory depends on both the history of its past states and the sequence of electrical inputs.

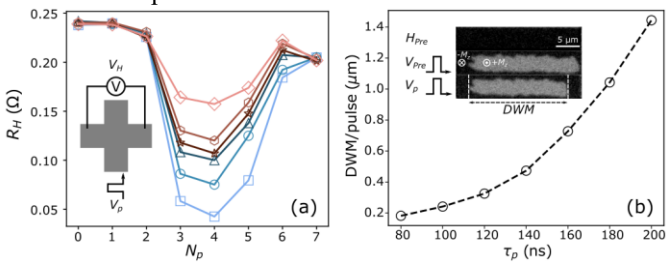

Fig. 1. Fading memory and non-linearity. (a) Variation of Hall resistance by 6 pulse trains with a different pulse amplitude only at the 3rd pulse. (b) Dependence of domain wall motion (DWM) per pulse on pulse width (τ_p).

B. Non-linearity

A working reservoir also requires non-linear dynamics. We study the SOT-driven domain evolution in a 5-μm microstrip using magneto-optic Kerr effect (MOKE) microscopy. The initial state is created using the same preset field (H_{pre}) and voltage (V_{pre}) for measurement. Then we count the total number of a fixed voltage pulse required to switch all $+M_z$ domains to $-M_z$. The extracted average domain wall motion (DWM) per pulse is shown in Fig. 1(b), which is unambiguously non-linear in different pulse widths (τ_p). Since R_H is proportional to $\pm M_z$, we expect R_H-τ_p to be non-linear as well.

C. Network and circuits

The physical reservoir is constructed in a three-step process as illustrated in Fig. 2(a). First, the input u_n is converted to a pulse train of fixed amplitude with the magnitude of u_n linearly mapped to τ_p. Second, the same u_n is mapped to different τ_p intervals – defined as the dynamic range (DR) – to generate distinct neurons. Finally, multiple neuron responses are stacked to generate a high-dimensional reservoir matrix X, followed by a standard readout using ridge regression (W_{out} is output weight matrix). We develop peripheral circuit on a customized PCB to support parallel processing. In Fig. 2(b), a total of 14 Hall bars

– as devices under test (DUT) – of highly homogeneous device properties are connected such that different neurons can be generated simultaneously, ensuring a high processing speed. The PCB is connected to a computer vias the USB port, and Python is used to run training algorithm and process data.

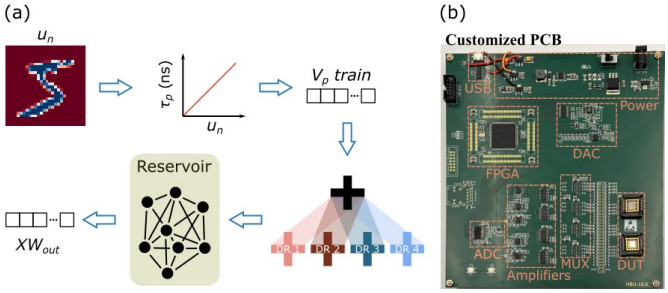

Fig. 2. Reservoir structure. (a) Construction of a physical reservoir from analog responses of Hall bars. (b) Customized PCB.

III. Performance

A. Classification

We first test our approach with a benchmark AI task – MNIST written digit recognition. We use the first 20,000 training samples from the database as our data pool then perform 10-fold cross validation. Fig. 3(a) shows the confusion matrix of a single trial with 20000 samples (N_s = 20k), achieving a high testing accuracy (A_{test}) of 0.903. Notably, our reservoir is making human-like mistakes since it confuses '5' with '3' and '9' with '4'. In Fig. 3(b), when we gradually increase the sample population, the training accuracy (A_{train}) decreases but A_{test} increases. This is because the reservoir can capture the features of a small dataset, but a sufficiently big dataset is required for generalization. The best performance achieved at N_s = 20000 are A_{train} = 0.916±0.001 and A_{test} = 0.880±0.001.

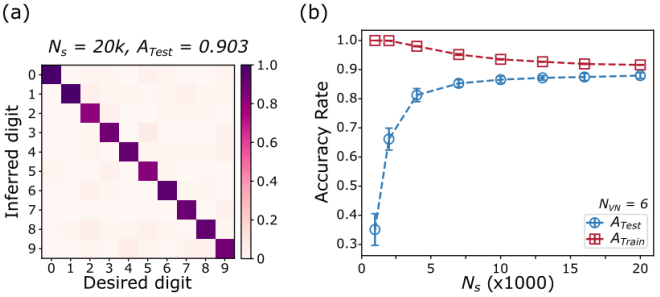

Fig. 3. Written digit classification. (a) Confusion matrix of a single trial. (b) Ttraining (A_{train}) and testing accuracy (A_{test}) from 10-fold cross validation.

B. Prediction

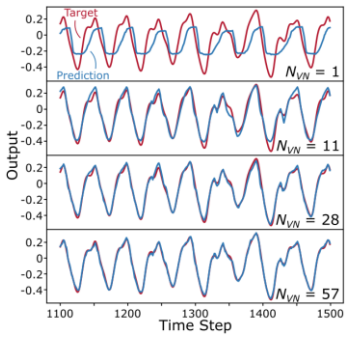

Fig. 4. Mackey-Glass chaotic time series prediction.

We also performed Mackey-Glass chaotic time series prediction, which is a specific benchmark task for RNN. We flush the first 100 data points and train the next 1000 data points to predict the 400 points after. Fig. 4 shows the effect of increasing the number of neurons (N_{VN}) on reservoir performance. When N_{VN} = 1, the prediction only captures the major frequency of the data, producing large discrepancies and phase difference. As N_{VN} increases, the discrepancies between prediction and target rapidly decreases, eventually producing highly superimposed traces at N_{VN} = 57. The normalized root mean square error achieved is 0.076, which is among the lowest.

C. Evaluation

We evaluate our physical reservoir by benchmarking its energy consumption against software implementation of RC in Table I. Our systems has a 99.6% and 94.8% of energy saving compared with a CPU implementation using state-of-the-art echo state network and feed forward neural network (FNN), respectively. We attribute this to the much less matrix multiplication in our system and operating energy of devices.

TABLE I. COMPARISON OF RC

Method	Network	Energy/sample	Energy/task
Spintronic memristor + CPU	RC	Device:1.16 μJ CPU: 0.212 mJ Total: 0.213 mJ	Device: 23.1 mJ CPU: 4.25 J Total: 4.27 J
CPU	RC	51.0 mJ	1.02 kJ
CPU	FNN	4.07 mJ	285 J

D. Scaling

We have successfully integrated the field-free strategy in our Hall bar with SOT-MTJ, which has a cell diameter of 240 nm, an endurance larger than 10^{12} and 100% switching probability at 10 ns[6]. The proposed domain-based reservoir construction strategy can be adapted to MTJ arrays where individual MTJ – being single-domain – switches on a probabilistic basis. The collective readout of MTJ arrays will be used in the place of R_H.

IV. Conclusion

We have demonstrated a non-volatile physical reservoir using non-linear dynamics of magnetic domains. The proposed method shows excellent AI task performance, and is scalable and adaptable to other nonvolatile memristors.

References

[1] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review,* vol. 3, no. 3, pp. 127-149, 2009.

[2] X. Liang, J. Tang, Y. Zhong, B. Gao, H. Qian, and H. Wu, "Physical reservoir computing with emerging electronics," *Nature Electronics,* vol. 7, no. 3, pp. 193-206, 2024.

[3] H. Jaeger, B. Noheda, and W. G. van der Wiel, "Toward a formal theory for computing machines made out of whatever physics offers," *Nature Communications,* vol. 14, no. 1, p. 4911, 2023.

[4] J. Zhou *et al.*, "Harnessing spatiotemporal transformation in magnetic domains for nonvolatile physical reservoir computing," *Science Advances,* vol. 11, no. 2, p. eadr5262, 2025.

[5] J. Zhou *et al.*, "Chiral Interlayer Exchange Coupling for Asymmetric Domain Wall Propagation in Field-Free Magnetization Switching," *ACS Nano,* vol. 17, no. 10, pp. 9049-9058, 2023.

[6] J. Zhou *et al.*, "Synergizing intrinsic symmetry breaking with spin–orbit torques for field-free perpendicular magnetic tunnel junction," *APL Materials,* vol. 12, no. 8, 2024.