Spatial analysis of physical reservoir computers

Jake Love,[1] Robin Msiska,[1,2] Jeroen Mulkers,[2] George Bourianoff,[3] Jonathan Leliaert,[2] and Karin Everschor-Sitte[1,*]

[1] *Faculty of Physics and Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Duisburg 47057, Germany*

[2] *DyNaMat, Department of Solid State Sciences, Ghent University, Ghent, Belgium*

[3] *Senior Principle Engineer, Intel Corp. (Retired)*

(Received 18 August 2023; accepted 26 September 2023; published 20 October 2023)

Physical reservoir computing is a computational framework that implements spatiotemporal information processing directly within physical systems. By exciting nonlinear dynamical systems and creating linear models from their state, we can create highly-energy-efficient devices capable of solving machine-learning tasks without building a modular system consisting of millions of neurons interconnected by synapses. To act as an effective reservoir, the chosen dynamical system must have two desirable properties: nonlinearity and memory. We present task-agnostic spatial measures to locally measure both of these properties and exemplify them for a specific physical reservoir based upon magnetic skyrmion textures. In contrast to typical reservoir-computing metrics, these metrics can be resolved spatially and in parallel from a single input signal, allowing for efficient parameter searches to design efficient and high-performance reservoirs. Additionally, we show the natural trade-off between memory capacity and nonlinearity in our reservoir's behavior, both locally and globally. Finally, by balancing the memory and nonlinearity in a reservoir, we can improve its performance for specific tasks.

DOI: 10.1103/PhysRevApplied.20.044057

I. INTRODUCTION

Reservoir computing (RC) is a unified supervised learning framework derived from recurrent neural networks (RNNs) that utilizes dynamical systems to perform computations [1,2]. A reservoir computer consists of two components: a fixed nonlinear dynamical system called a reservoir and a trainable linear readout layer. Input signals are fed into the reservoir, and the system's resulting transient states are observed. The observed states, referred to as readout values, are then fed through the linear readout layer, which can be trained to solve specific problems. By separating the complex nonlinear dynamics of the reservoir from the trainable readout layer, the connecting weights can be solved inexpensively using a least-squares model, often requiring far fewer training data than conventional RNNs [3]. Typically, reservoir computers are suited to pattern-recognition problems dealing with temporally or spatially correlated information, such as audio or image data [4].

The RC framework is not limited to mathematical models and can also be implemented with physical dynamical systems [5]. In physical RC, the input signal is fed into a physical substrate via time-dependent excitations. A finite representation of the system's transient states over the excitation is then observed via experimental methods, representing the readout values. Like the nonphysical case, the readout values are channeled into a trained linear model to perform the desired computation. By implementing reservoir computers physically, one can harness the inherent parallelism of physics and unlock opportunities for cost-effective energy-efficient computations that transcend the constraints of conventional CMOS hardware. Although one can use any physical dynamical system for reservoir computing, some are better than others. Solving complex tasks requires that the system has several fundamental properties: high dimensionality, nonlinearity, and fading memory [6]. A diverse set of physical dynamical systems based upon optics, spintronics [7], and even a bucket of water [8] have already been proposed to implement the reservoir computing paradigm, including magnetic [9] and magnetic skyrmion-based systems [10–12].

In RC, a balance must be achieved between the reservoir's ability to retain information from previous inputs over time (memory) and its ability to capture complex dynamic relationships between the inputs and outputs (nonlinearity). The trade-off between memory and nonlinearity is a critical and universal [13] factor in determining the computational capabilities of reservoir computers. Despite ample theoretical research [2,14,15], practical

*karin.everschor-sitte@uni-due.de

implications of these effects in physical RC implementations remain largely unexplored.

Existing literature provides global metrics to quantify the total memory capacity [16] and nonlinearity [6,14] of reservoir computers. While these global metrics are useful for the characterization of reservoir computers, they provide limited guidance in interpreting where the required properties, memory, and nonlinearity arise within a dynamical system.

This work introduces methods that measure the memory and nonlinearity locally for physical reservoirs with spatially correlated readouts. One can use these methods to understand which physical structures lead to memory and nonlinearity in the reservoir. By creating conditions that replicate such structures, one can parameterize future reservoirs in a directed manner to optimize a balance between nonlinearity and memory, as required by possible tasks. These metrics can also be applied to systematically determine optimal output contact-placement schemes for physical reservoirs, while minimizing the number of contacts needed. This is particularly relevant in practical applications, where only a portion of the extracted information is significant, such as in Internet-of-things scenarios, where transmitting all collected data to the cloud is not practical. Minimizing the number of contacts, without compromising performance, leads to an RC system with fewer parameters that capture the most significant features, thereby decreasing the chances of overfitting and enhancing generalization capabilities.

II. SPATIALLY RESOLVED METRICS

We introduce two spatially resolved metrics for memory and nonlinearity. Both metrics can be calculated in parallel from the same set of input and output data and are task agnostic, and hence, can be applied to various input sequences. This allows one to analyze how different input sequences, such as a sequence of uniform random numbers or series generated from a Mackey-Glass equation, couple to different structures in physical reservoirs.

A. Nonlinearity

As a reservoir is a nonlinear system with fading memory, its readout node values, y, can be modeled as a function of input to the system, u, using a Volterra series [17], such that

$$y(t) = h^{(0)} + \int_{\tau=0}^{\infty} h^{(1)}(\tau)u(t-\tau)d\tau + y^{\mathrm{NL}}(t). \quad (1)$$

Here, $h^{(0)}$ and $h^{(1)}$ are zero- and first-order Volterra kernels, respectively, and $y^{\mathrm{NL}}(t)$ represents truncated nonlinear terms of the Volterra series. For a linear system, the truncated terms of Eq. (1) are zero, and for a highly nonlinear system, these truncated terms dominate. To measure the nonlinearity of each readout node in the reservoir, we quantify how important these truncated terms are when modeling the individual readout as a Volterra series.

We first model each readout node, $y_n(t)$, in the reservoir as a discrete time-truncated Volterra series that can be estimated by the history of inputs, $u(t)$:

$$y_n^L(t) = h^{(0)} + \sum_{\tau=0}^{\infty} h_\tau^{(1)} u(t-\tau). \quad (2)$$

This model is then implemented as a linear estimator of the form

$$\hat{y}_n(t) = c + \sum_{\tau=0}^{k} w_\tau u(t-\tau), \quad (3)$$

where weights w and constant c are trained on the known input and output data, see Fig. 1(a). k is the time delay at which we truncate the Volterra kernel, which is picked to be longer than the relaxation time of the reservoir. The quality of the estimator in Eq. (3) is measured using the R^2 correlation coefficient (see Appendix A 1 for details) and can be considered a direct measure of how linear the response of each reservoir node is. Therefore, to act as a measure of nonlinearity, we present it as

$$\Phi_n^{\mathrm{NL}} = 1 - R^2[\hat{y}_n, y_n]. \quad (4)$$

The result is a measure of nonlinearity, NL_n, for each individual node of the reservoir bound by the interval $[0, 1]$, where zero indicates a linear relationship between the reservoir input and a readout node's value, and conversely, values close to one indicate a highly nonlinear relationship. The definition of nonlinearity proposed here is designed to be decoupled from the measure of memory in the system. Hence, it is possible to analyze systems with high memory and zero nonlinearity (a linear time-invariant filter) and vice versa (a high-order polynomial), and even systems with both low memory and low nonlinearity (physical pendulum with light displacement) or high memory and high nonlinearity (complex systems governed by nonlinear delay differential equations). This decoupling of memory and nonlinearity is in contrast to other methods, such as the quadratic memory capacity task seen in Refs. [18,19] and the n-bit parity task [20], that measure nonlinearity as a form of memory.

B. Memory

To spatially measure memory in our system, we use an adaption of the linear memory capacity metric [16]. In the original metric, a linear estimator is trained to recall the history of a reservoir's inputs, u, from all the readout values of all readout nodes. We use this metric to measure memory on a per-node basis by considering

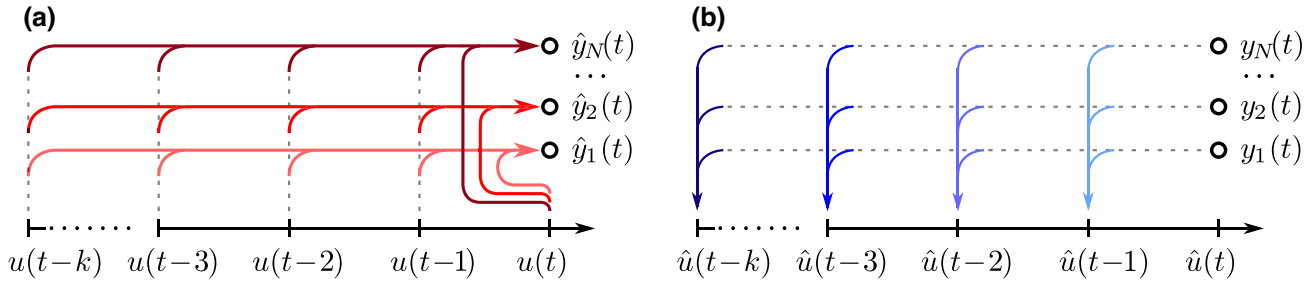


FIG. 1. Sketches illustrating the evaluation metrics of (a) nonlinearity, where multiple past inputs are used to predict the present value of individual readout nodes, and (b) memory capacity, wherein multiple readout nodes are employed to predict inputs at a specific time delay in the past.

the memory capacity of only the readout nodes within a threshold distance, in physical space, of the chosen node. The choice of threshold distance and the node-distance definition depends on the specific system under analysis.

For a chosen nth readout node, we create a vector, γ_n, of all other nodes of the reservoir that fall within a threshold distance, measured from n. We then use γ_n to create an estimator, $\hat{u}_n(t-\tau)$, that can recall the input at a time delay of τ:

$$\hat{u}_n(t-\tau) = c + \sum_i w_i \gamma_{n,i}(t), \tag{5}$$

see Fig. 1(b).

The memory capacity of the nodes in γ_n is then calculated by summing the correlation coefficient of all estimators up to a cutoff threshold, k, that is larger than the relaxation time of the reservoir:

$$\Phi_n^{\mathrm{MC}} = \sum_{\tau=1}^{k} R^2[\hat{u}_n(t-\tau), u(t-\tau)]. \tag{6}$$

C. Robustness

Robustness is an important attribute for physical reservoirs, forming the foundation for their reproducible operation. Following excitation by input signals, a robust reservoir effortlessly returns to a consistent (meta)stable state—a condition in which the reservoir state remains unaffected by minor perturbations. This ensures a consistent and distinctive response to a given input.

A convenient approach to characterize robustness in physical reservoir systems is to pass a random input signal into the reservoir. After the input signal dissipates, the reservoir is left to relax until it reaches a (meta)stable state. The robustness of the system is then determined by the difference between its input response state and its rest state, where a small (high) difference refers to high (low) robustness. The details of the difference measure can be adjusted to the specific system.

III. SPATIAL ANALYSIS OF MAGNETIC SKYRMION RESERVOIRS

To demonstrate the application of the above-introduced metrics, we apply them to a specific magnetic skyrmion-based implementation of physical reservoir computing [10,12]. Magnetic skyrmions are magnetic whirls with nontrivial topology. They occur, for example, in magnetic thin films where spins favor a canted alignment, often caused by the Dzyaloshinskii-Moriya interaction (DMI). Skyrmions host complex interactions with spin-polarized electric currents, defects in their host material, and other skyrmions, resulting in deformations with highly nonlinear behavior [11]. Skyrmions elastically deform back to their ground state upon removal of external forces, a condition that is necessary for fading memory.

In addition to the basic requirements, magnetic skyrmions are a competitive system for reservoir computing due to their nanoscale size, responsiveness to low-power excitations, and their ability to integrate with CMOS technology [21] and other potential magnetic devices [7,10,22,23].

For our specific model, we consider an inhomogeneous magnetic thin film with an interfacial DMI populated with magnetic skyrmions. We introduce grain inhomogeneities with varying magnetic anisotropy, resulting in a complex energy landscape. This causes specific locations on the film to be energetically favorable for skyrmions, referred to as pinning sites. As a result, the local energetic minimum state of the magnetic texture becomes an aperiodic arrangement of skyrmions, as shown in Fig. 2. The dynamics of the skyrmion reservoir are governed by the Landau-Lifshitz-Gilbert equation; see Appendix A 4 for more details.

We implement the reservoir input, $u(t)$, as a time-varying unidirectional electric current across the film characterized by the current density, $\mathbf{j}(t)$,

$$\mathbf{j}(t) = j_{\mathrm{in}} u(ft)\hat{\mathbf{x}}. \tag{7}$$

Here, $\hat{\mathbf{x}}$ is the current direction and $j_{\mathrm{in}}u(ft)$ is the time-dependent current strength. j_{in} provides control over input

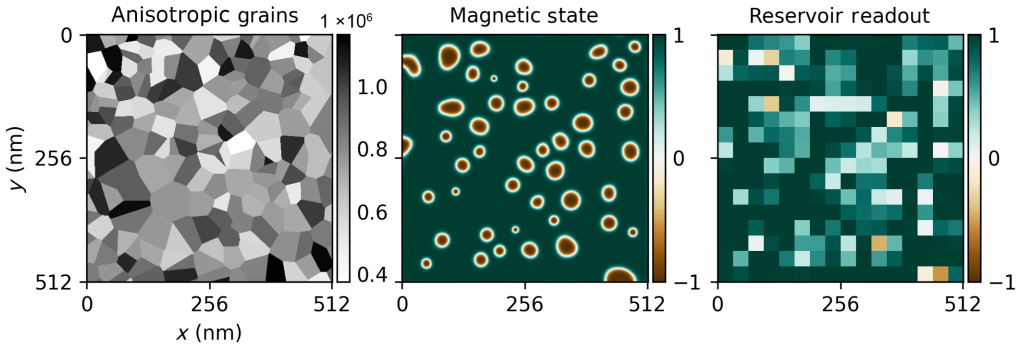
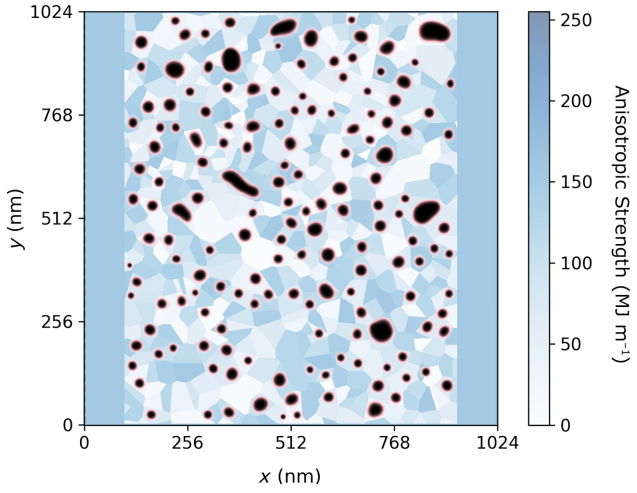


FIG. 2. Example distribution of grain inhomogeneities across a magnetic thin film. Left, each grain has a distinct magnetic anisotropy strength, as indicated by the level of grayscale color intensity. Middle, nonuniform anisotropy creates pinning sites for skyrmions, resulting in their aperiodic arrangements. Color code depicts the out-of-plane component of the local normalized magnetization. Right, for the readout, the out-of-plane projection of the magnetization field is taken and then discretized by taking its mean value over an area.

scaling, and factor f controls time scaling. Discrete-time inputs are fed into the reservoir as a step function, such that the reciprocal of f is the duration of each step. The electrical current interacts with magnetic skyrmions via spin torques, resulting in a time-dependent nonlinear deformation of the magnetic texture. Provided the current is within a certain threshold, the structure will relax back to its original state when the current is removed. See Video 1 for a demonstration of this process.

For readout nodes, we observe the changes in magnetic texture by taking a low-resolution snapshot of the film's out-of-plane magnetization component. We choose the resolution to be of similar magnitude to the sizes of individual skyrmions, as indicated in Fig. 2 (right panel).

For the memory metric, since our readout adopts a square-grid structure, we analyze the immediate Moore neighborhood at each node, resulting in a threshold value of ten. We implement the model in simulations using the MuMax3 software [24]; precise details on the simulation setup are given in Appendix A 2.



VIDEO 1. Animated visualization of the dynamic current-induced response of skyrmions in the reservoir, superimposed on the underlying magnetic anisotropy grain inhomogeneity pattern.

A. Results

We first analyze two distinct skyrmion reservoirs with different DMI strengths, 3.3×10^{-3} J m^{-2} and 3.6×10^{-3} J m^{-2}, but otherwise identical dynamics. These reservoirs can be visually distinguished by the sizes of the individual skyrmions present on the thin films, as seen in the magnetization shown in Fig. 3. Due to the effect of the DMI on the possible metastable states, the two reservoirs also have inherently different initial configurations.

Both reservoirs are tested against a discrete-time random input signal, $u(t)$. The signal has 1500 elements drawn from a random distribution with the bounds $[-1, +1]$. The input is coupled to the reservoirs with input scaling $j_{\text{in}} = 300 \times 10^9$ A m^{-2} and time scaling $f = 4$ GHz. The memory, nonlinearity, and robustness are spatially analyzed. The results are shown in Fig. 3. For the case of the skyrmion reservoir, we define robustness as the difference between the magnetic states before and after the random input signal has been applied. When applying the memory metric to this system, we set our threshold distance to 24, since we consider the next-nearest Moore neighborhood at each node on our readout square grid.

For both reservoirs, we can observe the well-studied trade-off between memory and nonlinearity, both spatially and globally. Locally, regions of high nonlinearity correspond closely to regions of low memory, as seen predominantly in Fig. 3(a) and less so in Fig. 3(b). However, the metrics are not entirely orthogonal, as regions with both memory and nonlinearity do coexist. Physically, these regions of high nonlinearity appear where changes to the magnetic texture are greatest. This is reflected in the robustness plots, where changes in the initial and final

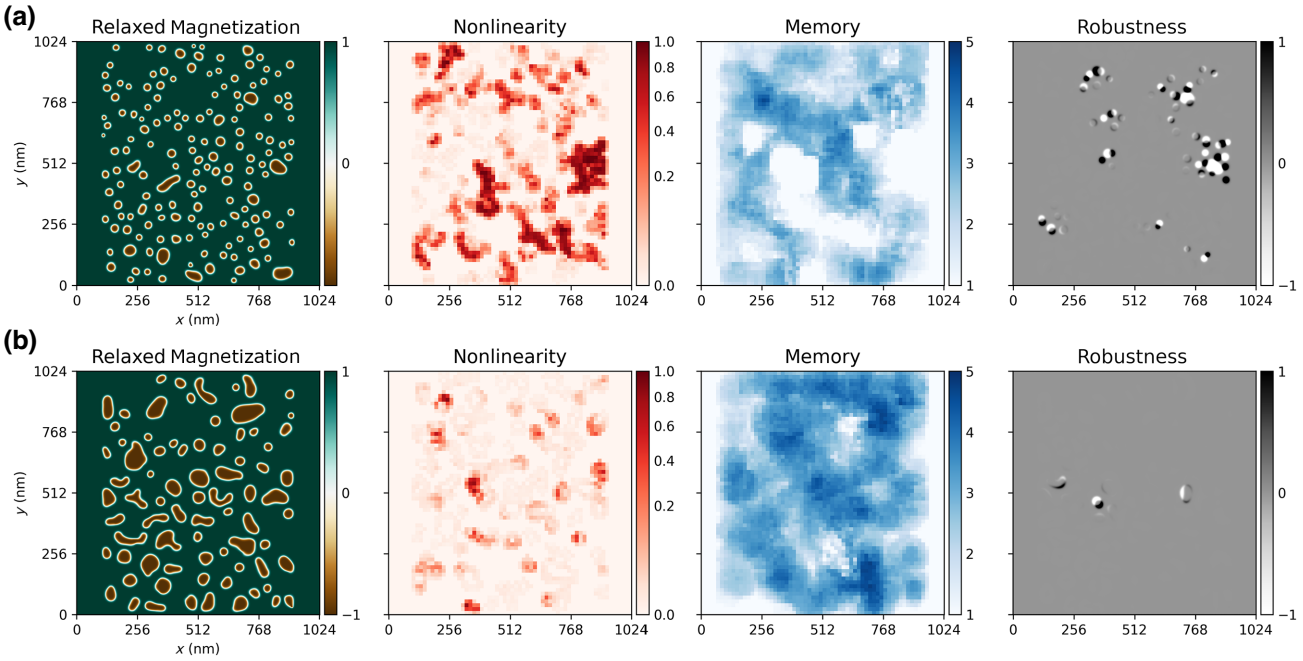


FIG. 3. Spatial analysis of two skyrmion reservoirs with (a) high and (b) low Dzyaloshinskii-Moriya interaction strengths. Relaxed magnetization shows the state of the reservoir when no input is applied. Nonlinearity and memory plots show the distribution said quantities over the reservoir. Robustness shows the difference between the reservoir's resting state and its stimulated state when measuring the nonlinearity and memory metrics.

states of the reservoir are observed. Although this results in the fading memory property being violated, the material can still act as a reservoir provided the global structure of the texture is maintained. In the skyrmion reservoir case, changes to local structure can be due to individual skyrmions switching between a small group of distinct pinning sites. However, this is not necessarily a problem as, once the reservoir is "warmed up," the dynamics of the reservoir may lead to skyrmions constantly switching between pinning sites. The final pinning site upon removal of the input is dependent on the local history of the input, and hence, the time at which the signal is removed from the reservoir.

Globally, our analysis suggests that the high DMI reservoir has an overall higher mean memory capacity and lower mean nonlinearity for input $u(t)$ than the low DMI reservoir. As both high memory capacity and high nonlinearity are desirable for reservoir computing, we next analyze a skyrmion "mixture reservoir" with a low-to-high DMI gradient. By mixing nonlinear and linear dynamics, we find that the skyrmion mixture reservoir has improved task performance over the low and high reservoirs. This complements the existing idea of an echo-state network-based mixture reservoir, where replacing small numbers of nonlinear nodes with linear nodes in the network improves their performance for time-series forecasting tasks [14].

We implement the skyrmion mixture reservoir by modifying our original reservoir to possess a DMI gradient from low (3.3×10^{-3} J m^{-2}) to high (3.6×10^{-3} J m^{-2}) along the axis parallel to the input-current direction. We extend the reservoir geometrically along the same axis by a factor of 4, so that the DMI gradient is shallow. In Fig. 4, we see that, as a result of the DMI gradient, there is also a gradient in the nonlinearity and memory metrics.

We next apply the reservoir to a chaotic time-series prediction task. We use a Mackey-Glass [25] system to generate a chaotic time series; see Appendix A 3 for more information. We input this series into the reservoir and train the readout to predict the value of the series at k steps in the future. The results for the prediction task are compared to the reservoirs with constant low and high DMIs, extended to the same physical size as the gradient reservoir. The outcomes of forecasting up to 50 future steps are illustrated in Fig. 5, revealing a minor performance improvement when using the gradient DMI mixture reservoir.

IV. SUMMARY

In this work, we demonstrate how two independent spatial metrics can be employed to locally quantify the properties of a reservoir computer in terms of nonlinearity and memory. Both metrics can be calculated in parallel using a small random dataset. This allows for a rapid evaluation of reservoir models and efficient screening of reservoir parameters for desired reservoir behavior and

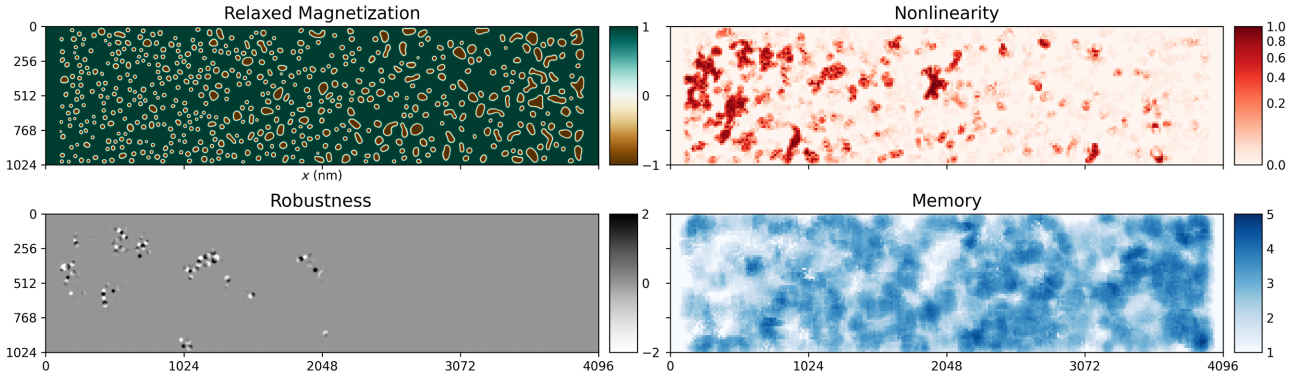


FIG. 4. Resolved memory, nonlinearity, and robustness for a system with a gradient in the Dzyaloshinskii-Moriya interaction strength along the axis parallel to the input current direction.

tuning of the reservoir performance to a specific purpose. For the particular reservoir computing model studied, the skyrmion reservoir, we find that dynamics in which both memory and nonlinear effects are present lead to improved task performance for the Mackey–Glass-series prediction task.

The metrics can also be used to better understand which situations lead to memory and nonlinearity in physical systems. By creating conditions that lead to desirable dynamics, one can parameterize reservoirs in a guided manner to optimize a balance between nonlinearity and memory. For example, in the case of the skyrmion reservoir, we achieve this by introducing a gradient in the DMI parameter.

It is important to mention that, although these metrics offer valuable insights into the spatial and emergent traits of a chosen reservoir system, they do not provide a holistic understanding of the reservoir's dynamics. This was confirmed in an additional investigation into the influence of reservoir material properties on the NL and MC metrics. For details, see Appendix A 5. Our findings indicate that, while the material properties govern skyrmion distribution and deformation, the dynamics of the skyrmion reservoir are contingent on several other crucial factors that are not adequately represented by the metrics proposed in this study. Nevertheless, our metrics demonstrate their value in various applications. As an illustration, these spatial metrics can provide a path to effectively choose a readout structure guided by local contributions to memory and nonlinearity. For instance, employing a smaller number of strategically positioned readout nodes can lead to a reduction in the quantity of model parameters, thereby enhancing the practicality of designing such devices.

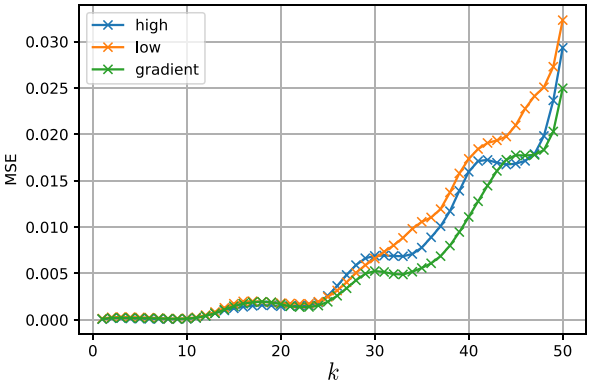


FIG. 5. Mean squared error for prediction of the Mackey-Glass series for k steps in the future. Results for the high, low, and gradient DMI reservoirs are shown.

ACKNOWLEDGMENTS

We acknowledge funding from the Emergent AI Centre funded by the Carl-Zeiss-Stiftung and the Deutsche Forschungsgemeinschaft under Project No. 320163632. J.L. was supported by the Fonds Wetenschappelijk Onderzoek (FWO-Vlaanderen) with senior postdoctoral fellowship No. 12W7622N. Part of the computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, the Research Foundation Flanders (FWO), and the Flemish Government Department EWI.

APPENDIX

1. Estimator Quality

To measure the quality of the linear estimators for both nonlinearity, Eq. (3), and memory, Eq. (5), we use the R^2 coefficient of determination. Consider the linear estimator,

$$\hat{y}(t) = \sum_i w_i u(t). \tag{A1}$$

We train the weights, w, of the estimator using the first 75% of our input, $u(t)$, and output, $y(t)$, samples, minimizing the residual sum-of-squares error. Using the remaining 25% of input samples, we compare the corresponding estimated

values, $\hat{y}(t)$, to the true value, $y(t)$, with

$$R^2[\hat{y},y] = \frac{\mathrm{cov}^2(\hat{y}(t),y(t))}{\sigma^2(\hat{y}(t))\sigma^2(y(t))}. \tag{A2}$$

Here, σ^2 is the variance and "cov" is the covariance. $R^2[\hat{y},y]$ is a measure for the estimator's quality, ranging from zero (no correlation) to one (perfect prediction). In this work, we use a total of 1000 data points for training and testing to measure estimator quality.

2. Simulation details for the skyrmion reservoir

The skyrmion reservoirs considered in this study were simulated using MuMax3 [24]. The homogeneous DMI reservoirs for the initial spatial analysis were modeled with a 1024×1024 cell grid with a cell size of 1 nm, which was smaller than the exchange length of the simulated material. The gradient reservoir, along with additional homogeneous DMI reservoirs for comparing Mackey-Glass prediction performance, were modeled with a 4096×1024 cell grid.

We simulated the Co layer of Pt/Co/MgO [26], with the following material parameters: saturation magnetization, $M_s = 580$ kA m^{-1}; exchange coupling constant, $A_{\mathrm{ex}} = 15$ pJ m^{-1}; mean uniaxial anisotropy, $K_u = 0.8$ MJ m^{-3}; temperature, $T = 0$ K; and electrical current polarization of 1. The DMI parameter is reservoir dependent, either 3.3×10^{-3} or 3.6×10^{-3} J m^{-2}. The anisotropic grain inhomogeneities were generated using a Voronoi tessellation [27] and had an average size of 40 nm and anisotropic variance of 20% [see Fig. 2(a) of the main text], similar to the experimentally observed grains in Ref. [28]. In the simulations, we did not include anisotropic magnetoresistance effects and instead assumed the material resistance to be constant.

We initialized the magnetic texture by first creating a skyrmion lattice and letting it relax to a metastable state. To further stabilize the system, we excited the texture with a random input current for 200 ns, then allowed it to relax. This was to remove any low-energy initial instabilities in the reservoir and to obtain a metastable state with a higher energy barrier suitable for RC.

3. Mackey-Glass details

The Mackey-Glass system is described by [25]

$$\frac{dx}{dt} = \frac{ax(t-\tau)}{1+x^n(t-\tau)} - bx(t), \tag{A3}$$

where $a = 0.2$, $b = 0.1$, $n = 10.0$, and $\tau = 23.0$. Since the dynamics of the series generated by Eq. (A3) are of a different timescale to the dynamics of the reservoir, we also scale time before using it as an input. This is done such that one unit of t in Eq. (A3) is scaled to 1 ns of physical input. The input is coupled to the reservoir with an input scaling of $j_{\mathrm{in}} = 3 \times 10^{11}$ A m^{-2}.

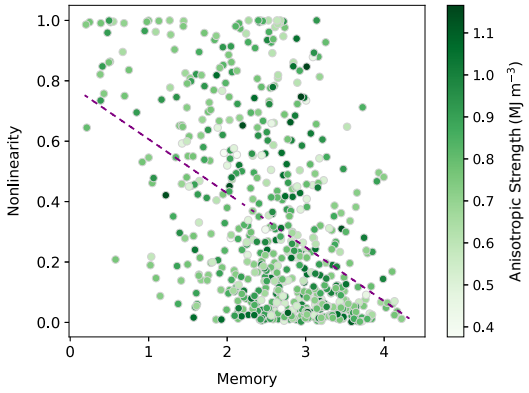


FIG. 6. Effect of magnetic anisotropy strength on memory capacity and nonlinearity metrics.

4. The micromagnetic model

We numerically solve for the magnetization dynamics by resolving the Landau-Lifshitz-Gilbert equation for the normalized magnetization, $\mathbf{m} = \mathbf{M}/M_s$, incorporating spin-transfer-torque effects by a spin-polarized-current term, $\mathbf{j}$, as follows:

$$\partial_t \mathbf{m} = -\gamma \mathbf{m} \times \mathbf{H}_{\mathrm{eff}} + \alpha \mathbf{m} \times (\partial_t + \frac{\beta}{\alpha}\xi\, \mathbf{j}[U,\mathbf{m}] \cdot \nabla)\mathbf{m} - (\xi\, \mathbf{j}[U,\mathbf{m}] \cdot \nabla)\mathbf{m}. \tag{A4}$$

Here, $\xi = P\mu_B/(eM_S)$, while M_s, P, μ_B, and e represent the saturation magnetization, current polarization, Bohr magneton, and electron charge, respectively. The effective field, represented by $\mathbf{H}_{\mathrm{eff}}$, is determined by the functional equation $\mathbf{H}_{\mathrm{eff}} = -M_S^{-1}\,[(\delta/\delta\mathbf{m})E[\mathbf{m}]]$, where the micromagnetic energy functional, $E[\mathbf{m}]$, is defined by the exchange stiffness, A_{ex}; anisotropy constant, K_u; and the DMI strength, D:

$$E[\mathbf{m}] = \int A_{\mathrm{ex}}(\nabla\mathbf{m})^2 + K_u(1-m_z^2) + D\mathbf{m}\cdot(\nabla\times\mathbf{m})\,\mathrm{d}V. \tag{A5}$$

5. Impact of reservoir material properties on nonlinearity and memory capacity

We investigated how the physical microstructure of the skyrmion reservoir material influenced the NL and MC metrics. To visualize this relationship, we generated scatter plots of the MC and NL values for individual readout nodes, utilizing the grain magnetic anisotropy strength as the color map, see Fig. 6. The dashed purple line in the figure is a trend line that further demonstrates the NL-MC trade-off.

The outcome of this investigation fell short of providing a comprehensive understanding of how microstructure impacts the balance between MC and NL. This is apparent from the findings presented in Fig. 6, which indicate

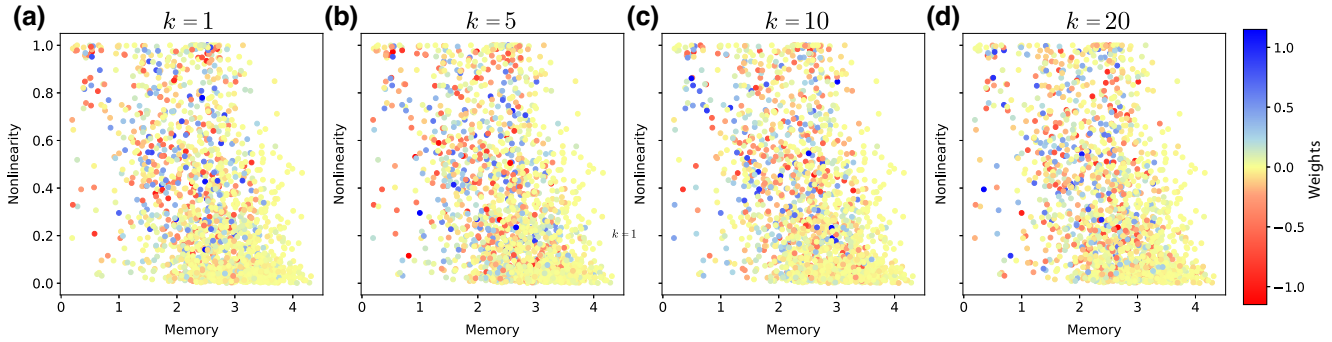


FIG. 7. Influence of memory capacity and nonlinearity metrics on regression weights across Mackey-Glass prediction delays: (a) $k = 1$, (b) $k = 5$, (c) $k = 10$, and (d) $k = 20$.

the absence of a simple correlation between NL, MC, and magnetic anisotropy strength. These results imply that the relationship is more intricate and nuanced than originally anticipated.

We also conducted a study to examine the contribution of regions characterized by various combinations of NL and MC levels (e.g., high NL and low MC, low NL and high MC) to the output regression weights at different prediction delays, k. We created scatter plots depicting the MC and NL values for individual readout nodes, using the regression weights of these nodes to define the color mapping. Figure 7 shows sample scatter plots for k steps of 1, 5, 10, and 20.

This study was conducted across all 50 k steps and no discernible pattern was found, as weights fluctuated dramatically at different k values.

While the NL and MC metrics offer a general understanding of the spatial distribution of respective quantities, the regression weights themselves are subject to significant changes at each step. These changes are driven by intricate microdynamics and, as a result, when the training weights are juxtaposed with the more stable measures of MC and NL, no striking correlation emerges. The interplay of these elements necessitates further investigation to gain a comprehensive understanding of their impact on regression in reservoir computing.

[1] M. Lukoševičius and H. Jaeger, Reservoir computing approaches to recurrent neural network training, Comput. Sci. Rev. **3**, 127 (2009).

[2] D. Verstraeten, B. Schrauwen, M. d'Haene, and D. Stroobandt, An experimental unification of reservoir computing methods, Neural. Netw. **20**, 391 (2007).

[3] H. Jaeger, *Tutorial on Training Recurrent Neural Networks, Covering BPPT, RTRL, EKF and the "Echo State Network" Approach* (GMD—German National Research Institute for Computer Science, Sankt Augustin, Germany, 2002).

[4] H. Jaeger and H. Haas, Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication, Science **304**, 78 (2004).

[5] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, Recent advances in physical reservoir computing: A review, Neural Netw. **115**, 100 (2019).

[6] J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, Information processing capacity of dynamical systems, Sci. Rep. **2**, 1 (2012).

[7] R. Nakane, G. Tanaka, and A. Hirose, in *2019 International Joint Conference on Neural Networks IJCNN* (IEEE, Budapest, Hungary, 2019), p. 1.

[8] C. Fernando and S. Sojakka, in *European Conference on Artificial Life* (Springer, Dortmund, Germany, 2003), p. 588.

[9] T. Furuta, K. Fujii, K. Nakajima, S. Tsunegi, H. Kubota, Y. Suzuki, and S. Miwa, Macromagnetic simulation for reservoir computing utilizing spin dynamics in magnetic tunnel junctions, Phys. Rev. Appl. **10**, 034063 (2018).

[10] D. Prychynenko, M. Sitte, K. Litzius, B. Krüger, G. Bourianoff, M. Kläui, J. Sinova, and K. Everschor-Sitte, Magnetic skyrmion as a nonlinear resistive element: A potential building block for reservoir computing, Phys. Rev. Appl. **9**, 014034 (2018).

[11] G. Bourianoff, D. Pinna, M. Sitte, and K. Everschor-Sitte, Potential implementation of reservoir computing models based on magnetic skyrmions, AIP Adv. **8**, 055602 (2018).

[12] D. Pinna, G. Bourianoff, and K. Everschor-Sitte, Reservoir computing with random skyrmion textures, Phys. Rev. Appl. **14**, 054020 (2020).

[13] L. Gonon and J.-P. Ortega, Reservoir computing universality with stochastic inputs, IEEE Trans. Neural Netw. Learn. Syst. **31**, 100 (2020).

[14] M. Inubushi and K. Yoshimura, Reservoir computing beyond memory-nonlinearity trade-off, Sci. Rep. **7**, 1 (2017).

[15] D. Verstraeten, J. Dambre, X. Dutoit, and B. Schrauwen, in *The 2010 International Joint Conference on Neural Networks IJCNN* (IEEE, Barcelona, Spain, 2010), p. 1.

[16] H. Jaeger, in *GMD Technical Report 152* (GMD—German National Research Institute for Computer Science, Sankt Augustin, Germany, 2002).

[17] S. Boyd and L. Chua, Fading memory and the problem of approximating nonlinear operators with volterra series, IEEE Trans. Circuits Syst. **32**, 1150 (1985).

[18] F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, All-optical reservoir computing, Opt. Express **20**, 22783 (2012).

[19] L. Grigoryeva, J. Henriques, L. Larger, and J.-P. Ortega, Optimal nonlinear information processing capacity in delay-based reservoir computers, Sci. Rep. **5**, 1 (2015).

[20] S. Tsunegi, T. Taniguchi, K. Nakajima, S. Miwa, K. Yakushiji, A. Fukushima, S. Yuasa, and H. Kubota, Physical reservoir computing based on spin torque oscillator with forced synchronization, Appl. Phys. Lett. **114**, 164101 (2019).

[21] G. Finocchio, M. Di Ventra, K. Y. Camsari, K. Everschor-Sitte, P. K. Amiri, and Z. Zeng, The promise of spintronics for unconventional computing, J. Magn. Magn. Mater. **521**, 167506 (2021).

[22] J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, and A. Fukushima, *et al.*, Neuromorphic computing with nanoscale spintronic oscillators, Nature **547**, 428 (2017).

[23] M. Romera, P. Talatchian, S. Tsunegi, F. A. Araujo, V. Cros, P. Bortolotti, J. Trastoy, K. Yakushiji, A. Fukushima, and H. Kubota, *et al.*, Vowel recognition with four coupled spin-torque nano-oscillators, Nature **563**, 230 (2018).

[24] A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sanchez, and B. Van Waeyenberge, The design and verification of mumax3, AIP Adv. **4**, 107133 (2014).

[25] M. C. Mackey and L. Glass, Oscillation and chaos in physiological control systems, Science **197**, 287 (1977).

[26] X. Wang, H. Yuan, and X. Wang, A theory on skyrmion size, Commun. Phys. **1**, 1 (2018).

[27] J. Leliaert, B. Van de Wiele, A. Vansteenkiste, L. Laurson, G. Durin, L. Dupré, and B. Van Waeyenberge, Current-driven domain wall mobility in polycrystalline Permalloy nanowires: A numerical study, J. Appl. Phys. **115**, 233903 (2014).

[28] M. Baćani, M. A. Marioni, J. Schwenk, and H. J. Hug, How to measure the local Dzyaloshinskii-Moriya interaction in skyrmion thin-film multilayers, Sci. Rep. **9**, 1 (2019).