Optimizing Reservoir Computing Based on an Alternating Input-Driven Spin-Torque Oscillator

Xuezhao Wu,[1,2,*] Zihan Tong,[1,3] and Qiming Shao[1,2,3,4,5,†]

[1] *Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong SAR, China*

[2] *State Key Laboratory of Advanced Displays and Optoelectronics Technologies, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong SAR, China*

[3] *ACCESS – AI Chip Center for Emerging Smart Systems, InnoHK Centers, Hong Kong Science Park, Hong Kong SAR, China*

[4] *Department of Physics and IAS Center for Quantum Technologies, The Hong Kong University of Science and Technology, Hong Kong SAR, China*

[5] *Guangdong-Hong Kong-Macao Joint Laboratory for Intelligent Micro-Nano Optoelectronic Technology, The Hong Kong University of Science and Technology, Hong Kong SAR, China*

(Received 23 July 2022; revised 19 May 2023; accepted 7 August 2023; published 28 August 2023)

Neurons in the brain show nonlinear oscillatorlike behaviors, inspiring the development of neuromorphic computing with spin-torque oscillators for their low power consumption and high integration density. The time-multiplexing technique enables utilizing only one device to achieve neuromorphic computing, further reducing power consumption and simplifying device fabrication. However, optimizing such systems remains a challenge and needs further development. A controversial hypothesis claims that dynamic systems possess the optimal computational capacity to address the information at the dynamical phase transition. Here, we apply this hypothesis to improve a reservoir computing system based on an alternating input-driven spin-torque oscillator. The input-driven property of reservoir computing allows us to exploit the dynamics of a spin-torque oscillator by manipulating the input stream. By tuning the ac amplitude in inputs, the system can exhibit diversified dynamic regimes indicated by the largest Lyapunov exponent and echo-state property spectra. Our findings demonstrate that the proper configuration of ac inputs for reservoir computing based on a single spin-torque oscillator can adjust its information-processing capacity profile and enhance its computational performance for specific tasks.

DOI: 10.1103/PhysRevApplied.20.024069

I. INTRODUCTION

Along with the rapid development and flourishing of artificial intelligence and big-data technologies, the demands of computing are increasing and expanding tremendously. Traditional computing systems face challenges when handling data-centric tasks, suffering from large power consumption and delays in data transfer due to the separation of memory and processing units in the conventional von Neumann structure. A revolutionary solution is to integrate these two units to break the bottleneck, giving rise to emerging nonvolatile memory devices and alternative computing paradigms. In other words, we aim to enable memory devices to process, i.e., compute, the saved information directly to cut off the requirement of information shuffling [1].

Spintronic devices are promising candidates for implementing in-memory logic and neuromorphic computing for their natural nonvolatility, high-speed, low power consumption, high integration density, and rich nonlinear dynamics [2]. Recently, there has been significant interest in computing systems that utilize different spintronic devices, including skyrmion [3], spin wave [4,5], domain wall devices [6], spin-torque oscillators (STOs) [7,8], and typical magnetic tunnel junctions (MTJ) [9,10]. Among these devices, STOs have gained popularity due to their nonlinear oscillating behavior, which is similar to neurons in the brain, making it a natural choice to construct a neuromorphic system with crossbar arrays of nonlinear oscillators [7,11]. It has been reported that the amplitude, frequency, and even phase of STOs can all be modulated for computing purposes [12]. Moreover, the effect of magnetic or electronic interaction, couple strength, and the thermal noises between STOs have also been studied [13]. Although the crossbar structure has achieved many successes in various scientific fields, it

*xwubx@connect.ust.hk
†eeqshao@ust.hk

simultaneously introduces some challenges, such as parasitic resistance and sneaky-path current, which may limit their application.

Recently, the proposal of the time-multiplexing method provides opportunities for implementing reservoir computing (RC) on a single STO, strongly reducing and even eliminating the drawbacks of fabrication, power consumption, and secondary effects [14]. RC [15] is a special recurrent neural network (RNN) with fixed input and recurrent layers, possessing a lower training cost than the typical RNNs. The reservoir, comprising many randomly connected neurons, can be constructed using various physical systems, including STOs [16]. However, optimizing RC with a single device is still a challenge. A general idea of improving the computing system for complex information processing is tuning the dynamic systems to operate in the so-called "edge of chaos," a unique critical dynamic regime between orders and chaos. Several reports [13,17–19] claim the best computational performance is achieved near the edge of chaos. However, recent findings indicate that the underlying principle of this enhanced capacity is generated at the edge of the echo-state property (ESP) [20–22]. Section IV will discuss these details further.

In this paper, we first introduce the concept of STO and create a nonlinear STO numerically by solving the Landau-Lifshitz-Gilbert (LLG) equation with the spin-transfer torque (STT) effect. We then utilize the time-multiplexing technique to the amplitude of a single STO, achieving the physical reservoir computing based on a single device. We employ short-term memory (STM) and parity-check (PC) tasks to quantify its computational performance. We demonstrated that diversified dynamical regimes of our computational system, indicated by the largest Lyapunov exponent (LLE) and ESP spectra, can be tuned by adjusting the ac configurations of inputs. The information-processing capacity (IPC) analysis shows that RC based on a single STO can perform well for STM and PC tasks by tuning the ac amplitude to a dynamic transition regime. Our findings will pave the way to the future design of neuromorphic applications based on STOs.

II. SPIN-TORQUE OSCILLATOR

STO is one type of MTJ whose oscillation is driven by charge currents with spin-torque effects, such as STT [11,23]. The typical structure of MTJ consists of two ferromagnetic layers (pinned layer and free layer) and a tunnel barrier, as shown in Fig. 1(a). The magnetization of the pinned layer is fixed in a specific direction, which can be done through the exchange bias effect from the interface of antiferromagnetism and ferromagnetism [24]. Oppositely, the magnetization vector in the free layer can be controlled by the direction of charge current through the STT effect [25]. The magnetization in the free layer is governed by the LLG equation incorporating spin-transfer dampinglike torque, shown in Eq. (1).

$$\frac{\partial \mathbf{m}}{\partial t} = -\gamma \mathbf{m} \times \mathbf{H}_{\text{eff}} + \alpha \mathbf{m} \times \frac{\partial \mathbf{m}}{\partial t} + \gamma a_j \mathbf{m} \times (\mathbf{m} \times \mathbf{e_x}). \tag{1}$$

Here, γ and α are gyromagnetic ratios and damping factors, respectively. $\mathbf{m} = \mathbf{M}/M_s$ is the normalized magnetization vector of the free layer. Due to the tunnel magnetoresistance effect, an MTJ will show high (low) resistance when the magnetization in the free layer is antiparallel (parallel) to the pinned layer. a_j is the STT magnetic field induced by charge current density J, including the dc and ac components in Eq. (2). Equation (3) depicts the relationship between STT magnetic fields a_j and the charge current density J. P, t_f, and M_s are the spin polarization, thickness of the free layer, and magnetization saturation, determined by materials and fabrication process. $\hbar$ is the reduced Planck constant, e is the elementary charge, and μ_0 is the vacuum permeability. We always set the magnetization vector of the pinned layer along the x axis and presume that the effective field H_{eff} includes only the external field H_{ext} and demagnetization field, as expressed in Eq. (4). From the view of dynamics, the first and second terms of the LLG equation [Eq. (1)] will lead the $\mathbf{m}$ to precess and gradually move toward the direction of the effective field, repectively. A continuous precession of magnetization in the free layer will occur when the amplitude of spin-transfer dampinglike torque induced by charge current is balanced with the Gilbert-damping torque [26]. Besides, we can obtain diversified dynamics of STO by adding an ac component to the dc charge current, as shown in the three insets of Fig. 1(a). The purple, blue, and red curves indicate the ordered, chaotic, and transitional dynamics of STO [27]. For the sake of narration, we directly encode the inputs as STT effective field a_j in the construction of our STO-based computing system. The two STT DC field values a_{dc} in Table I correspond to the binary inputs, which render the STO oscillate in two different trajectories. As a result, the responses of STO are nonlinear for the given input signal, which is a crucial requirement of RC [3]. We utilize the fourth Runge-Kutta method to establish the MTJ model with the physical parameters in Table I from Ref. [27]. The corresponding values of current densities used in our work can be calculated (see Appendix A).

$$a_j = a_{\text{dc}} + a_{\text{ac}} \cos(\omega_{\text{ac}} t). \tag{2}$$

$$a_j = \frac{\hbar P J}{2e\mu_0 M_S t_f}. \tag{3}$$

$$\mathbf{H}_{\text{eff}} = H_{\text{ext}} \mathbf{e_x} - 4\pi M_s m_z \mathbf{e_z}. \tag{4}$$

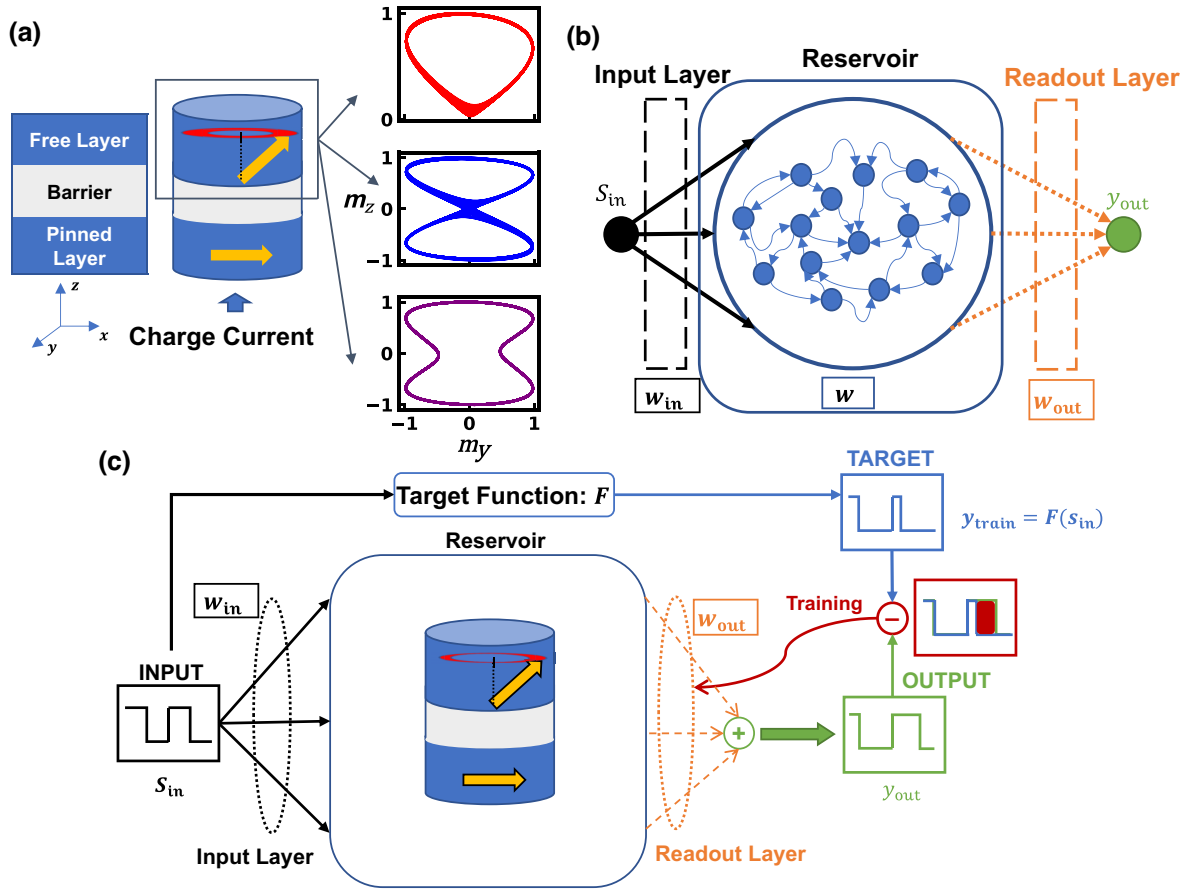

FIG. 1. (a) The illustration of a typical magnetic tunnel junction. We generate spin-transfer torque (field) by applying the charge current. It exhibits oscillating behavior when the amplitude of the spin-transfer torque gets balanced with the dampinglike torque. As depicted in the three insets, the different ac and dc field configurations of inputs will induce diversified dynamics. (Purple curve, $a_{dc} = 200$ Oe, $a_{ac} = 0$ Oe; red curve, $a_{dc} = 269$ Oe, $a_{ac} = 26$ Oe; blue curve, $a_{dc} = 260$ Oe, $a_{ac} = 26$ Oe.) (b) Schematic of classical reservoir computing framework. W_{in}, W, W_{out} are weights of the input layer, recurrent layer (reservoir), and readout layer, respectively. In RNN, all of them need to be adjusted. In RC, we train only W_{out} to reduce the difference between outputs and targets. The solid lines indicate the fixed weights. (c) Diagram of reservoir computing framework based on a single spin-torque oscillator. The magnetization amplitude of the spin-torque oscillator roles as a reservoir in this system.

III. PHYSICAL RESERVOIR COMPUTING BASED ON A SINGLE SPIN-TORQUE OSCILLATOR

RC is generated from RNN with fixed input and recurrent layer, proposed by Herbert Jaeger [15] and confirmed by the later discovery that the dominant weights changes are located in the readout layer during the training of RNN [28]. Figure 1(b) shows the classical structure of reservoir computing. Based on this specific structure, an exciting idea is proposed to replace the reservoir with physical systems to achieve alternative computing. Typically, we require the physical reservoir to satisfy two main properties, ESP and nonlinear response of given inputs [22]. ESP refers to the fact that the reservoir should gradually fade the influence of the initial values. This feature can be controlled by the spectral radius of the weight matrix in Eq. (5). W and W_{in} are the weight matrix of recurrent and input layers. f is the activation function, generating the nonlinear mapping of the previous reservoir state $X(n)$ and the current input $U(n+1)$. A general idea to establish the STO-based reservoir computing system (STO RC) is to employ the amplitude of the oscillator's responses [7]. This

TABLE I. Parameters of spin-torque oscillator in simulation.

Parameter	Symbol	Value[a]
Gilbert-damping constant	α	0.02
Gyromagnetic ratio	γ	1.7×10^7 Oe^{-1} s^{-1}
External field	H_{ext}	200 Oe
Demagnetization field	H_{deg}	8400 Oe
ac frequency	ω_{ac}	32 GHz
STT field induced by current	a_{dc}	200 Oe/100 Oe

[a]These values are from Ref. [27]

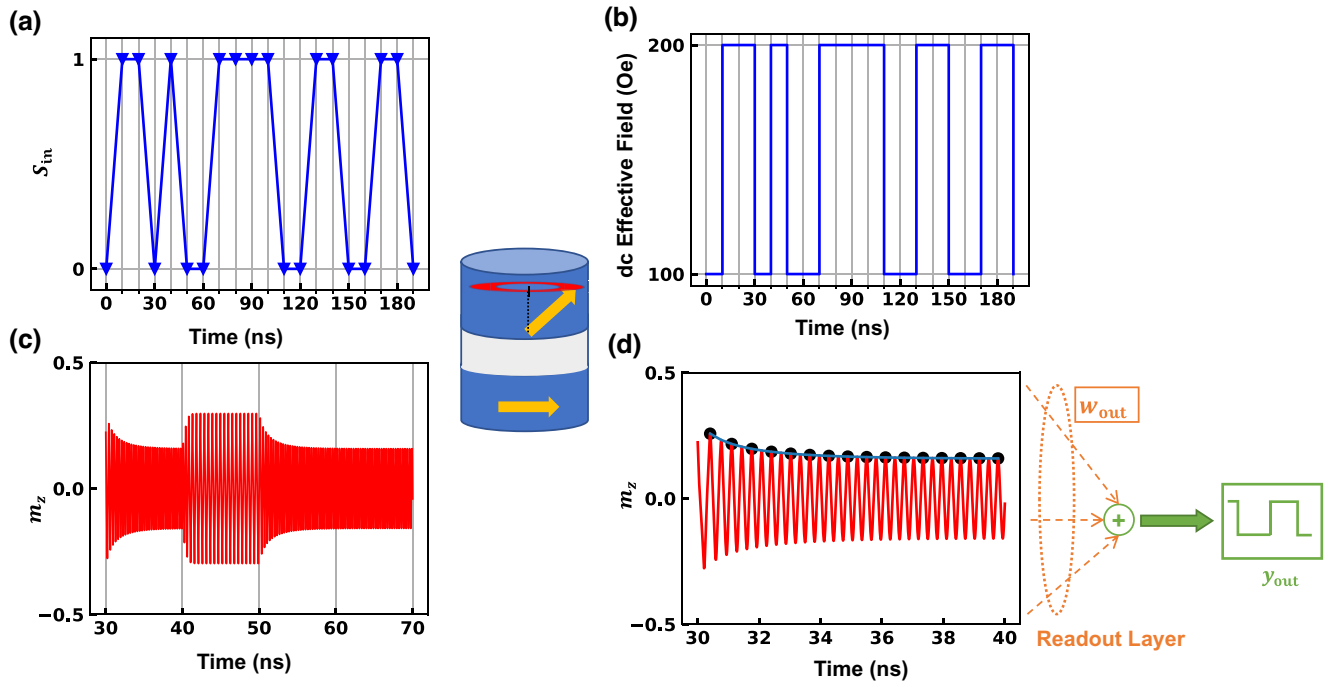

FIG. 2. Building reservoir computing based on a single spin-torque oscillator. (a) A random input signal S_{in} will be generated with a specific distribution. (b) Then it is encoded as the dc effective field a_{dc} with constant input widths. (c) The magnetization m_z in the free layer shows diversified oscillation behaviors corresponding to the effective field a_j, and (d) its amplitudes will be utilized to build the reservoir states X (the black dots) through time multiplexing. The number of nodes is 16 and the input width is 10 ns here.

experimental configuration works as Fourier transformation, introducing the memory and influencing the nonlinear mapping of STO externally. The binary inputs induce the STO to switch between two trajectories, generating the dependency of the inputs' history in STO's amplitude. The complete computational framework is shown in Fig. 1(c).

Figure 2 depicts the details of constructing the reservoir states from one STO. We encode each binary random input $S_{in}(n)$ as the STT dc effective field a_{dc}. The STT effective field a_j will then be created by ac field a_{ac} and dc field a_{dc} through Eq. (3), in which the ac field a_{ac} is independent to the input series S_{in}. The STO will generate its responses for the given STT magnetic fields a_j, following the rule of the LLG equation with STT effect [Eq. (1)]. We collect the amplitude of z axis component of magnetization m_z of STO free layer under given STT magnetic fields a_j and sample it periodically with a constant time interval. The sampling results act as the reservoir state $X(n)$ corresponding to the transient input $S_{in}(n)$. This method is called the time-multiplexing technique [14,29], replacing the chain reactions of neurons with a chain reaction in the time of a single STO to simplify the hardware cost. Finally, we can generate the reservoir matrix $X = [X(1), X(2), \ldots, X(n)]$ based on the total input $S_{in} = [S_{in}(1), S_{in}(2), \ldots, S_{in}(n)]$ and obtain the output Y_{out}. The reservoir matrix X and Y_{train} will be used to update the weight in readout layer. The matrix Y_{train} is determined by the specific computational task and delay time T_{delay}, which will be described clearly later. During the training, we utilize the ridge regression algorithm [Eq. (8)] to reduce the mean square error [L_{MSE}, Eq. (7)] between the expected output Y_{train} and the actual output Y_{out}. One decisive advantage of this algorithm is that $Y_{train}X^T$ and XX^T are irrelevant to the length of inputs, providing access for us to observe all middle weight states during the training [30].

$$X(n+1) = f(WX(n) + W_{in}U(n+1)). \tag{5}$$

$$Y_{out}(n) = \sum_{i=1}^{N} W_i x_i(n) + b = W_{out}(X(n) + b). \tag{6}$$

$$L_{MSE} = \frac{1}{T}\sum_{n=1}^{T}(Y_{train}(n) - Y_{out}(n))^2. \tag{7}$$

$$W_{out} = Y_{train}X^T(XX^T)^{-1}. \tag{8}$$

We use STM and PC tasks to evaluate the capacity of nonlinear mapping and memory effect of our RC system. An example of these two benchmark tasks is illustrated in Fig. 3. We obtain the expected output Y_{train} by using Eqs. (9) and (10) and define the correlation coefficient (Cor) and memory capacity (MC) to quantify our computing system's performance through Eqs. (11) and (12).

Figure 4 shows the results of MC values with different numbers of nodes versus the delay time T_{delay} in four input configurations. The input setting of Fig. 4(a) involves

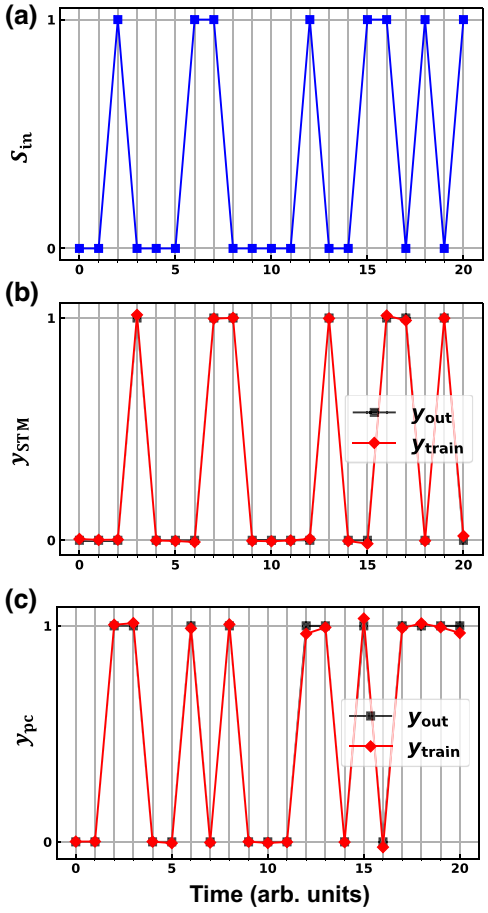

FIG. 3. Schematic of short-term memory and parity-check tasks with delay time $T_{delay} = 1$. (a) Random input S_{in} will be generated. Then, the corresponding train signal Y_{train} and output Y_{out} of STO RC from (b) short-term memory [Eq. (9)] and (c) parity-check tasks [Eq. (10)] are shown.

only the dc field without any ac STT field component. In Figs. 4(b)–4(d), different ac STT field a_{ac} will be applied to the STO all the time. We set the value of the maximum T_{delay} as 15 to preserve the accuracy of our memory capacity measurement. The results of Cor square values versus the different delay time T_{delay} are displayed in Fig. 5. It can be observed that the setting of the maximum delay time $T_{delay,max} = 15$ for the calculation of MC in STM and PC tasks is reasonable and sufficient because the Cor^2 is really close to zero from $T_{delay} \geq 8$. One property evaluated by the STM and PC tasks is nonlinear mapping, depending on the number of neurons in the neural network. An abstract explanation is that each input will induce particular reaction chains between neurons during the classification, and more neurons have more reaction chains, increasing the classification ability. Since we utilize only the chain reaction within one STO and the nonlinearity of its amplitude is finite, it can be observed from the results in Fig. 4 that the MC will saturate after a specific number of nodes. The nodes in a single device correspond to the neurons in the general neural network, sampled periodically from its nonlinear responses. Since we cannot generate infinite inputs to tune the readout layer, the actual computational performance of a trained STO RC will significantly vary due to the random inputs. We draw the error bar to describe this unavoidable phenomenon.

$$Y_{train,\ STM}(n) = S_{in}(n - T_{delay}). \tag{9}$$

$$Y_{train,\ PC}(n) = S_{in}(n - T_{delay}) \oplus S_{in}(n - T_{delay} + 1) \oplus \cdots \oplus S_{in}(n). \tag{10}$$

$$Cor^2(T_{delay}) = \frac{Cov^2(Y_{train}(T_{delay}), Y_{out}(T_{delay}))}{Var(Y_{train}(T_{delay}))Var(Y_{out}(T_{delay}))}. \tag{11}$$

$$C_{MEM} = \sum_{T_{delay}=0}^{T_{delay,\ max}} Cor^2(T_{delay}). \tag{12}$$

IV. EDGE OF CHAOS INDICATED BY THE LARGEST LYAPUNOV EXPONENT AND ECHO-STATE PROPERTY INDICATED BY SYNCHRONIZATION INDEX

It is common to improve computational performance by adjusting the neural-network parameters, such as the number of neurons, as mentioned in Fig. 4. However, this method cannot effectively find the best parameter configuration and thus enormously increases the training cost for searching parameters.

Recently, a more exciting viewpoint is to improve RC by adjusting the input streams. It can increase the tolerance margin for device fabrications and strongly influence computational performance without deliberate control of physical systems [20,31–33]. As shown in Fig. 4, the ac component is a significant factor for the total MC. The underlying principle is related to the dynamics of STO, which is strongly sensitive to the amplitude of ac current. To quantify the difference between the dynamic state of our system, we calculate the largest Lyapunov exponents and ESP index of reservoir states. As shown in Eq. (13), the LLE describes the exponential convergence or divergence of two trajectories with slightly different initial states, predicting the influence of perturbation on a system [34]. γ_0 is the initial perturbation between two trajectories, and γ_t is the difference as a function of evolution time. When $\lambda < 0$, the trajectories will converge towards each other over time, indicating a stable, nonchaotic system. When $\lambda \approx 0$, the system is in a critical transition regime between chaos and order. When $\lambda > 0$, the trajectories will diverge from each other over time, indicating a chaotic system.

$$\lambda = \lim_{t\to\infty} \frac{1}{t} \ln \frac{\gamma_t}{\gamma_0}. \tag{13}$$

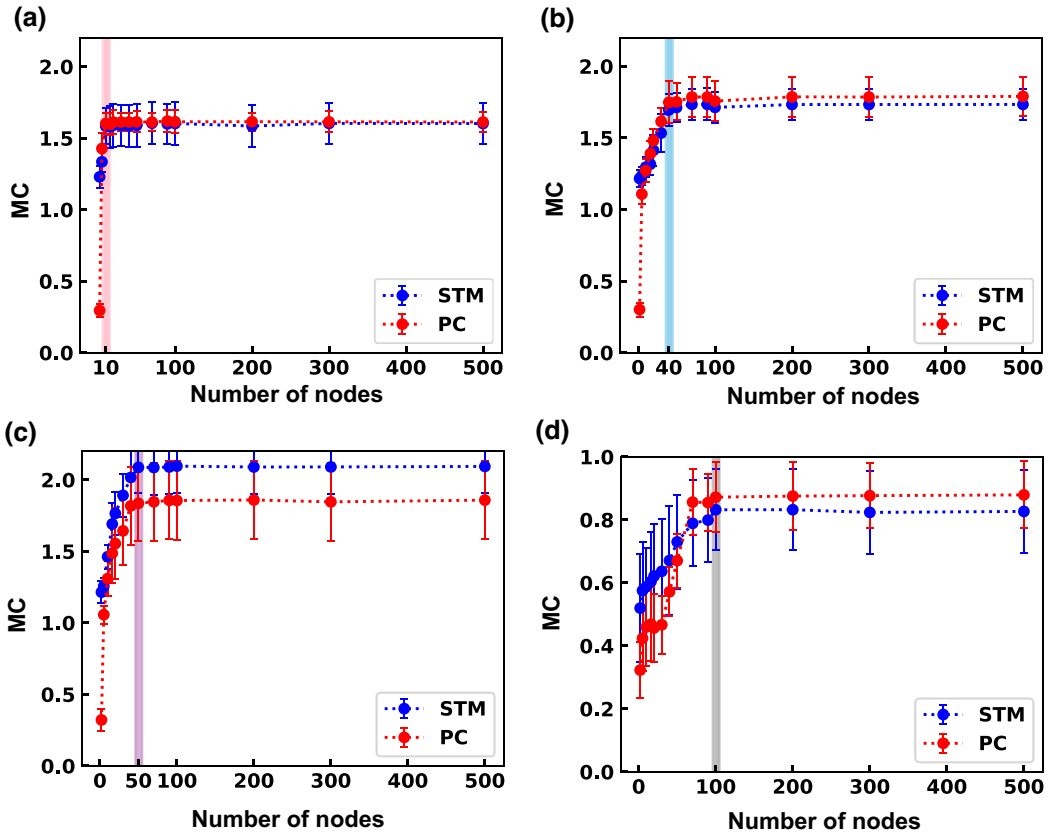

FIG. 4. The influence of nodes on the memory capacity of STO RC with different ac components (a) $a_{ac} = 0$ Oe, (b) $a_{ac} = 30$ Oe, (c) $a_{ac} = 50$ Oe, and (d) $a_{ac} = 100$ Oe. Each task will be calculated 10 times for different random inputs. The maximum delay time $T_{delay,\,max}$ is 15, and the input width is 4 ns. The standard errors are depicted as error bars for short-term memory and parity-check tasks.

The reservoirs should maintain their stability for different initializations due to the untrained input and recurrent layers [16]. In other words, the physical system acting as reservoirs should be insensitive to the different initial values and generate the synchronization between inputs and outputs, confirmed by the phenomena of common-signal

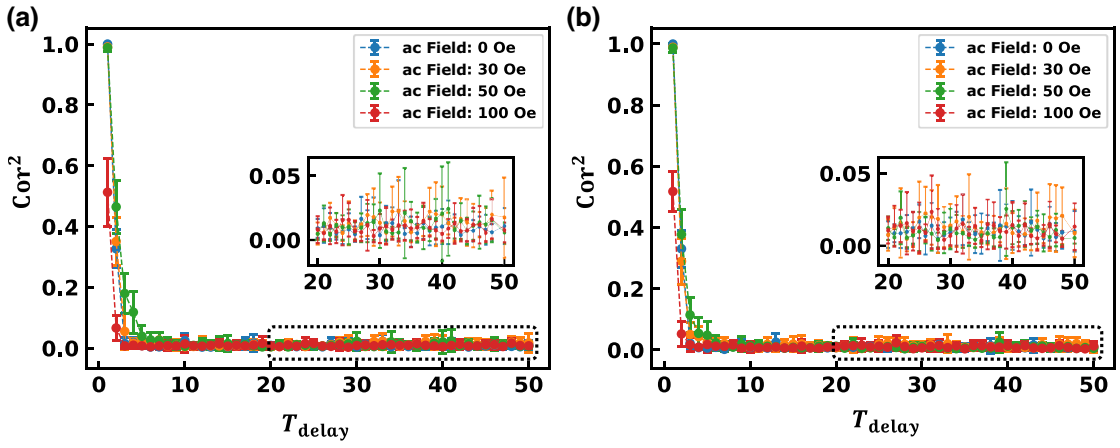

FIG. 5. The correlation coefficient square value of STO RC versus STM (a) and PC (b) tasks with the different delay time T_{delay}. Here, we show four input configurations (ac: 0, 30, 50, and 100 Oe). The Cor^2 values decrease with the increment of delay time T_{delay}, and the four colored curves are overlapped and close to zero starting at $T_{delay} = 8$.

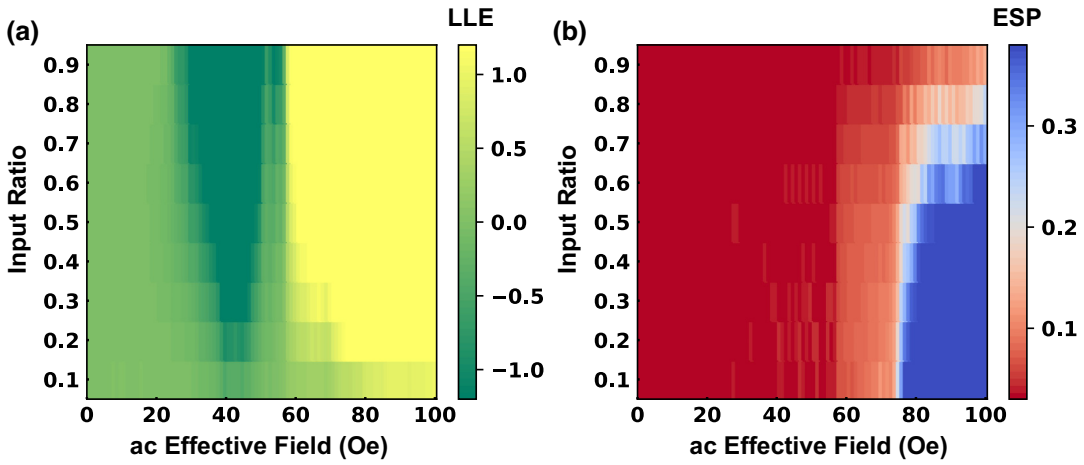

FIG. 6. The largest Lyapunov exponent spectra of STO RC and (b) results of the echo-state property with the different ac effective fields and different input ratios. The ESP value indexes are normalized to 1. The input ratio is defined as the ratio of 100 Oe in inputs (the probability P in Bernoulli distribution).

synchronization [35]. Otherwise, the reservoir will be useless and impractical if its performance depends strongly on the initialization. Recently, the quantified methods of ESP for the input-driven system have been proposed and utilized [20,31,35–37]. Moreover, the results of them are called the ESP index or synchronization index, defined as Eq. (14). Here, m_1 and m_2 are two STOs with common inputs but different initial values. The reservoir possesses the ESP only when the ESP index F is approximately equal to zero [31]. Additionally, the boundary between $F = 0$ and $F > 0$ is defined as the edge of the ESP [21].

$$F(T,L) = \frac{1}{L-T}\int_T^L d(m_1(t), m_2(t))dt. \tag{14}$$

The rich dynamics of STO have already been reported in the previous literature. For example, Refs. [27,33,38] revealed that an STO works at diversified dynamics regimes when adjusting the input configuration. In this paper, the input configuration is changed by tuning the magnitude of the ac current applied. According to the Poincare-Bendixson theorem, chaotic behavior is prohibited in a system with less than 3 degrees of freedom [31,32]. However, since the norm of m is fixed to 1, there are only 2 degrees of freedom in a dc-driven STO (two components among m_x, m_y, and m_z in Cartesian coordinates or θ and ϕ in spherical coordinates). By introducing ac components into the input, an extra degree of freedom can be obtained. Thus, an STO can be driven into a chaotic regime by increasing ac amplitude. The results of LLE spectra with different input ratios are depicted in Fig. 6(a), and the calculation details are as follows:

(1) Two identified STO RC modules are created with the same initial values and then evolved with the same time duration.

(2) A perturbation is then added to one of the STO RCs. Since the magnetization vector in the free layer satisfies $\|\mathbf{m}\| = 1$, the difference between two trajectories can be expressed as the Euclidean norm of the difference between two magnetization vectors. Thus, it is defined that $\gamma_t = \left\|\mathbf{m_t} - \mathbf{m_{p,t}}\right\|$ in which $\mathbf{m_t}$ and $\mathbf{m_{p,t}}$ refers to the magnetization vector of the unperturbed and perturbed STO at timestep t, respectively. Similarly, we also define the difference between the two reservoir states as $\gamma_t^{'} = \left\|\mathbf{X_t} - \mathbf{X_{p,t}}\right\|$. The initial values of γ_t and $\gamma_t^{'}$, denoted as γ_0 and $\gamma_0^{'}$, are also calculated at this time.

(3) A series of random inputs is generated to drive both STO RC systems. The systems will evolve for T time steps, during which $\ln(\gamma_t^{'}/\gamma_0^{'})$ is calculated and recorded.

(4) When the evolution is finished, the perturbated STO will be reinitialized to restrict the expansion rate of LLE [32,39], according to the following equation: $\mathbf{m_{p,T+1}} = \mathbf{m} + \gamma_0/\gamma_T(\mathbf{m_{p,T}} - \mathbf{m_T})$.

(5) Steps (3) and (4) will be iterated many times until the average value of $\gamma_t^{'}$ in each iteration converges. The average of $\gamma_t^{'}$ in the final iteration will be the LLE value calculated.

The calculation details of the ESP index for our input-driven STO RC are below. Firstly, we randomly generate the input sequence $S_{\text{in}} = [S_1, S_2, S_3, \ldots, S_L]$ and 51 initial values $[\mathbf{m_0}, \mathbf{m_1}, \mathbf{m_2}, \mathbf{m_3}, \ldots, \mathbf{m_{50}}]$ of magnetization in the free layer of STO. A transient time T is set, and its value should be smaller than the length of input sequence L. Then, the input sequences S_{in} will be continuously applied on the STO with each initial value $\mathbf{m_i}$, while the corresponding outputs of them will be collected as $[\mathbf{O_0}, \mathbf{O_1}, \mathbf{O_2}, \mathbf{O_3}, \ldots, \mathbf{O_{50}}]$. We define the outputs of STO as the evolution of reservoir states $\mathbf{O_i} = [\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3}, \mathbf{X_4}, \ldots, \mathbf{X_L}]$. Also,

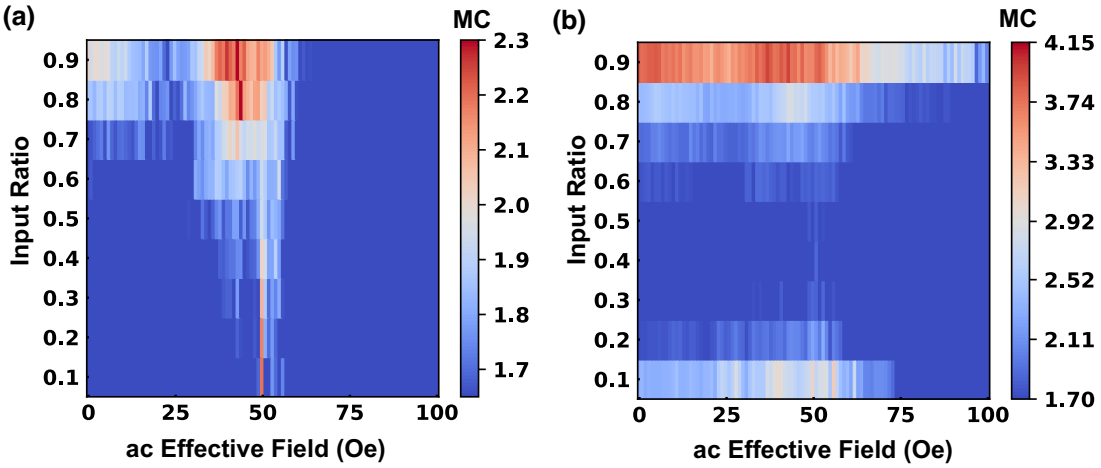

FIG. 7. The computational performance of STO RC for short-term-memory (a) and parity-check (b) tasks with different input ratios and ac components. The colors indicate the magnitude of memory capacity following Eq. (12) and the maximum delay time $T_{\text{delay,max}}$ used to calculate is 15. The number of nodes in STO RC is 100, and the constant input width is 4 ns. Each point is the average value of 10 times calculations.

We define the trajectories distance as the difference $\zeta_i = \sum_{j=1}^{L-T} \|\mathbf{O_i}(\mathbf{T}+\mathbf{j}) - \mathbf{O_0}(\mathbf{T}+\mathbf{j})\|$ for all reservoirs except the $\mathbf{O_0}$. Finally, the average result $F = 1/50\sum_{i=1}^{50}(\zeta_i/L - T)$ will be the ESP index. The result of the ESP index of STO RC is shown in Fig. 6(b).

To ensure the accuracy of our calculations, we have explored the LLE and ESP results with varying numbers of nodes (see Fig. 10 at Appendix B and Fig. 11 at Appendix C). Our results reveal that as the ac effective field a_{ac} increases, the STO RC gradually transits from orders to chaos in the STO RC. The amplitude of LLE is influenced globally by the input ratio, possibly because the inputs always follow the Bernoulli distribution. Both the LLE and ESP index results indicate that the STO RC can exit diverse dynamic regions by adjusting the ac effective field. The LLE and ESP index are two related, yet distinct measures, as highlighted in a recent study [21] that examined STO dynamics. Specifically, while the zero LLE of STO is well known, the same system can exhibit a nonzero ESP index, demonstrating the subtle differences between these two measures. Much research [17,40,41] has suggested that computational systems may perform optimally at or near the dynamical transition regime. However, not all computational tasks may benefit from it [42], and there is no conclusive evidence that it is a necessary condition for optimal performance [22]. Another hypothesis argues that the optimal conditions may be more closely aligned with the edge of the ESP, rather than the edge of chaos, which possibly explains the previous discoveries [21]. Our results reveal a significant overlap between the two edges of STO RC, suggesting that the echo-state property of STO RC is maintained before the dynamics become chaotic. The ESP of this system is primarily determined by the dynamics of STO, which is controlled by the ac effective field.

V. INFORMATION PROCESSING CAPACITY ANALYSIS FOR RESERVOIR COMPUTING BASED ON A SINGLE SPIN-TORQUE OSCILLATOR DRIVEN BY DIFFERENT INPUT STEAM

We have found that a range of ac field values can drive the system towards diverse dynamical regions. The boundary between orders and chaos is known to be a region where several basic information processing properties, such as information dynamics [41,43], susceptibility [17], sensitivity, synaptic variability [44], and correlation length [42], obtain their maximum values. The result in Fig. 7 shows computational performances of STO RC handling with the STM and PC tasks, quantified by the MC in Eq. (12). The STO RC performs best for these two tasks near the edge of the ESP in all input ratios, although the specific optimal ac points can vary for different input ratios. For instance, in the STM results, the 50-Oe ac effective field provides the largest MC in the input ratio range from 0.1 to 0.6, but this best ac point shifts to smaller ac field regions in the ratio range of 0.7 to 0.9. Similar phenomena are observed in the PC task. In general, ac effective fields slightly before the transitional region provide the STO RC with better computational ability than other regions. Furthermore, our analysis shows that there is a steeper decay of performance when the system moves from the edge of the ESP to the chaotic region compared to the movement to the ordered region. This behavior is consistent with the active information storage measurement of the echo-state network [41] and the information transfer measurement of random Boolean networks [45]. Additionally, the STO RC performs better in the ordered region than the chaotic region for both tasks. These behaviors can be attributed to the disappearance of ESP [36], which means that the STO

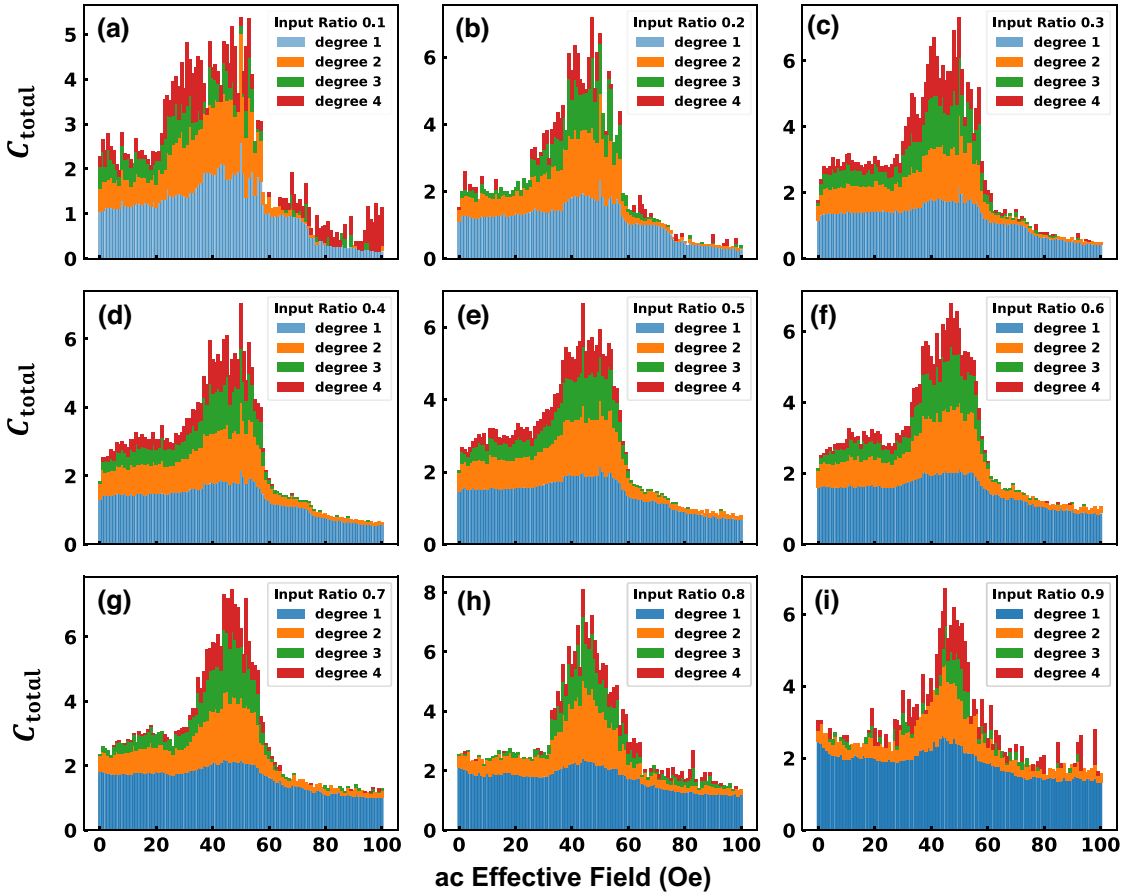

FIG. 8. Significant IPC values C_{total} of reservoir states in STO RC driven by different input ratios (a)–(i) and ac components. Each degree d_i indicates the sum of $[[d_i, \zeta_{d_i,j}]]$ IPC terms for all delay setting $\zeta_{d_i,j}$.

cannot effectively serve as a reservoir in the chaotic region and thus leads to poor performance.

In order to further validate our findings, we calculated the IPC profile of our STO RC and displayed the results in Figs. 8 and 9. The IPC is a variable tool for assessing the ability of dynamic systems to process information, providing a detailed analysis of the unknown process within the black-box systems [46,47]. The presence of ESP means that a system's dynamic state will ultimately become independent of its initial states after an infinite number of input streams, indicating that the dynamic state is solely determined and controlled by its input history. The IPC method decomposes the dynamic states of a system into a series of basis vectors generated by the input history, with the corresponding squared coefficients of those vectors defined as the IPC values. Additionally, the ideal outputs of all time-series tasks should be determined by the input signals, which can also be divided into several input-history-based vectors. Therefore, a system can be considered to have excellent computational capacity for a specific time-series task if its dynamic states consistently possess high IPC values for corresponding components of the ideal expected outputs [20]. In our case, serval orthogonal polynomials z_i are created based on the probabilistic distribution of inputs (Bernoulli distribution), delay setting ζ_{d_i}, and degree d_i. (See Appendix D for the generation of orthogonal polynomials z_i.) The reservoir states X were then decomposed using these orthogonal polynomials z_i, as expressed in Eq. (15).

$$X = \sum_{i=1}^{\infty} c_i z_i. \tag{15}$$

Here, X is the reservoir state. c_i and z_i are the ith coefficient vector and orthogonal polynomials, respectively. The typical way of measuring the IPC is to evaluate how well reservoirs can generate outputs that match the previous inputs [16]. This quantity is commonly used in echo-state networks with nonlinear activation functions in both continuous-time [48] and discrete-time configurations [44,49].

The IPC profile of STO RC is calculated as follows. The 10000-length independent identically distributed (IID) random signal S_{in} with a specific distribution (Bernoulli

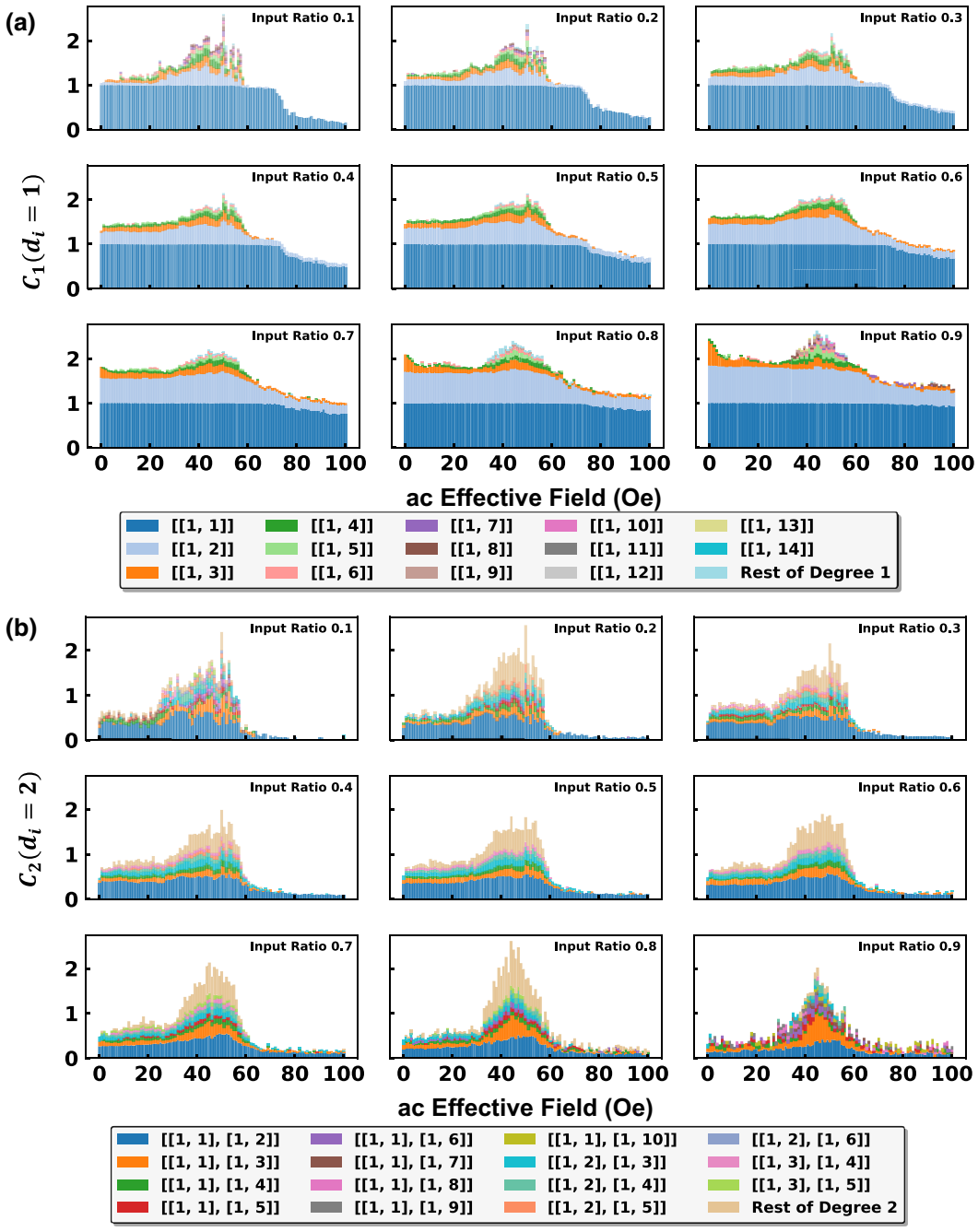

FIG. 9. Significant IPC values with degree $d_i = 1$ (a) and $d_i = 2$ (b) of reservoir states in STO RC driven by different input ratios and ac components. The larger IPC value C_i indicates the better ability to process the information. The degree and delay settings $[[d_i, \zeta_{d_i,j}]]$ are shown in the list. The components with IPC values less than 0.05 (0.1) will be summarized as the rest of the degree $d_i = 1$ ($d_i = 2$).

distribution) is created and written into the STO with a constant input width (4 ns in our case). The degree d_i and delay set $\zeta_{d_i} = \left[\zeta_{d_i,1}, \zeta_{d_i,2}, \zeta_{d_i,3}, \ldots, \zeta_{d_i,d_i}\right]$ will be configured first. The length of delay set ζ_{d_i} is equal to the value of the degree d_i. According to the degree d_i and delay set ζ_{d_i}, the orthogonal polynomials will be generated by inputs S_{in} as $z_i = \prod_{j=1}^{d_i} S_{\text{in}}(t - \zeta_{d_i,j}) - \langle S_{\text{in}} \rangle$ through the Gram-Schmidt orthogonalization procedure [37]. Then the reservoir states X are extracted by using the single value decomposition (SVD) to obtain the r-normalized matrix X^r. The average value of reservoirs $1/i \sum X_i$ will be removed from reservoirs X before the SVD process to remove the constant-induced capacity. The r is the rank of reservoir states X. Equation (15) will be rewritten as Eq. (16), showing that

the r-normalized reservoirs X^r can be expanded with the basis of normalized polynomials $z_i/||z_i||$.

$$X^r = \sum_{i=1}^{\infty} c_i \frac{z_i}{\|z_i\|}, \tag{16}$$

$$C_i = \|c_i\|^2, \tag{17}$$

$$C_{\text{total}} = \sum C_i = r. \tag{18}$$

The ith IPC is defined as the norm of the coefficient vector in polynomials [Eq. (17)], while the sum of the compete IPC should be equal to the rank value of reservoir states [Eq. (18)]. IPC value of C_i quantifies the computational ability of STO RC. To avoid the numerical error of IPC calculation from the finite length of the input series, the shuffle surrogate procedure will be applied to all the IPC values [37,46]. After obtaining the reservoir states X with inputs, we will randomly shuffle the input signals but not transfer them to the STO RC anymore. The alternative orthogonal polynomials z_i will be put in Eq. (16) to calculate the reference IPC (C_{shuffle}). We repeatedly calculate this procedure 200 times to get 200 C_{shuffle} with the same reservoir states X but different input sequences S_{in}. The maximum values of C_{shuffle} will be used to generate the threshold value of IPC results through Eq. (19). The scale factor σ_{thresold} will be set as 1.2–3 in our experiments. Note that all IPC results shown in the figures are significant IPC ($C_{\text{significant}}$).

$$C_{\text{significant},i} = \begin{cases} C_i, & \sigma_{\text{thresold}} * C_{\text{shuffle, max}} < C_i \\ 0, & \text{otherwise} \end{cases}. \tag{19}$$

Figure 8 illustrates the IPC profile of reservoir states for STO RC driven by different ac effective fields and input ratios. As described in Eq. (18), the completeness property of IPC ensures that the total IPC should be equal to the rank of the reservoir states' matrix r ($r = 46$ in our case). However, due to the presence of a threshold [Eq. (19)], IPC values below the threshold are disregarded, leading to lower total IPC values in Fig. 8 than the corresponding ideal values [50]. Additionally, the loss of the ESP would also decrease the total IPC of reservoir states. The fact that the total IPC of reservoir states is less than the number of nodes (100) indicates that some reservoir states are not linearly independent. It is worth noting that the total IPC of output in a fixed reservoir remains constant (see a good example at Appendix E), and the additional ac effective field provides a way to modify its IPC profile. By adjusting the IPC profile of the system's output to match the target, we can effectively increase its computational capacity. As described in Eq. (6), the readout layer connecting the reservoir states with the output can be trained, enabling us to extract the computing capacity from the reservoir states to the output.

We find that the sum of significant IPC values C_{total} is largest near the transition between order and chaos for all input ratios, indicating the rich computing capacity of STO RC induced by the edge of the ESP. Specifically, the STO RC driven by an ac STT effective field of around 50 Oe exhibits better computational ability than the condition without ac or with extremely high ac fields, consistent with the memory capacity results of STM and PC tasks in Fig. 7. According to Eq. (9), the IPC profile with degree $d_i = 1$ strongly supports the accuracy of our memory capacity results in Fig. 7 and the reasonability of the configuration of maximum delay time $T_{\text{delay,max}} = 15$ in the results from Figs. 4 and 6. Furthermore, we observe that the reduction in significant IPC is more robust when the system moves from the edge of the ESP to the chaotic region compared to the transfer to the ordered regime. Figure 9 shows the IPC values for different input ratios with degree $d_i = 1$ and $d_i = 2$. Similarly, we observe that the significant IPC decays faster as it moves from the transition point to the chaotic region.

Interestingly, we find a trade-off that the significant IPC value increases with $d_i = 1$ but decreases with $d_i = 3$ as the input ratio increases. This suggests that smaller input signals ($a_{\text{dc}} = 100$ Oe) tend to drive the STO's dynamic states into the low-rank level, suppressing the nonlinear mapping ability of the reservoir. In addition, the optimal ac field points with different input ratios in Fig. 8 exhibit some inconsistency with the edge of chaos regimes (and the edge of ESP), particularly for high input ratios. For instance, the STO RC driven by an input stream with a high input ratio (from 0.7 to 0.9) achieves its highest significant IPC in the ordered regime. This discrepancy between dynamical transition and the best optimization has been previously reported and observed in software-based continuous-time [48] or discrete-time [49] reservoir computing neural networks, as well as in hardware-based on-chip neuromorphic computing systems [42]. The shift is attributed to the static or dynamic suppression of chaos due to the input noise. The stochastic external input can alter the transition of the edge of chaos by forcing the system to move to different unstable phase spaces. However, our input streams do not include stochastic signals and consist only of ac and dc components. Nonetheless, ac current-induced phase-locking behavior has been reported in STO, where it is controlled by the input strength and ac frequency [51]. More components around the external ac frequency would occur while the STO operates outside the phase-locking regimes by adjusting the dc amplitude [27]. Therefore, we suggest that configurations of dc effective fields a_{dc} may cause the shift of the optimal point in high-input ratios due to the presence of different redundant responses of the STO. The impact of the dc effective field on the IPC profile can be observed by comparing the IPC results for different input ratios. For example, the IPC of $[[d_i, \zeta_{d_i,j}]] = [[1, 2]]$ increases with the input ratio, and more types of

significant IPC occur when the input ratio is either 0.1 or 0.9.

As shown in Fig. 8, not all IPC components with different $\zeta_{d_i,j}$ attain their maximum values at the edge of the ESP, despite their sum being the highest at that location. For example, the IPC values of component $[[d_i, \zeta_{d_i,j}]] = [[1, 1]]$ remain constant at most ac fields, especially for cases with an input ratio of 0.9. Furthermore, the IPC of component $[[d_i, \zeta_{d_i,j}]] = [[1, 3]]$ achieves the best values in the ordered region instead of the edge of the ESP in the input ratio range from 0.7 to 0.9. Similar failures have been reported in previous studies [18,42,52] for some simple tasks in which the performance at the dynamical transition regime is either equal or worse than that in the ordered region. The computation may be disturbed if the input information stored in the reservoir is inconsistent with the task requirements, i.e., the system retains an abundance of task-irrelevant information [41]. Based on our significant IPC results, it is easy to explain this task-related property. For instance, our STO RC would not benefit from tuning the ac field when addressing the simplified STM task with only $T_{\text{delay}} = 2$ in the input ratio range from 0.7 to 0.9. In fact, the STO RC without any ac component performs even better for this simplified task in the input ratio range from 0.7 to 0.9. Although the sum of IPC is the largest and IPC components are the most abundant at the edge of the ESP, the task-irrelevant IPC term would not contribute to the information processing of the STO RC. Another example is the simplified STM task with only $T_{\text{delay}} = 1$, for which the STO RC cannot achieve higher computational performance at the edge of the ESP. The similar phenomena can be found in the significant IPC results of degree $d_i = 3$ and $d_i = 4$ in Appendix D. It is commonly believed that the dynamical transition point is where a system performs best when handling complicated tasks [18,44]. This is because complicated tasks may require systems with more and different IPC components, and the edge of the ESP, which includes the richest IPC components, should benefit the dynamic system when addressing such tasks. Moreover, the abundant components of IPC at the edge of the ESP demonstrate the potential of spintronics in neuromorphic computing.

To confirm our discovery, we utilize our STO RC to emulate the nonlinear autoregressive-moving-average (NARMA) model defined in Ref. [53], which is often used to quantify computing capacity [46,50,54]. It is worth noting that we still use binary random inputs to handle NARMA tasks, as we did in the IPC calculation, STM, and PC tasks. The testing results and IPC analysis of NARMA2, NARMA3, and NARMA5 are also consistent and described in Appendix E. For all the tasks mentioned in this paper, the IPC spectra accurately quantify or predict the corresponding results of the STO RC.

These results demonstrate that the edge of the ESP can be a guideline for optimizing the STO RC. We demonstrate that not only stochastic signals but also ac signals can be used to approach the edge of the ESP of STO-based computing systems. In our STO RC, the two edges of chaos and ESP almost overlap, and the best computational performance for STM and PC tasks is not aligned with these two edges. We suggest that this phenomenon is generated by the dc configuration and the dynamic properties of the STO itself. To further clarify the relationships between these factors, we need more powerful tools to gain accurate insights in the future.

VI. CONCLUSION

We establish an RC system based on a single STO by solving the LLG equation with the STT term and employing time multiplexing. We show that tuning the ac input streams can adjust the dynamic states of STO, as indicated by the LLE and ESP index, and enhance the computational performance of the STO RC for STM and PC tasks. Our IPC analysis reveals that the additional ac signals allow us to adjust the IPC profile of the STO RC, indicating that using extra ac inputs can improve computational performance of specific tasks. We hope our findings can contribute to the future design and applications of neuromorphic computing with STOs.

ACKNOWLEDGMENTS

This research is partially supported by the State Key Laboratory of Advanced Displays and Optoelectronics Technologies (HKUST). We acknowledge funding support from the Shenzhen-Hong Kong-Macau Science and Technology Program (Category C, Grant No. SGDX2020110309460000), RGC General Research Fund (Grant No. 16303322) and the Research Fund of Guangdong-Hong Kong-Macao Joint Laboratory for Intelligent Micro-Nano Optoelectronic Technology (Grant No. 2020B1212030010).

APPENDIX A: THE CHARGE CURRENT DENSITY CORRESPONDING TO THE STT MAGNETIC FIELD USED IN THE SPIN-TORQUE OSCILLATOR SIMULATION

In our configuration, the random inputs are first encoded as STT field a_j induced by the charge current density J. To prove the experimental reasonability of our simulation, we calculate the current density values corresponding to the STT dc effective field by utilizing the parameters shown in Table II. The calculation results of current density are 3.84×10^{10} A/m^2 for $a_{\text{dc}} = 100$ Oe and 7.68×10^{10} A/m^2 for $a_{\text{dc}} = 200$ Oe, which is normal for current-drven STO [55].

TABLE II. Parameters used in the calculation of charge current density.

Parameters	Symbols	Values[a]
Spin polarization	P	0.58
STT magnetic field	a_j	200 Oe/100 Oe
Vacuum permeability	μ_0	1.2566×10^{-6} H/m
Saturation magnetization	M_s	6.68×10^5 A/M
Free layer thickness	t_f	1.1 nm

[a]These values are from Ref. [56]

APPENDIX B: THE LARGEST LYAPUNOV EXPONENT WITH DIFFERENT NUMBERS OF NODES AND INPUT WIDTHS

Figure 10 shows the relationship between the LLE and the numbers of nodes or input widths. The LLE is irrelative to the number of nodes, demonstrating that the LLE results are converged and reasonable to present the dynamic regions of STO RC. The dependency on input width (evolution time) further verifies that the SOT RC is an input-driven system. The input width change should significantly influence the RC system's dynamics and performance.

APPENDIX C: THE ECHO-STATE PROPERTY SPECTRA WITH DIFFERENT NUMBERS OF NODES

Figure 11 is the result of the ESP index of STO RC with different numbers of nodes. To present the results more clearly, the fake ESP (FSP) values will also be calculated. During the ESP calculation, we define the results of the ESP index as FSP if the input sequences are different for different initial values m_i [37]. The FSP can work as an ESP reference because the 51 STOs with different initializations are irrelative in the calculation [37]. Therefore, the reservoir will possess ESP if the FSP value is far away from the ESP index. Otherwise, the RC should not have any ESP. The results depicted in Fig. 11 demonstrate that the number of nodes does not have any influence on the ESP index, similar to the results of LLE in Appendix B. It is reasonable that the ESP originates from the reservoir itself and is irrelevant to the RC configurations.

APPENDIX D: THE INFORMATION PROCESSING CAPACITY (IPC) RESULTS IN HIGHER RANK

The IPC values of STO RC reservoirs with degree $d_i = 3$ and $d_i = 4$ are depicted in Fig. 12. The IPC results with different input ratios also support our conclusions that the STO RC shows the best computational performance near the edge of the ESP. Similar to the results of degree $d_i = 1$ and $d_i = 2$, the chaotic region decays the significant IPC of STO RC more strongly. The edge of the ESP gives rise to the more and richer IPC components of the dynamics system, which are more remarkable in degree $d_i = 3$ and $d_i = 4$ than the cases in degree $d_i = 1$ and $d_i = 2$. The condition when input ratio of 0.1 shows some difference. The ordered dynamics could provide the almost equivalent IPC at the edge of the ESP, although the largest one is still in the transitional regions. Besides, STO RC still possesses the significant IPC values in chaotic regions in input ratios of 0.1, 0.8, and 0.9, indicating the influence of dc effective field configuration in STO RC.

$$z_{d_i} = S_{\text{in},t} - \sum_{j=0}^{d_i-1} C_j z_j, \tag{D1}$$

$$C_j = \frac{\langle z_j, S_{\text{in},t} \rangle}{\langle z_j, z_j \rangle}. \tag{D2}$$

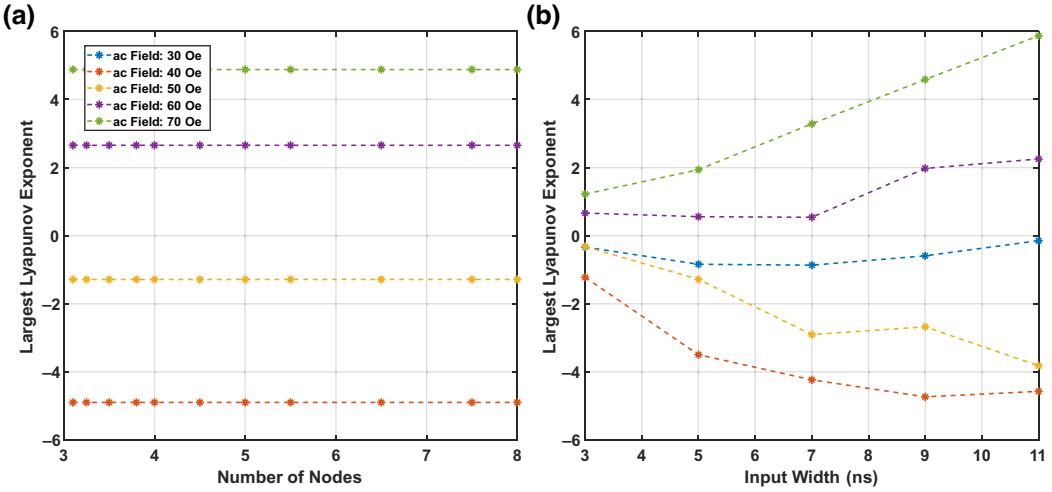

FIG. 10. The LLE values under different numbers of nodes (a) and input widths (b) with different ac effective fields.

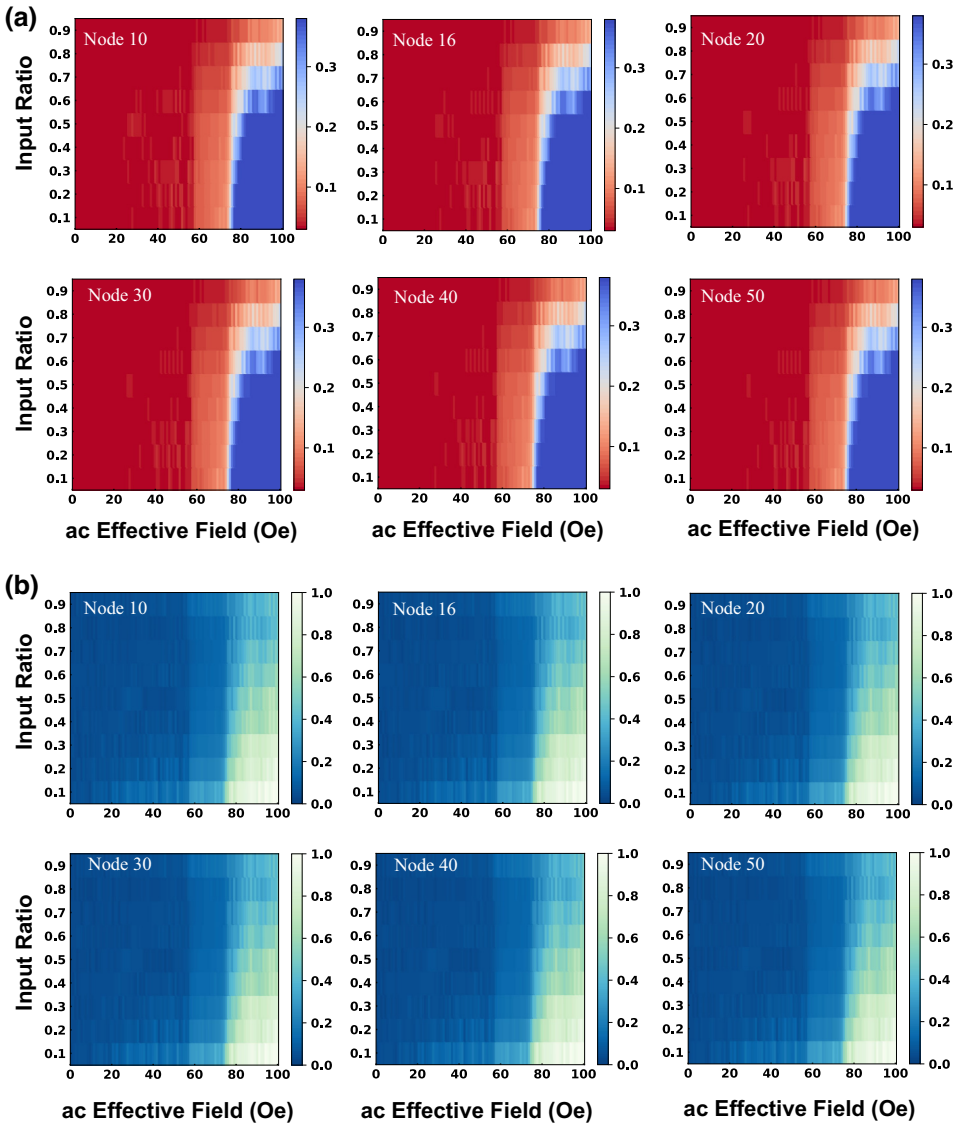

FIG. 11. (a) ESP results with the different numbers of nodes. (b) Results of the ratio between the ESP and FSP index. During the ESP calculation, we will get the FSP results if we use different input sequences for different initial values m_i. The reservoir loses its ESP if the FSP is close to the ESP index value. The input width is 4 ns for the calculation.

During our calculations of significant IPC for STO RC, the orthogonal polynomials z_i are determined by inputs s_{in} based on the Gram-Schmidt orthogonalization procedure following Eqs. (D1) and (D2). The Gram-Schmidt orthogonalization procedure is suitable for generating arbitrary polynomials for any probability distribution [46]. To describe our calculation more clearly, Table III shows examples of orthogonal polynomials z_i for Bernoulli distribution ($z_0 = [1, \ldots, 1]$).

TABLE III. Examples of orthogonal polynomials built by the Gram-Schmidt procedure.

Degree d_i	Delay $\zeta_{d_i,j}$	Orthogonal polynomials z_i
1	[1]	$S_{\text{in}}(t-1) - < S_{\text{in}} >$
1	[2]	$S_{\text{in}}(t-2) - < S_{\text{in}} >$
2	[1,2]	$(S_{\text{in}}(t-1) - < S_{\text{in}} >)(S_{\text{in}}(t-2) - < S_{\text{in}} >)$
2	[2,3]	$(S_{\text{in}}(t-2) - < S_{\text{in}} >)(S_{\text{in}}(t-3) - < S_{\text{in}} >)$
3	[1,2,3]	$(S_{\text{in}}(t-1) - < S_{\text{in}} >)*$ $(S_{\text{in}}(t-2) - < S_{\text{in}} >)(S_{\text{in}}(t-3) - < S_{\text{in}} >)$

APPENDIX E: THE STO RC PROCESSING NARMA2, NARMA3, AND NARMA5 TASKS

We utilize our STO RC to address the NARMA2, NARMA3, and NARMA5 tasks. The corresponding performance is displayed in Fig. 13. The definitions of NARMA tasks follow Eq. (E1) (for the NARMA2 task)

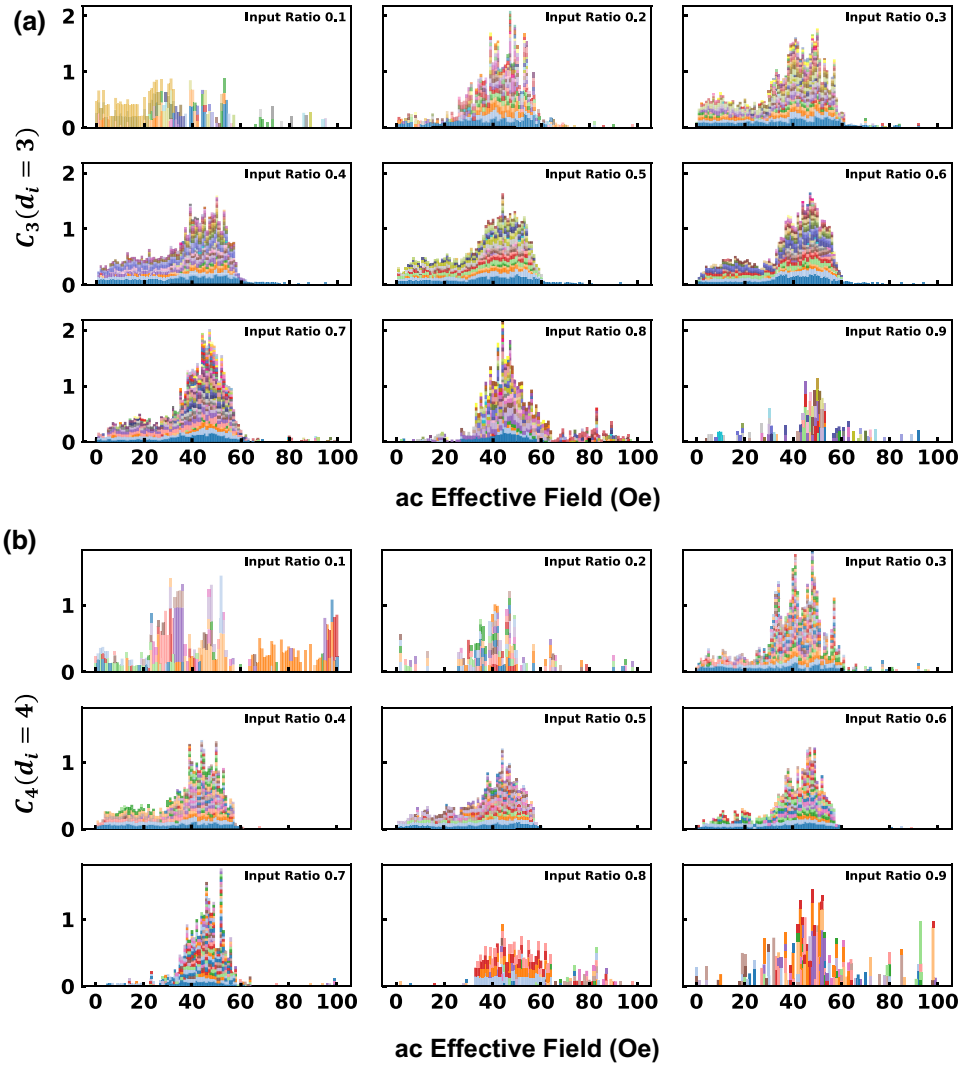

FIG. 12. The information processing capacity decomposition of STO RC reservoirs for degree $d_i = 3$ (a) and $d_i = 4$ (b). Each color bar in the figures corresponds to the specific different degree and delay settings $[[d_i, \zeta_{d_i,j}]]$. The input width is 4 ns, and the number of nodes is 100. Note that the IPC values less than 0.01 will be removed.

and Eq. (E2) (for n-rank NARMA tasks) [53,54]. The computational performance is quantified by normalized mean-square errors (NMSEs), as depicted in Eq. (E3). Besides, each ideal IPC component of these tasks will be calculated and shown in Fig. 14 to analyze these tasks. As described in Fig. 14(a), the degree $d_i = 2$ IPC component could be ignored in the NARMA2 task, indicating that NARMA2 task is not sensitive to the change in our input distribution (Bernoulli distribution). However, the requirement of reservoirs become more complicated due to high-order terms in the NARMA3 and NARMA5 tasks. The binary input s_{in} in these tasks is normalized to 0 and 0.2 to avoid numerical divergence. The IPC profile of these NARMA tasks presents the completeness property of IPC itself, indicating that the change of input influence only the IPC profile, rather than the constant total IPC. The slight difference of total IPC is generated by the configuration of threshold during the IPC calculation, as mentioned in Sec. V.

$$y_{k+1} = 0.4y_k + 0.4y_k y_{k-1} + 0.6S^3_{\text{in},k} + 0.1. \tag{E1}$$

$$y_{k+1} = 0.3y_k + 0.04y_k \sum_{j=0}^{n-1} y_{k-j} + 1.5S_{\text{in},k-n+1}S_{\text{in},k} + 0.1. \tag{E2}$$

$$L_{\text{NMSE}} = \frac{\sum_k (y_k - \widehat{y_k})^2}{\sum_k y_k^2}. \tag{E3}$$

The IPC analysis in Fig. 14(d) reveals that the dynamic system requires four IPC components to handle the NARMA2 task. The IPC profile with degree $d_i = 1$ is shown in Fig. 9. We find that the STO RC performs best for this task near the edge of the ESP for input ratios

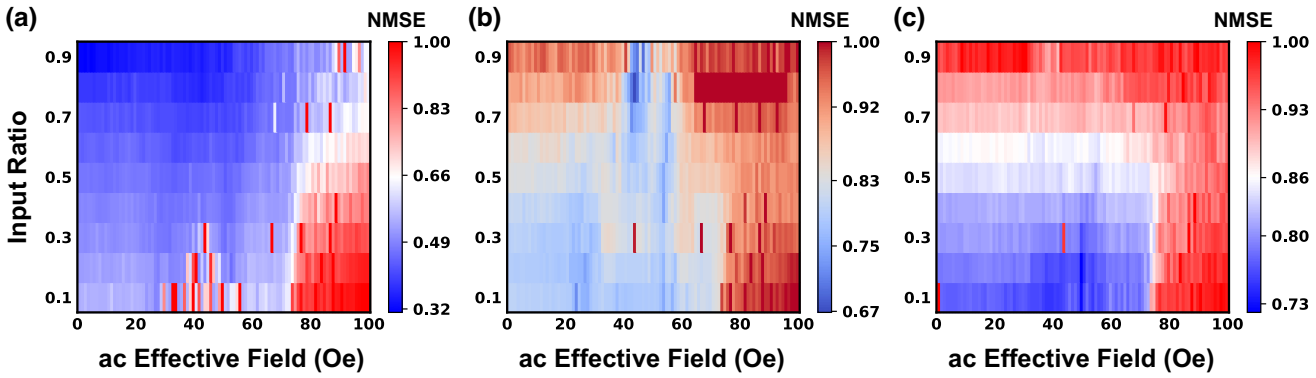

FIG. 13. Results of (a) NARMA2, (b) NARMA3, and (c) NARMA5 tasks. The normalized mean-square error quantifies the computational performance [NMSE, Eq. (S5-3)]. The smaller values of NMSE indicate better computational performance. The number of nodes is 100, and the input width is 4 ns.

range from 0.1 to 0.6, as this region corresponds to the largest values of the four required IPC terms. However, we observe that the $[[d_i, \zeta_{d_i,j}]] = [[1, 3]]$ IPC value gradually increases in the ordered region with an increment of the input ratio. As a result, the best performance of the STO RC is achieved in the ordered region for the input ratio range from 0.8 to 0.9, which is consistent with the results shown in Fig. 13(a). The optimal point in the input ratio of 0.7 is difficult to determine, as the sum of $[[d_i, \zeta_{d_i,j}]] = [[1, 1]]$, $[[d_i, \zeta_{d_i,j}]] = [[1, 2]]$, and $[[d_i, \zeta_{d_i,j}]] = [[1, 3]]$ IPC components is equivalent in most ac regions. Additionally, while the $[[d_i, \zeta_{d_i,j}]] = [[1, 4]]$ IPC component of the STO RC is still the largest at the edge of the ESP, its contribution to the NARMA2 task is minimal.

The IPC analysis also provides insight into the performance of the STO RC in the NARMA3 and NARMA5 tasks. Specifically, we find that the $[[d_i, \zeta_{d_i,j}]] = [[1, 1]]$ and $[[d_i, \zeta_{d_i,j}]] = [[1, 3]]$ IPC values of the STO RC are too small to meet the requirements of the NARMA3 task, which is driven by high ac effective fields. Moreover, for the input ratio of 0.9 with small ac regions, the required IPC mainly comes from degree 2, whose value is very low in STO RC, as shown in Fig. 9. Consequently, the STO RC can handle only this task well near the edge of the ESP

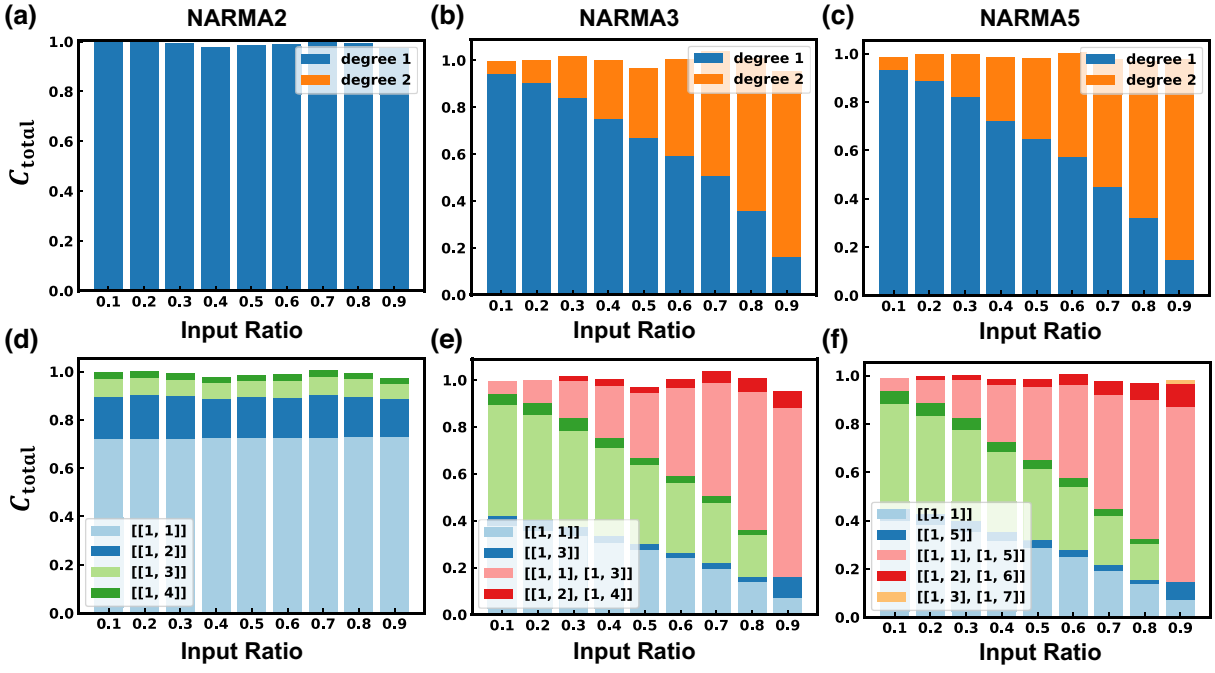

FIG. 14. Ideal IPC components of training signals y_{train} for different NARMA tasks. (a),(d) for NARMA2. (b),(e) for NARMA3. (c),(f) for NARMA5. Note that the finite input length and numerical error may cause the fluctuation of values.

with a high input ratio. Similar results are observed for the NARMA5 task. It is worthwhile to note that the systems cannot be used to handle the NARMA tasks if the NMSEs are too large. Nevertheless, the quantification of the STO RC performance still peaks at the edge of the ESP in these complex tasks.

[1] A. Sebastian, M. Le Gallo, R. Khaddam-Aljameh, and E. Eleftheriou, Memory devices and applications for in-memory computing, Nat. Nanotechnol. **15**, 529 (2020).

[2] Q. Shao, Z. Wang, Y. Zhou, S. Fukami, D. Querlioz, J. J. Yang, Y. Chen, and L. O. Chua, Spintronic memristors for computing , ArXiv:2112.02879.

[3] W. Jiang, L. Chen, K. Zhou, L. Li, Q. Fu, Y. Du, and R. H. Liu, Physical reservoir computing using magnetic skyrmion memristor and spin torque nano-oscillator, Appl. Phys. Lett. **115**, 192403 (2019).

[4] L. Li, L. Chen, R. Liu, and Y. Du, Recent progress on excitation and manipulation of spin-waves in spin Hall nano-oscillators, Chin. Phys. B **29**, 117102 (2020).

[5] S. Watt and M. Kostylev, Reservoir Computing Using a Spin-Wave Delay-Line Active-Ring Resonator Based on Yttrium-Iron-Garnet Film, Phys. Rev. Appl. **13**, 034057 (2020).

[6] R. V. Ababei, M. O. A. Ellis, I. T. Vidamour, D. S. Devadasan, D. A. Allwood, E. Vasilaki, and T. J. Hayward, Neuromorphic computation with a single magnetic domain wall, Sci. Rep. **11**, 15587 (2021).

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

[8] 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).

[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] Q. Zheng, X. Zhu, Y. Mi, Z. Yuan, and K. Xia, Recurrent neural networks made of magnetic tunnel junctions, AIP Adv. **10**, 025116 (2020).

[11] J. Grollier, D. Querlioz, K. Y. Camsari, K. Everschor-Sitte, S. Fukami, and M. D. Stiles, Neuromorphic spintronics, Nat. Electron. **3**, 360 (2020).

[12] D. Marković, N. Leroux, M. Riou, F. Abreu Araujo, J. Torrejon, D. Querlioz, A. Fukushima, S. Yuasa, J. Trastoy, P. Bortolotti, and J. Grollier, Reservoir computing with the frequency, phase, and amplitude of spin-torque nano-oscillators, Appl. Phys. Lett. **114**, 012409 (2019).

[13] T. Kanao, H. Suto, K. Mizushima, H. Goto, T. Tanamoto, and T. Nagasawa, Reservoir Computing on Spin-Torque Oscillator Array, Phys. Rev. Appl. **12**, 024052 (2019).

[14] M. Riou, F. A. Araujo, J. Torrejon, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, and J. Grollier, in *2017 IEEE International Electron Devices Meeting (IEDM)* (2017), p. 36.3.1.

[15] H. Jaeger, *The "echo State" Approach to Analysing and Training Recurrent Neural Networks.*, Tech. Rep. 148 (institution German National Research Center for Information Technology GMD Technical Report, Bonn, Germany, 2001).

[16] K. Nakajima, Physical reservoir computing—an introductory perspective, Jpn. J. Appl. Phys. **59**, 060501 (2020).

[17] J. Hochstetter, R. Zhu, A. Loeffler, A. Diaz-Alvarez, T. Nakayama, and Z. Kuncic, Avalanches and edge-of-chaos learning in neuromorphic nanowire networks, Nat. Commun. **12**, 4008 (2021).

[18] N. Bertschinger and T. Natschläger, Real-time computation at the edge of chaos in recurrent neural networks, Neural Comput. **16**, 1413 (2004).

[19] R. Legenstein and W. Maass, Edge of chaos and prediction of computational performance for neural circuit models, Neural Netw. **20**, 323 (2007).

[20] S. Tsunegi, T. Kubota, A. Kamimaki, J. Grollier, V. Cros, K. Yakushiji, A. Fukushima, S. Yuasa, H. Kubota, K. Nakajima, and T. Taniguchi, Information processing capacity of spintronic oscillator, Advanced Intelligent Systems, 2300175 (2023),.

[21] T. Yamaguchi, S. Tsunegi, K. Nakajima, and T. Taniguchi, Computational capability for physical reservoir computing using a spin-torque oscillator with two free layers, Phys. Rev. B **107**, 054406 (2023).

[22] K. Nakajima and I. Fischer, eds., *Reservoir Computing: Theory, Physical Implementations, and Applications*, Natural Computing Series (Springer Singapore, Singapore, 2021).

[23] S. I. Kiselev, J. C. Sankey, I. N. Krivorotov, N. C. Emley, R. J. Schoelkopf, R. A. Buhrman, and D. C. Ralph, Microwave oscillations of a nanomagnet driven by a spin-polarized current, Nature **425**, 380 (2003).

[24] M. Acet, Ll. Mañosa, and A. Planes, in *Handbook of Magnetic Materials*, Vol. 19, edited by K. H. J. Buschow (Elsevier, 2011), p. 231.

[25] J. C. Slonczewski, Current-driven excitation of magnetic multilayers, J. Magn. Magn. Mater **159**, L1 (1996).

[26] D. Ralph and M. Stiles, Spin transfer torques, J. Magn. Magn. Mater **320**, 1190 (2008).

[27] Z. Li, Y. C. Li, and S. Zhang, Dynamic magnetization states of a spin valve in the presence of dc and ac currents: Synchronization, modification, and chaos, Phys. Rev. B **74**, 054417 (2006).

[28] U. D. Schiller and J. J. Steil, Analyzing the weight dynamics of recurrent learning algorithms, Neurocomputing **63**, 5 (2005).

[29] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, Information processing using a single dynamical node as complex system, Nat. Commun. **2**, 468 (2011).

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

[31] N. Akashi, T. Yamaguchi, S. Tsunegi, T. Taniguchi, M. Nishida, R. Sakurai, Y. Wakao, and K. Nakajima,

Input-driven bifurcations and information processing capacity in spintronics reservoirs, Phys. Rev. Res. **2**, 043303 (2020).

[32] T. Yamaguchi, N. Akashi, K. Nakajima, S. Tsunegi, H. Kubota, and T. Taniguchi, Synchronization and chaos in a spin-torque oscillator with a perpendicularly magnetized free layer, Phys. Rev. B **100**, 224422 (2019).

[33] T. Yamaguchi, N. Akashi, S. Tsunegi, H. Kubota, K. Nakajima, and T. Taniguchi, Periodic structure of memory function in spintronics reservoir with feedback current, Phys. Rev. Res. **2**, 023389 (2020).

[34] M. Balcerzak, D. Pikunov, and A. Dabrowski, The fastest, simplified method of Lyapunov exponents spectrum estimation for continuous-time dynamical systems, Nonlinear Dyn. **94**, 3053 (2018).

[35] Z. Lu, B. R. Hunt, and E. Ott, Attractor reconstruction by machine learning, Chaos **28**, 061104 (2018).

[36] C. Gallicchio, in *ES2019-76* (Bruges (Belgium), 2019).

[37] K. Tanaka, Y. Tokudome, Y. Minami, S. Honda, T. Nakajima, K. Takei, and K. Nakajima, Self-organization of remote reservoirs: Transferring computation to spatially distant locations, Adv. Intell. Syst. **4**, 2100166 (2022).

[38] Z. Yang, S. Zhang, and Y. C. Li, Chaotic Dynamics of Spin-Valve Oscillators, Phys. Rev. Lett. **99**, 134101 (2007).

[39] J. C. Sprott, *Chaos and Time-Series Analysis* (Oxford University Press, Oxford, 2003).

[40] L. M. Pecora and T. L. Carroll, Synchronization in Chaotic Systems, Phys. Rev. Lett. **64**, 821 (1990).

[41] J. Boedecker, O. Obst, J. T. Lizier, N. M. Mayer, and M. Asada, Information processing in echo state networks at the edge of chaos, Theory Biosci. **131**, 205 (2012).

[42] B. Cramer, D. Stöckel, M. Kreft, M. Wibral, J. Schemmel, K. Meier, and V. Priesemann, Control of criticality and computation in spiking neuromorphic networks with plasticity, Nat. Commun. **11**, 2853 (2020).

[43] L. Barnett, J. T. Lizier, M. Harré, A. K. Seth, and T. Bossomaier, Information Flow in a Kinetic Ising Model Peaks in the Disordered Phase, Phys. Rev. Lett. **111**, 177203 (2013).

[44] T. Toyoizumi, Beyond the edge of chaos: Amplification and temporal integration by recurrent networks in the chaotic regime, Phys. Rev. E **84**, 051908 (2011).

[45] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, The information dynamics of phase transitions in random Boolean networks, ALife **XI**, 374 (2008).

[46] T. Kubota, Unifying framework for information processing in stochastically driven dynamical systems, Phys. Rev. Res. **3**, 043135 (2021).

[47] J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, Information processing capacity of dynamical systems, Sci. Rep. **2**, 514 (2012).

[48] J. Schuecker, Optimal Sequence Memory in Driven Random Networks, Phys. Rev. X **8**, 041029 (2018).

[49] T. Haruna and K. Nakajima, Optimal short-term memory before the edge of chaos in driven random recurrent networks, Phys. Rev. E **100**, 062312 (2019).

[50] T. Kubota, Y. Suzuki, S. Kobayashi, Q. H. Tran, N. Yamamoto, and K. Nakajima, Temporal information processing induced by quantum noise, Phys. Rev. Res. **5**, 023057 (2023).

[51] W. H. Rippard, M. R. Pufall, S. Kaka, T. J. Silva, S. E. Russek, and J. A. Katine, Injection Locking and Phase Control of Spin Transfer Nano-oscillators, Phys. Rev. Lett. **95**, 067203 (2005).

[52] M. Mitchell, P. Hraber, and J. P. Crutchfield, Revisiting the Edge of Chaos: Evolving Cellular Automata to Perform Computations, Complex Syst. 89 (1993).

[53] A. Atiya and A. Parlos, New results on recurrent network training: Unifying the algorithms and accelerating convergence, IEEE Trans. Neural Netw. **11**, 697 (2000).

[54] K. Goto, K. Nakajima, and H. Notsu, Twin vortex computer in fluid flow, New J. Phys. **23**, 063051 (2021).

[55] J. D. Costa, S. Serrano-Guisan, B. Lacoste, A. S. Jenkins, T. Böhnert, M. Tarequzzaman, J. Borme, F. L. Deepak, E. Paz, J. Ventura, R. Ferreira, and P. P. Freitas, High power and low critical current density spin transfer torque nano-oscillators using MgO barriers with intermediate thickness, Sci. Rep. **7**, 7237 (2017).

[56] K. Zhang, D. Zhang, C. Wang, L. Zeng, Y. Wang, and W. Zhao, Compact modeling and analysis of voltage-gated spin-orbit torque magnetic tunnel junction, IEEE Access **8**, 50792 (2020).