Spin waves propagating through a stripe magnetic domain structure and their applications to reservoir computing

Ryosho Nakane [1], Akira Hirose [1], and Gouhei Tanaka [1,2]

[1]*Department of Electronic Engineering and Information Systems, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-8656, Japan*

[2]*International Research Center for Neurointelligence (IRCN), The University of Tokyo, 7-3-1 Hongo Bunkyo-ku, Tokyo 113-0033, Japan*

Spin waves propagating through a stripe domain structure and reservoir computing with their spin dynamics have been numerically studied, focusing on the relation between physical phenomena and computing capabilities. Our system utilizes a spin-wave-based device that has a continuous magnetic garnet film and one-input/72-output electrodes on top. To control spatially distributed spin dynamics, a stripe magnetic domain structure and amplitude-modulated triangular input waves were used. The spatially arranged electrodes detected spin vector outputs with various nonlinear characteristics that were leveraged for reservoir computing. By moderately suppressing nonlinear phenomena, our system achieves 100% prediction accuracy in temporal exclusive-OR problems with a delay step up to 5. At the same time, it shows perfect inference in delay tasks with a delay step more than 7 and its memory capacity has a maximum value of 21. This study demonstrated that our spin-wave-based reservoir computing has a high potential for edge-computing applications and also can offer a rich opportunity for further understanding the underlying nonlinear physics.

I. INTRODUCTION

Reservoir computing is a computational framework which is originally based on recurrent neural networks [1,2]. It is realized with a system having a reservoir part and a readout part. In this computational framework, the role of the reservoir part is to nonlinearly transform time-series input data to high-dimensional spatiotemporal signals, which allows us to optimize the readout part with linear regression while the reservoir part is unchanged. Owing to its unique feature, reservoir computing models require much less training cost than deep neural networks. Thus, it is promising for machine-learning-based edge computing [3,4] that can perform energy-efficient information processing of a time-series data obtained from mobile devices and sensors. Recently, it has been demonstrated that reservoir computing systems can be realized with reservoirs based on physical phenomena. Such reservoir computing, called physical reservoir computing [5], has the above-described valuable features as well as feasibility for machine-learning electronic devices: Since the role of the reservoir can be realized by appropriately controlled physical phenomena in feasible hardware structures, it is unnecessary to implement large and complex structures, such as a network structure with massive numbers of internal units and connections. Thus, physical reservoir computing has

a high potential for revolutionary hardware technologies in the next-generation Internet of Things era.

The following examples are some of the typical systems and devices used for physical reservoir computing: optelectronic systems [6,7], optical systems [8,9], an electronic circuit [10], memristive resistance networks [11,12], a soft material [13], water in a bucket [14], a silicon beam [15], a ferroelectric metal-oxide-semiconductor field-effect transistor [16], spin torque oscillators [17–19], a magnetic tunnel junction [20], nanomagnets [21], and spin waves [22,23]. Although the relation between physical properties and computation capabilities in these physical systems has not been fully understood, physical reservoir computing can be performed well when a physical system possesses rich dynamics including high-dimensionality, nonlinearity, input-history dependency, and fading memory property (echo state property) [5]. In particular, high capabilities in extremely efficient information processing are expected for excitable continuous medium reservoirs utilizing propagation of waves triggered by stimulation inputs [14], without internal wiring. In these reservoirs, a high dimensionality can be realized by large numbers of spatially arranged inputs and/or detectors for input/output signals. To take advantage of this characteristic, it is primarily important to excite, control, and detect waves to have rich dynamics.

In our recent papers [22,23], a spin-wave-based reservoir computing device has been proposed as an *on-chip* excitable continuous medium reservoir and its application to machine-learning computation has been demonstrated. The spin-wave-based reservoir computing device is fairly attractive for applications in edge domains, since it can be feasibly realized on a chip with electrical wires just for input/output
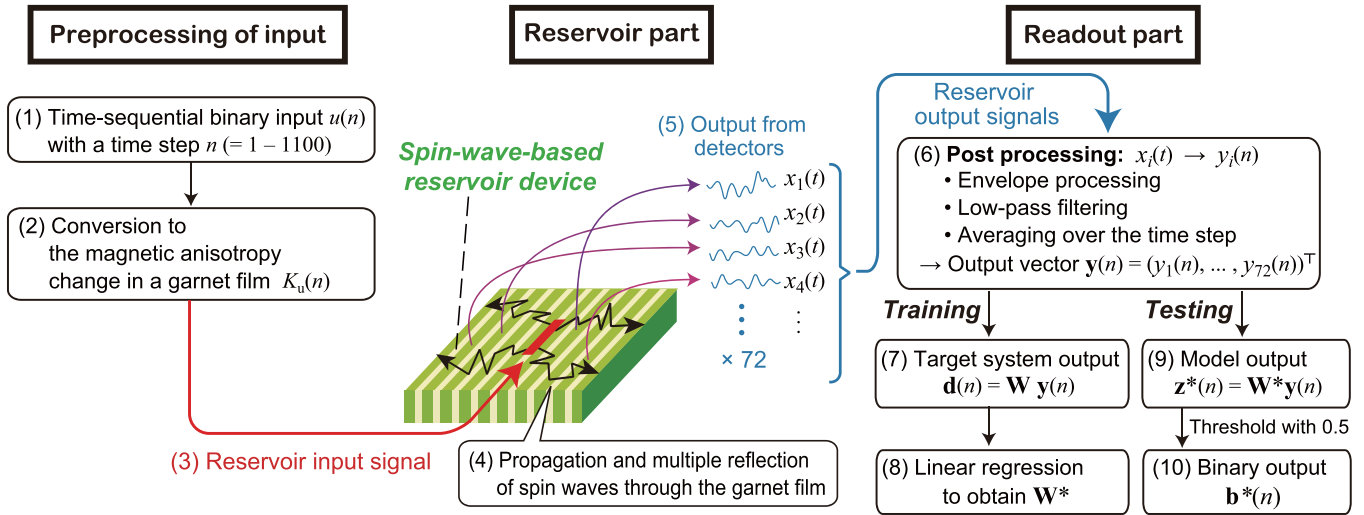
FIG. 1. Spin-wave-based reservoir computing system that is composed of an input preprocessing, a reservoir part, and a readout part. The computing is executed following the order of (1)–(8) in the training, whereas it is executed following the order of (3)–(6) and then (9)–(10) in the testing. In a spin-wave-based reservoir device in the reservoir part, a red area is the input exciter, a zebra pattern with pale and deep greens represents a stripe magnetic domain structure, and bended black arrows schematically illustrate the expected propagation of spin waves with multiple reflections by the magnetic domains. Blue curves schematically illustrate time-series output wave forms obtained at 72 detectors in the device

electrodes and its power consumption is expected to be low because of the signal transmission using spin waves [24]. Moreover, since spin waves originate from the dynamical change in the spin distribution that is one of the electronic properties, highly tolerant computing is expected. Toward practical applications, steady progress in computing capabilities can be made by starting with simple benchmark tasks: temporal exclusive-OR (XOR) problems and delay tasks for evaluating memory capacity. For the reservoir system, the former task requires both nonlinearity and short-term memory, whereas the latter task requires only short-term memory. Since these characteristics play crucial roles in reservoir computing, the achievements of high computing capabilities for these tasks clearly show the high potential of the system. Furthermore, a comprehensive study on spin-wave-based reservoir computing can provide deep insight into the relation between physical phenomena and computing capabilities, which is one main interest in the field of physical reservoir computing.

In this paper, we numerically study reservoir computing using a spin-wave-based reservoir device. To efficiently realize spatially distributed rich dynamics of spin waves with nonlinear phenomena, including nonlinear propagation and multiple reflections, amplitude-modulated waves representing a bit sequence are used for input signals and a stripe magnetic domain structure is introduced in a continuous magnetic garnet film where spin waves propagate. It appears that various wave forms are obtained at different positions in the magnetic garnet film by controlling nonlinear phenomena with input and material parameters. Then, using the resultant wave forms under various parameter conditions, benchmark tasks are solved to reveal what features of spin waves are effective for high-capability reservoir computing. In temporal XOR problems with various interval time steps, it is found that the prediction accuracy strongly depends on the above two parameters. Under moderate suppression of nonlinear phenomena

in spin waves, 100% prediction accuracy is achieved for relatively large delay steps up to 5. At the same time, the memory capacity estimated using delay tasks with various delay time steps has a maximum value of 21. Finally, the relation between physical phenomena and computing capabilities is discussed.

II. RESERVOIR COMPUTING SYSTEM

Spin-wave-based reservoir computing was performed with simulator-calculated spin waves. Figure 1 shows a schematic of our reservoir computing system that consists of a signal preprocessing part, a spin-wave-based reservoir part, and a readout part. In this study, the discrete time step is expressed by $n (= 1, 2, \cdots, 1100)$ and the input data is expressed by $u(n)$ that has a bit value (0 or 1) changing randomly with n.

First, in the preprocessing part, $u(n)$ is transformed to a reservoir input signal that is the change in the uniaxial magnetic anisotropy $K_u(n)$ of the magnetic garnet film in the reservoir device. When a time-series reservoir input signal $K_u(n)$ is fed into the input exciter of the magnetic garnet film (the red line), spin waves are excited and then propagated. Since the magnetic garnet film has a stripe domain structure illustrated by pale and very-pale green stripes, as shown later, it is expected that spin waves reflected multiple times interfere with each other, as illustrated by black bended arrows. Reservoir output signals $x_i(t)$ $(i = 1, 2, \cdots, 72)$ in response to the reservoir input signal are obtained by detecting spin waves at the ith detector on the magnetic garnet film, where t represents continuous time. To use steady response, the output signals in the first 100 time steps (corresponding to the first 100 bits) are discarded in the readout processing.

In the readout part, $x_i(t)$ are converted to $y_i(n)$ through some signal processing: an envelope processing, a low-pass filtering, and an averaging over the time step range. Then, an optimization of the readout part is performed based on a linear

regression [25]. By collecting $y_i(n)$ for $i = 1, 2, \ldots, N_y$ with $N_y \leqslant 72$, a reservoir output vector $\mathbf{y}(n)$ is given by

$$\mathbf{y}(n) = (y_1(n), \ldots, y_{N_y}(n))^\mathsf{T} \in \mathbb{R}^{N_y}. \tag{1}$$

The system output at the kth node is expressed by $z_k(n)$ and then a system output vector $\mathbf{z}(n)$ is given by

$$\mathbf{z}(n) = (z_1(n), \ldots, z_{N_z}(n))^\mathsf{T} \in \mathbb{R}^{N_z}. \tag{2}$$

The system output is expressed by

$$\mathbf{z}(n) = \mathbf{W}^{\text{out}}\mathbf{y}(n), \tag{3}$$

where $\mathbf{W}^{\text{out}}$ is an output weight matrix:

$$\mathbf{W}^{\text{out}} = \left(w_{ki}^{\text{out}}\right) \in \mathbb{R}^{N_z \times N_y}. \tag{4}$$

After the reservoir outputs $\mathbf{y}(n)$ from $n = 101$ to 1100 are divided into two parts, 500 steps in the first part are used for training. Then, the time steps from $601 + L_0$ to 1098 (i.e. totally L time steps) in the second part are used for testing, where L_0 corresponds to the maximum delay step in each task. The initial L_0 steps are discarded to use the common $\mathbf{y}(n)$ for all the delay steps and the last two steps are discarded to avoid artifacts in the envelope processing. In the training, the target signal $d_k(n)$ for the kth system output in each task is generated using $u(n)$ and then the target of the system output vector $\mathbf{d}(n)$ is given by

$$\mathbf{d}(n) = (d_1(n), \ldots, d_{N_z}(n))^\mathsf{T} \in \mathbb{R}^{N_z}. \tag{5}$$

In the training, to determine the optimum weight matrix $\mathbf{W}^*$, $\mathbf{Y}$ and $\mathbf{D}$ matrices are given by column-wise collections of $\mathbf{y}(n)$ and $\mathbf{d}(n)$, respectively, as follows:

$$\mathbf{Y} = [\mathbf{y}(101), \ldots, \mathbf{y}(600)] \in \mathbb{R}^{N_y \times 500}, \tag{6}$$

$$\mathbf{D} = [\mathbf{d}(101), \ldots, \mathbf{d}(600)] \in \mathbb{R}^{N_z \times 500}. \tag{7}$$

Then, $\mathbf{W}^*$ is calculated by using the pseudoinverse matrix $\mathbf{Y}^\dagger$:

$$\mathbf{W}^* = \mathbf{D}\mathbf{Y}^\dagger. \tag{8}$$

In the testing, a model output signal vector $\mathbf{z}^*(n)$ is calculated using $\mathbf{W}^*$ and $\mathbf{y}(n)$ as follows:

$$\mathbf{z}^*(n) = \mathbf{W}^*\mathbf{y}(n). \tag{9}$$

Then, a binary output signal $b_k^*(n)$ at the kth node is calculated from $z_k^*(n)$ with a threshold of 0.5 as follows:

$$b_k^*(n) = \begin{cases} 1 & (z_k^*(n) \geqslant 0.5) \\ 0 & (z_k^*(n) < 0.5). \end{cases}$$

For evaluating the computing capability with the system output at the kth node, the prediction accuracy of a computing task in the testing is calculated using the normalized Hamming distance as follows:

$$\text{acc} = 1 - \sum_n \frac{|y_k(n) - b_k^*(n)|}{L}, \tag{10}$$

where the second term is the normalized Hamming distance. Here, L for temporal XOR problems is 474 and that for delay tasks is 426.
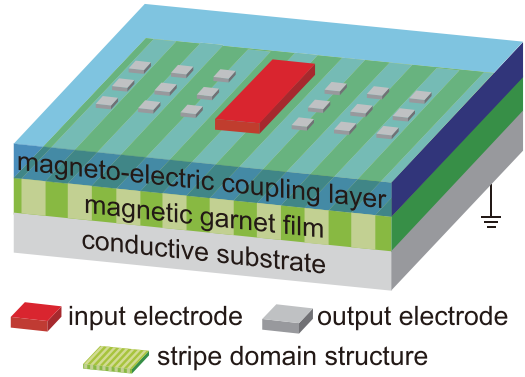

FIG. 2. Schematic illustration of a spin-wave-based reservoir computing device that is composed of (from the bottom to top) a conductive substrate, a magnetic garnet film, a magnetoelectric coupling layer, and input (a red cuboid)/output (gray cylinders) electrodes. The stripes in the magnetic garnet film denote a stripe magnetic domain structure. In the operation of an actual device, spin waves in the magnetic garnet film are excited by an input voltage, they propagate through the magnetic garnet film, and their dynamics beneath the output electrodes are detected by output voltages.

III. SPIN-WAVE-BASED RESERVOIR COMPUTING DEVICE

A. Device structure

Figure 2 shows a schematic device structure that is composed of (from the bottom to top) a conductive substrate, a magnetic garnet film, a magnetoelectric (ME) coupling layer, and input (a red cuboid)/output (gray cuboids) electrodes. It is basically the same as that in our previous paper [22] and the main differences are the shape of the input electrode, the arrangement of the output electrodes, and the stripe magnetic domain structure. The magnetic garnet film/ME coupling bilayer structure is typically called a synthetic multiferroic structure that converts from magnetic to electrical properties, and vice versa.

In the operation of an actual device, spin waves in the magnetic garnet film are excited by an input voltage, they propagate through the magnetic garnet film, and their dynamics beneath the output electrodes are detected by output voltages. The exchanges from spin to voltage, and vice versa, are realized through the function of the synthetic multiferroic structure, such as voltage controlled magnetic anisotropy (VCMA) [26].

In numerical experiments in this study, since spin dynamics in the magnetic garnet film is simulated, input and output are expressed by the change in the properties of the magnetic garnet film near the surface, as a typical ME coupling in synthetic multiferroic structures. The input is expressed by the change in the magnetic anisotropy near the surface. The output is expressed by the out-of-plane component of the averaged spins over a small volume beneath each output electrode since the output voltage is expected to be proportional to the out-of-plane component of spins near the surface.

B. Key ideas for high performance in reservoir computing

The following are our considerations on what physical phenomena are specifically needed in our spin-wave-based reservoir computing device:

(1) High-dimensionality: Different wave forms are obtained at different positions in space.

(2) Nonlinearity: Nonlinear phenomena, such as nonlinear propagation and interference, occur in response to a time-series input signal.

(3) Input-history dependency and fading memory property: Spin waves excited by an input signal that corresponds to one bit information causes fluctuations at the same position for a certain time length.

In physical reservoir computing, the reservoir part appropriately designed in advance is unchanged. Thus, the input transformation from a time-series data to time-series physical quantities and the configuration of a reservoir device are essential for the realization of rich physical dynamics. Our key ideas for these two building blocks are described below.

Regarding input signals, our idea is to use amplitude-modulated waves corresponding to bit sequence data. This is because transient nonlinear responses to input signals are effective for reservoir computing [10] and such responses excited by forced oscillations can be reliably measured in the signal envelope, as in spin torque oscillators [17]. When the input $u(n)$ is transformed to a reservoir input signal, the time period of the modulation signal is an important parameter, since it is expected that nonlinear transient phenomena strongly depend on it. Thus, this study examines how spin dynamics changes depending on this time period.

Regarding the configuration of a reservoir device, our idea is to realize unstable spin waves in a stripe domain structure that is frequently obtained in magnetic garnet films under near-zero external DC magnetic field H^{EX} [27]. Here, unstable spin waves is used to indicate a situation that long-range dipole interactions are not negligible due to large polar angles in spin precessions and the precession axis for each spin dynamically changes. From previous papers by other groups [28], unstable spin waves can lead to rich dynamics, including period-doubling bifurcations and chaotic behavior, when a high-power radio-frequency input is used. Instead of this method, when a stripe domain structure with a small H^{EX} is used, the following features are expected: Unstable spin waves result in space-varying spin dynamics, they show nonlinear phenomena, and they are split and reflected multiple times in the presence of domains (and domain boundaries). These can lead to the aforementioned physical characteristics needed for reservoir computing. Since the properties of the magnetic garnet film are important parameters, this study examines how spin dynamics changes depending on the damping factor at the boundaries of the garnet film, as will be specifically explained in Sec. III C. It is noteworthy that a small H^{EX} can be realized by an available technology for on-chip devices [29], which is advantageous for feasible implementation.

C. Simulation procedure and material parameters

Figures 3(a) and 3(b) show top-schematic and cross-sectional-schematic views of a magnetic garnet film used for the simulation, respectively, where red and blue regions are
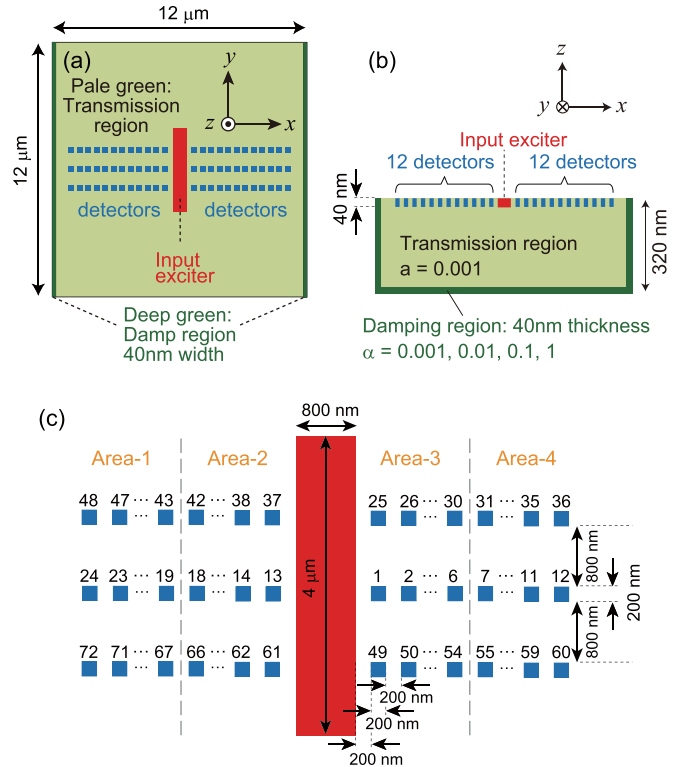

FIG. 3. (a) Top and (b) cross-sectional views of the spin-wave-based reservoir computing device utilizing a magnetic garnet film, where a Cartesian coordinate is defined in each figure, pale-green and deep-green regions are the transmission and damping regions of spin waves in the garnet film, respectively, and red and blue regions are the input exciter and detectors, respectively. The damping constant α_0 in the transmission region is 0.001, whereas the damping constant α in the damping regions is varied ($\alpha = 0.001, 0.01, 0.1,$ and 1) to examine the change in spin dynamics, as described in Sec. III B. The x-y plane and thickness along the z axis of the magnetic garnet film is $12 \times 12~\mu\mathrm{m}^2$ and 320 nm, respectively. The depths of the input exciter and detectors are 40 nm from the surface, as shown in (b). (c) Detailed top-view structure of the input exciter and detectors on the top of the garnet film, where detectors with $200 \times 200~\mathrm{nm}^2$ are arranged with 400 nm and 1000 nm in pitch along the x and y axes, respectively, and the number is defined for each detector. All the detectors are divited into four groups: areas 1, 2, 3, and 4, each of which has 18 detectors.

the input and output regions, respectively, and a Cartesian coordinate system is defined in each figure. Hereafter, the input and output region are called the input exciter and detector, respectively. The x-y plane has an area of $12 \times 12~\mu\mathrm{m}^2$ and a thickness of 320 nm along the z direction, where pale-green and deep-green regions are the transmission and damping regions of spin waves, respectively, and the transmission region includes the input exciter and detectors. The upper and lower boundaries in Fig. 3(a) are connected for reducing the simulation time, which is expected to have no influence on detected spin waves since the detectors are arranged far from these boundaries. Figure 3(c) shows the detailed top-view arrangement of the input exciter (a red rectangle) and detectors (blue rectangles) in the magnetic garnet film, respectively. The input exciter has an area of 800 nm $\times$ 4 μm centered

at the origin of the x-y plane and has a depth of 40 nm along the z axis [Fig. 3(b)]. The detectors have an area of 200 nm $\times$ 200 nm and a depth of 40 nm along the z axis [Fig. 3(b)], they are arranged at a 400 nm pitch interval along the x axis and at a 1000 nm pitch interval along the y axis, respectively, and the detector number is defined in Fig. 3(c). The detectors are divided into four groups, areas 1, 2, 3, and 4, by the broken-gray boundaries along the y axis and each group has 18 detectors. The damping regions at the left and right boundaries are expected to fully damp the spin waves propagating from the input region at the center, i.e., there is no reflection of spin waves from the boundaries to the center. The reflection of spin waves will be verified later.

We used a micromagnetic simulator Mumax3 [30] for numerical experiments on spin dynamics in the magnetic garnet film in Figs. 3(a) and 3(b). The simulation and material parameters are as follows: The mesh unit defined by the xyz axes is cubic with 40 nm on one side, a spin (magnetic moment) is located at every mesh corner, the simulation temperature is 0 K, the saturation magnetization M_s is 190 kA/m, the stiffness constant A_{EX} is 3.7×10^{-12} J/m, the uniaxial magnetic anisotropy K_u along the z axis is $K_u^H = 5$ kJ/m^3 in the damping region and the transmission region except the input exciter, and the damping constant α_0 is 0.001 in the transmission region. The M_s value is that of a bulk $Y_3Fe_5O_{12}$ (YIG) at a few K [31], the A_{EX} value is identical with the value for a liquid-phase epitaxy grown YIG film on a $Gd_3Ga_5O_{12}$ (GGG) substrate [32], and the K_u^H is similar to the values for pulsed laser deposition grown YIG films on GGG substrates [33]. In the input exciter, K_u along the z axis was changed in proportion to an input signal, as will be explained in Sec. III E. To study fundamental properties and reservoir computing with various types of spin dynamics, α in the damping region was varied (= 0.001, 0.01, 0.1, and 1). During the simulation, a constant external magnetic field H^{EX} was applied along the $+y$ direction. The time step in the simulation was 1×10^{-12} s, whereas averaged spin values for a specific detector were recorded at every 5×10^{-11} s. Hereafter, the x, y, and z components of the normalized spin are expressed by s_x, s_y, and s_z, respectively.

It should be noted that the α_0 value 0.001 in the transmission region is larger by one order of magnitude than the experimentally estimated values in epitaxial YIG films with high quality [34–37]. The reasons are simply to reduce simulation time and to eliminate the influence of spin waves reflected back from the left and right boundaries by intentionally reducing the amplitudes of spin waves while propagating (the latter will be discussed in VI A). Thus, since this α_0 value just leads to smaller decay lengths of propagating spin waves, consistent results with those in this paper will be obtained when simulation is performed using a YIG film having a larger x-y plane and a smaller α_0 value.

D. Magnetic domain structure

The initial magnetic domain structure for spin dynamics simulation was formed using the following procedure: First, $\mu_0 H^{EX} = 0.05$ T was applied, then it was reduced to 0.01 T in step of 0.01 T and, finally, it was reduced to 0.005 T in step of 0.001 T. At each step, the spin distribution was relaxed.
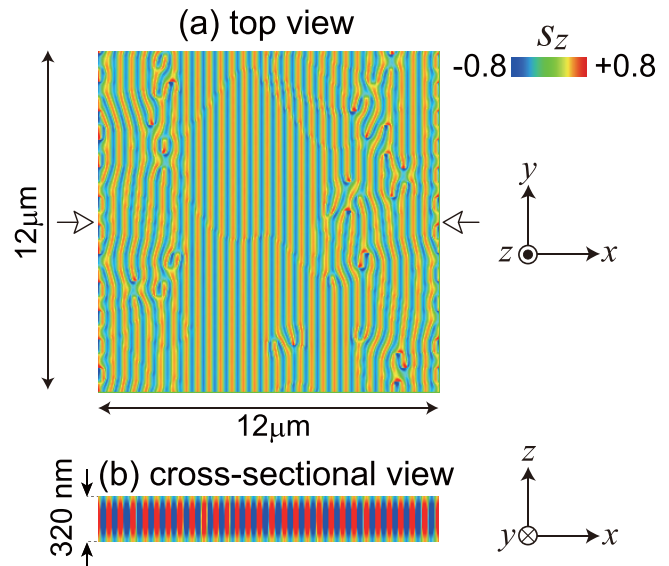

FIG. 4. (a) Top and (b) cross-sectional views of the distributions of s_z in the magnetic garnet film after the relaxation under a magnetic field $\mu_0 H^{EX} = 0.005$ T along the $+y$ direction, where the magnitude is defined by a color bar. The cross-sectional view is along the x axis between two open arrows and it is extended along the z axis for visibility.

Figures 4(a) and 4(b) show top and cross-sectional views of the distributions of s_z in the magnetic garnet film after the final step, respectively, where the magnitude is defined by a color bar and the cross-sectional view is extended along the z axis for visibility. The magnetic domain has a checkered pattern composed of red and blue lines along the y axis, i.e., a stripe domain structure with two spin directions along the $\pm z$ directions. The deviation from the complete stripe structure increases as the x position approaches the left and right boundaries. This feature is preferable for reservoir computing because it can lead to various wave forms at the detectors.

E. Reservoir input signals

In the preprocessing part in Fig. 1, $u(n)$ was transformed to the reservoir input signal $K_u(n)$ as follows. First, the time lengths of a triangular pulse and the time step length X were set at 1 ns and a multiple of 1 ns, respectively. As described in Sec. III B, X is the time period of the modulation signal, which is an important parameter for controlling spin dynamics. To analyze the effect of X on spin dynamics as well as on the capabilities in reservoir computing, responses to various X values were examined: $X = 1, 2, 4, 6,$ and 8 ns. The repetition number of triangular pulses for one bit can be calculated such that X ns is divided by 1 ns. For example, when $X = 4$ ns, the repetition number is 4. Next, to change the amplitude of ΔK_u following the input data, $K_u^M = 4.5$ and $K_u^L = 4$ kJ/m^3 were prepared with a use of the maximum $K_u^H = 5.0$ kJ/m^3 as the baseline. As an example, Fig. 5(a) shows a reservoir input signal with $X = 4$ ns, in which 0 and 1 in $u(n)$ are expressed by $\Delta K_{u,0} = K_u^H - K_u^L$ and $\Delta K_{u,1} = K_u^H - K_u^M$, respectively. Finally, time-series reservoir input signals with various X were prepared using the same $u(n)$ for $n = 1, \ldots, 1100$.
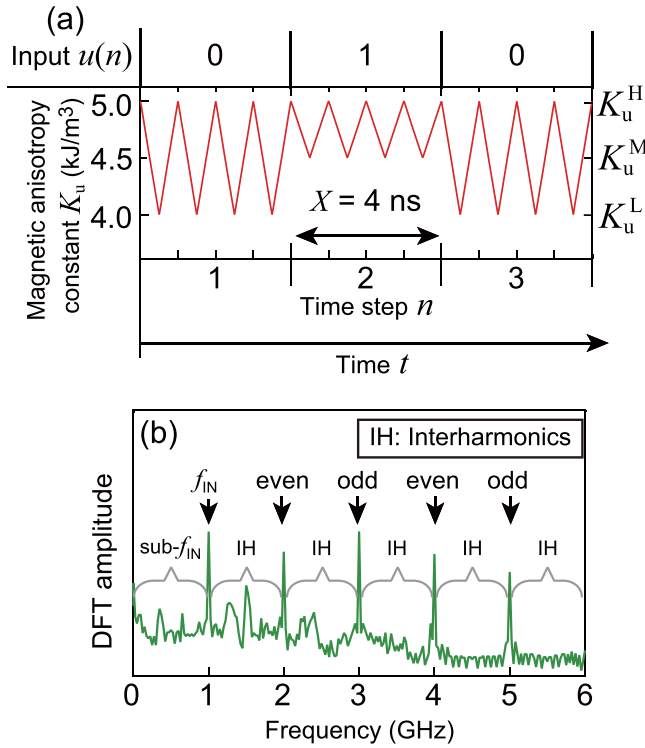
FIG. 5. (a) Example of a reservoir input signal $K_u(n)$ with the time step length $X = 4$ ns: 0 and 1 in $u(n)$ are expressed by $\Delta K_{u,0} = K_u^H - K_u^L$ and $\Delta K_{u,1} = K_u^H - K_u^M$, respectively, with a use of K_u^H as the baseline, where K_u^H, K_u^M, and K_u^L are 5.0, 4.5, and 4.0 kJ/m^3, respectively. To analyze the effect of X on the properties of spin waves as well as on the capability of the reservoir computing, reservoir input signals with various X values were prepared; $X = 1, 2, 4, 6,$ and 8 ns. The repetition number of triangular pulses for one bit can be calculated such that X ns is divided by 1 ns. (b) Schematic illustration of a DFT spectrum, where f_{IN}, odd, and even denote the pulse repetion frequency (1 GHz), the odd-number harmonic frequencies (3 and 5 GHz), and the even-number harmonic frequencies (2 and 4 GHz), respectively, and sub-f_{IN} and IH denote interhamonics below f_{IN} and interharmonics between the interger-number harmonic frequencies, respectively.

F. Analysis of spin dynamics

To clarify the relation between physical phenomena and computing capabilities, it is necessary to analyze how spin dynamics are changed by the values of the two parameters α and X. Hence, reservoir output wave forms at the detectors were observed and they were also characterized by discrete Fourier transform (DFT) spectra that were calculated using output wave forms in the n range from 101 to 1100. In input signals, since a triangular pulse with a period of 1 ns is repeated, namely, a triangular wave with 1 GHz is used as a carrier wave, a DFT spectrum always has a component at 1 GHz. Hereafter, the input signal frequency 1 GHz is denoted by f_{IN}. In the analysis, a DFT spectrum is divided into four parts with different names, following each feature [38], as shown in Fig. 5(b):

(1) Odd-number harmonics: Waves characterized by DFT components at 1, 3, and 5 GHz that are the fundamental frequencies for a triangular pulse.

(2) Even-number harmonics: Waves characterized by DFT components at 2 and 4 GHz.

(3) Interharmonics: Waves characterized by DFT components between the odd- and even-number harmonic frequencies, including the fractional-order harmonics at 1.5, 2.5, 3.5, and 4.5 GHz.

(4) Sub-f_{IN}: Waves characterized by DFT components below 1 GHz, including subharmonics, and supersubharmonics. In a broad definition, this part is included in interharmonics.

Since triangular pulses were used in input signals, the odd-number harmonics are linear responses, whereas the others are nonlinear responses due to nonlinear phenomena. Hereafter, the odd-number harmonics and even-number harmonics are sometimes collectively called the integer-number harmonics. To obtain accurate components at the integer-number and fractional-number harmonic frequencies in DFT spectra, the window function in the calculation was rectangle and the time length was the maximum common multiples of the time periods of these harmonics in the n range from 101 to 1100. The reason why the rectangle window function was used is that a time-varying mean value is hardly subtracted from the output signal and it possibly has important components for reservoir computing. It is well recognized that this DFT procedure can lead to components at frequencies other than these harmonic frequencies and that quantitative analyses on sub-f_{IN} and interharmonics have not yet been established [38]. Hence, the analysis with DFT spectra is phenomenological based on how the spectrum feature changes with the parameters X and α.

IV. PHYSICAL PROPERTIES OF SPIN DYNAMICS AT DETECTORS

Figure 6 shows wave forms and DFT spectra obtained with $\alpha = 0.1$ and various X ($= 1, 2, 4, 6,$ and 8 ns), where the unit of the horizontal axis for the wave forms is time step n, red-line plots are results for the reservoir input signal with $X = 4$ ns, and blue-line plots are results for reservoir output signals at detector 1. In the wave forms, the green sections correspond to input value of 0 ($\Delta K_{u,0}$), whereas the rest of the white sections correspond to input value of 1($\Delta K_{u,1}$). In the DFT spectrum of the reservoir input signal in the upper-right panel, whereas the peaks at 1.0, 3.0, and 5.0 GHz are dominant, the background amplitude shows a gradual decrease with increasing frequency as well as a small oscillation with a constant period. The odd-number harmonics are consistent with the Fourier series expansion of sequential triangular pulses and the gradual decrease is simply interpreted as the random change in ΔK_u with n. On the other hand, the period of the small oscillation is 0.25 GHz that corresponds to $1/X = 1/4$ ns^{-1}. Since this relation in oscillation was also confirmed in the DFT spectra for input signals with other X values (not shown here), it is attributable to the change in ΔK_u with the step of X.

In the reservoir output signals plotted with the blue lines, the following features are seen. The prominent feature is that nonlinear phenomena were obtained under all the conditions examined, since all the DFT spectra have peaked components at the even-number harmonic frequencies and components at the sub-f_{IN} and interharmonic frequencies. At $X = 1$ ns, the amplitude of the wave form changes frequently with n and its outline does not apparently respond to the reservoir
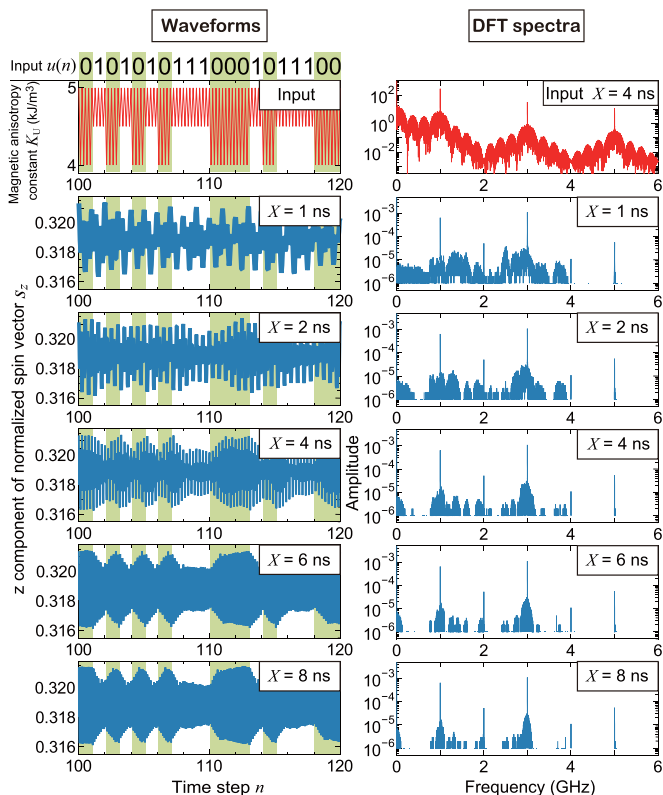
FIG. 6. Red lines in the top two figures: A wave form and a DFT spectrum of input triangular pulses with two modulated amplitudes corresponding to bit sequence data, for which the transformation rule from the bit to amplitude of K_U is shown in Fig. 5(a). On the left-hand side, the unit of the horizontal axis is time step n, i.e., the time normalized by X ns. On the right-hand side, the DFT spectrum calculated for $X = 4$ ns is shown. Blue lines: Wave forms and DFT spectra of the resultant spin waves detected at detector 1 when α = 0.1 and various time steps X (= 1, 2, 4, 6, and 8 ns). In the wave forms, the green ranges correspond to the bit value 0 ($\Delta K_{u,0}$), whereas the rest white ranges correspond to the bit value 1 ($\Delta K_{u,1}$). Note that the widths of the lines for $X = 1$ and 2 ns are broadened compared to those of the others to clearly see the outlines.

input signal at the present time step n. As X is increased, the outline of the wave form more clearly responds to the reservoir input signal at the present time step, in such a way that the amplitude of the wave form increases (decreases) after the bit transition from 1 to 0 (from 0 to 1). In the n range of 107–114 (corresponds to an input sequence of [0, 1, 1, 1, 0, 0, 0]), the wave forms for $X = 2$, 4, 6, and 8 ns show an almost complete change from the maximum to minimum amplitudes, and vice versa, with similar relaxation times. In the n range of 100–107 (corresponds to an input sequence of [0, 1, 0, 1, 0, 1, 0]), the outlines of the wave forms slightly differ from each other. Considering the similar relaxation times in the n range of 107–114, the difference in the outline of the wave form between $X = 2$, 4, 6, and 8 ns mainly arises from the difference in the time step length X. On the other hand, the DFT spectrum changes with X. As X is increased, the sub-f_{IN} and interharmonic components decrease, whereas the integer-number harmonic components remain almost unchanged. These features indicate that the most change in spin
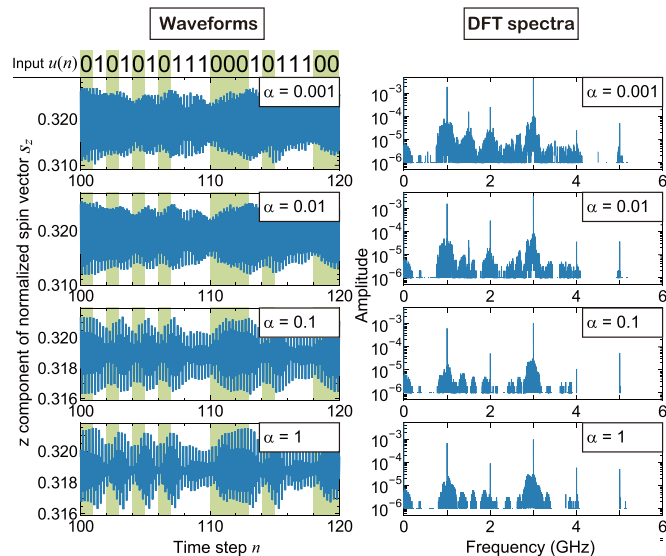

FIG. 7. Wave forms and DFT spectra of reservoir output signals obtained at detector 1 when various α values (= 0.001, 0.01, 0.1, and 1) and $X = 4$ ns.

dynamics originates from nonlinear phenomena characterized by the sub-f_{IN} and interharmonic components.

Figure 7 shows wave forms and DFT spectra obtained with various α values (= 0.001, 0.01, 0.1, and 1) and $X = 4$ ns, where blue lines are the results for reservoir output signals at detector 1. In the n range of 100–107 (corresponds to an input sequence of [0, 1, 0, 1, 0, 1, 0]), the outline of the wave form responds to the present input signal more clearly with increasing α, while the amplitude becomes smaller. In the DFT spectra, when the sub-f_{IN} components and the relative ratios between the integer-number harmonic components are observed, they change little with α. In contrast, the interharmonic components greatly increase with decreasing α, which is similar to the feature with decreasing X in Fig. 6. However, the spectrum for $\alpha = 0.001$ and $X = 4$ ns in Fig. 7 differs from that for $\alpha = 0.1$ and $X = 1$ ns in Fig. 6. Thus, the effects of decreasing α and X on the reservoir output signal are roughly similar, but they are not identical. This indicates that the change in the spin dynamics by α is different from that by X.

Reservoir output signals at various detectors were observed to confirm the realization of various wave forms. Figure 8 shows wave forms and DFT spectra at detectors 1, 13, and 37 obtained for $\alpha = 0.1$ and $X = 4$ ns. Whereas the three detectors are located at the nearest x position from the input exciter, the y positions are different, as shown in Fig. 3(c). The three wave forms are roughly similar, but not identical. The difference is also characterized by the DFT spectra, where each spectrum has a unique feature that is most simply confirmed in the relative ratios between the integer-number harmonic components. Thus, it was found that the three detectors have different output wave forms even though they are located in the stripe domain structure without significant deviation, as seen in Fig. 4(a). In the same manner, when wave forms and DFT spectra at all the detectors were observed, they were unique for each detector. To understand such
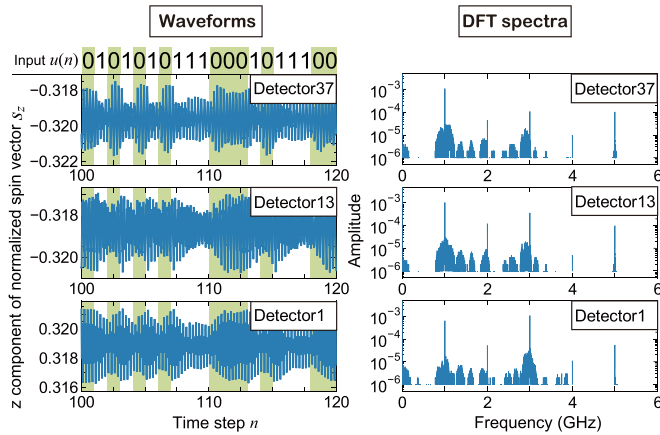
FIG. 8. Wave forms and DFT spectra of reservoir output signals obtained at detectors 37, 13, and 1 when $\alpha = 0.1$ and $X = 4$ ns.

detector-dependent wave forms, the spatial distributions of s_z at different n are also helpful, as shown in Appendix A.

Important phenomena were also found when wave forms at detectors in the same line along the x axis were observed. Figure 9 shows wave forms and DFT spectra at detectors 1, 7, and 12 when $\alpha = 0.1$ and $X = 4$ ns. With the increase of the detector number from 1 to 7 (the increase of the distance from the input exciter), the wave forms more ambiguously respond to the present input signal, which are confirmed by the valleys and peaks of the outline. Moreover, even when the wave form at detector 1 is shifted along the horizontal axis and its amplitude is reduced, a wave form identical with that at detector 7 cannot be created. Thus, it is very probable that the spin waves are nonlinearly changed while propagating due to nonlinear phenomena and reflection with multiple times. In the wave form at detector 12, long-period beats with a width of $\sim$4 ns appear and the relation between the wave form and input signal is unclear. On the other hand, in the DFT spectra, as the detector number increases, the components at higher frequencies preferentially decrease and finally the components at around 1 GHz are present at detector 12. Thus, the change in the wave form with increasing distance from the input exciter partially originates from the feature that spin
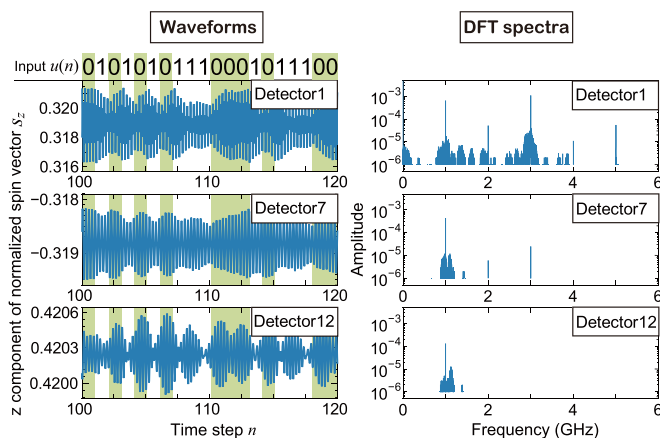

FIG. 9. Wave forms and DFT spectra of reservoir output signals obtained at detectors 1, 7, and 12 when $\alpha = 0.1$ and $X = 4$ ns.

waves propagating in long distances are inevitably determined by the stripe domain structure.

V. POSTPROCESS OF RESERVOIR OUTPUT SIGNALS

In the readout part in Fig. 1, the reservoir output signals $x_i(t) (i = 1, \cdots, 72)$ are converted to a 72-dimensional output vector $\mathbf{y}(n)$ at each time step n. More specifically, $x_i(t)$ in a time step period are summarized into one value that reflects a characteristic response to the input signals from near past to n. As seen in the DFT spectra, the reservoir output signals have oscillations whose fundamental time period is smaller than X and their envelopes are expected to reflect characteristic responses. Moreover, instantaneous changes should be excluded since they can be noise that does not reflect characteristic responses. Based on these considerations, the following procedure of the post signal processing was used, as shown in Figs. 10(a)–10(d). First, an envelope processing and a subsequent low-pass filtering were performed for the reservoir output signals $x_i(t) (i = 1, \cdots, 72)$, as shown in Figs. 10(b) and 10(c). In the low-pass filtering, the cutoff frequency f_c defined by -10 dB is 1.3 GHz. Then, the resultant signal at each detector was averaged over the time period of n, i.e., X, as shown in Fig. 10(d). Finally, the averaged value at detector i was assigned to $y_i(n)(i = 1, \cdots, 72)$ in $\mathbf{y}(n)$. Clearly, the post signal processing extracts information in the low-frequency range $\leqslant f_c$. Nonetheless, spin waves in a wide frequency range are probably related to $\mathbf{y}(n)$ since they can dynamically interact with each other through nonlinear phenomena. It should be noted that the above postprocess procedure is feasible for implementation since each signal processing is a widely used one.

VI. RESERVOIR COMPUTING

In this section, reservoir computing is performed with two main purposes. The first purpose is to find parameter sets of α and X values that achieve high computing capabilities for two different tasks, which is needed for an adjustable reservoir computing system. The second purpose is to explore how the computing capabilities for the two tasks change with the physical properties of the spin waves, for deep understanding of physical reservoir computing. Whereas the specifications (1) and (2) in Sec. III B were basically characterized in Sec. IV, the specification (3) can be evaluated only by computing tasks. Here, what can be revealed by computing tasks is briefly described below.

(1) Temporal XOR problems [39]

These problems are linearly inseparable tasks that need both nonlinearity and fading memory. Since the relation between nonlinear physical phenomena and nonlinearity for computing is unclear, we can discuss what nonlinear physical phenomena are effective for high computing capabilities.

(2) Delay tasks [40]

When the computing capability is plotted against delay step, the curve is directly related to the time length of fading memory in specification (3). Using the index memory capacity estimated from the plot, we can qualitatively discuss how long input information is preserved in the reservoir. Moreover, by combining the computing capability for temporal XOR

(a) Output waveform

(b) Envelope processing

(c) Low-pass filtering
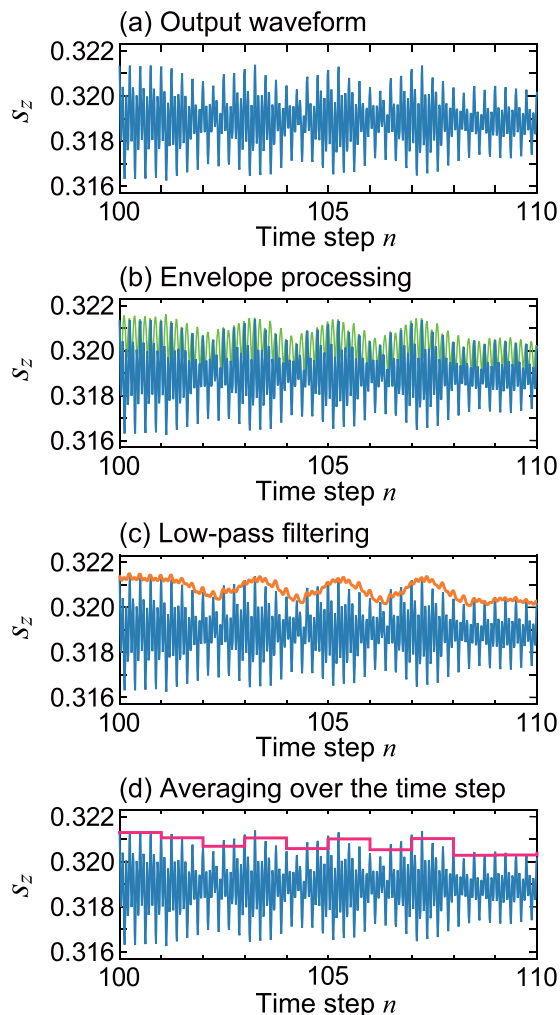

(d) Averaging over the time step

FIG. 10. Postprocessing of a reservoir output signal, which is executed in the order of (a) reservoir output signal, (b) an envelope processing, (c) a low-pass filtering, and (d) averaging over the time step X. Blue lines in (a)−(d) are the same signal obtained at detector 1 when $\alpha = 0.1$ and $X = 4$ ns. Colored lines in (b)−(d) are the signals after the processing. The $x_i(n)$ value in (d) is assigned to $y_i(n)(i = 1, \cdots, 72)$ in $\mathbf{y}(n)$.

problems under a certain parameter set of α and X, it is possible to discuss the relation between nonlinearity and memory capacity, which is one of the interests in the research field of reservoir computing [41].

A. Temporal XOR problems

Reservoir computing was performed using 72-dimensional $\mathbf{y}(n)$ for various conditions with α (= 0.001, 0.01, 0.1, and 1) and X (= 1, 2, 4, 6, and 8 ns). Hereafter, $\mathbf{y}(n)$ has 72 dimensions unless otherwise noted. To evaluate the device capability for nonlinear transformation of the input signal to $\mathbf{y}(n)$, temporal XOR problems were solved with the target output:

$$d_k(n) = \text{XOR}(u(n-1), u(n-k))$$
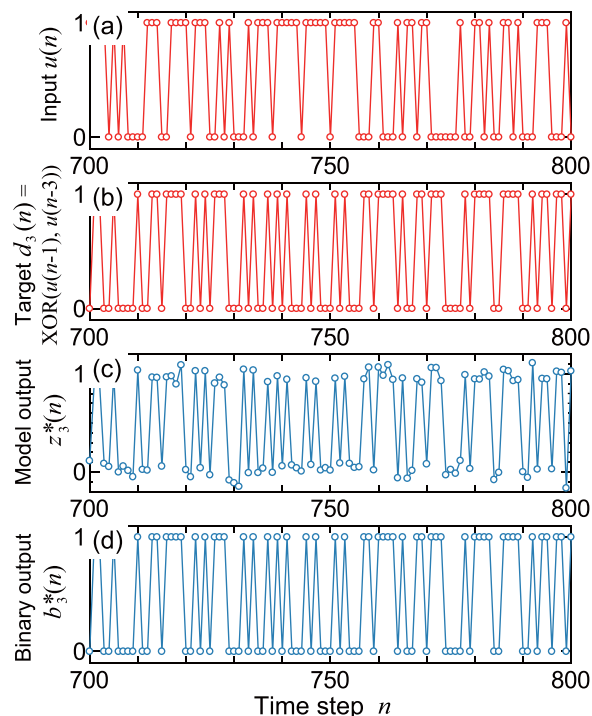
$$\text{for delay } k = 2, \dots, 24. \quad (11)$$

FIG. 11. Singals in solving a XOR problem using 72-dimensional $\mathbf{y}(n)$ with the target output $d_3(n) = \text{XOR}(u(n-1), u(n-3))$ (corresponding to delay $k = 3$) when $\alpha = 0.1$ and $X = 4$ ns. Red open circles are the input binary data and blue open circles represent the calculated values in the readout part. (a) Input $u(n)$ that has a bit value (0 or 1) changing randomly with n. (b) Target signal $d_3(n) = \text{XOR}(u(n-1), u(n-3))$. (c) Model output $z_3^*(n)$. (d) Binary output $b_3^*(n)$ calculated from $z_3^*(n)$ with a threshold of 0.5.

To solve the temporal XOR problem with delay k, the signal transformation by the reservoir requires nonlinearity and a memory of past k inputs. We assume $k \geqslant 2$ to eliminate the trivial case of $k = 1$ where $d_k(n) = 0$ for all n. Here, to evaluate the capability, the index CP is defined by the maximum value of k such that 100% prediction accuracy (acc = 1) is maintained when k is increased. First, signals in the computation flow were observed by one example. Figures 11(a)–11(d) show the input $u(n)$, the target output $d_3(n)$, the model output $z_3^*(n)$, and the binary output $b^*(n)$, respectively, where $\alpha = 0.1$, $X = 4$ ns, open circles are the values at n, and colored lines connecting the nearest-neighbor open circles are guides. In this case, $d_3(n) = b_3^*(n)$ in the entire n range, meaning acc = 1. Figures 12(a)–12(d) show acc calculated for $\alpha = 0.001$, 0.01, 0.1, and 1, respectively, plotted against k, where black, red, blue, brown, and green dots represent the results for $X = 1, 2, 4, 6$, and 8 ns, respectively. When $\alpha = 0.001$ in Fig. 12(a), all the curves do not have acc = 1 and immediately decrease to the minimum acc = 0.5 with increasing k, in which the decreasing slope changes little with X. When α is increased from 0.001 to 0.01, acc for each X becomes larger at $k = 2$ and the decreasing slope in each curve becomes more steep except $X = 1$ ns, as shown in Fig. 12(b). However, since all the curves do not have acc = 1, CP was not obtained. When α is increased from 0.01 to 0.1, the features of the curves are

FIG. 12. Prediction accuracy plotted against delay k, which were estimated by solving XOR problems using 72-dimensional $\mathbf{y}(n)$ with $d_k(n) = \mathrm{XOR}(u(n-1), u(n-k))$ for delay $k = 2, \cdots, 24$: α = (a) 0.001, (b) 0.01, (c) 0.1, and (d) 1, where black, red, blue, brown, and green dots represent the results for $X = 1, 2, 4, 6$, and 8 ns, respectively, and colored lines connecting nearest-neighbor dots are guides.
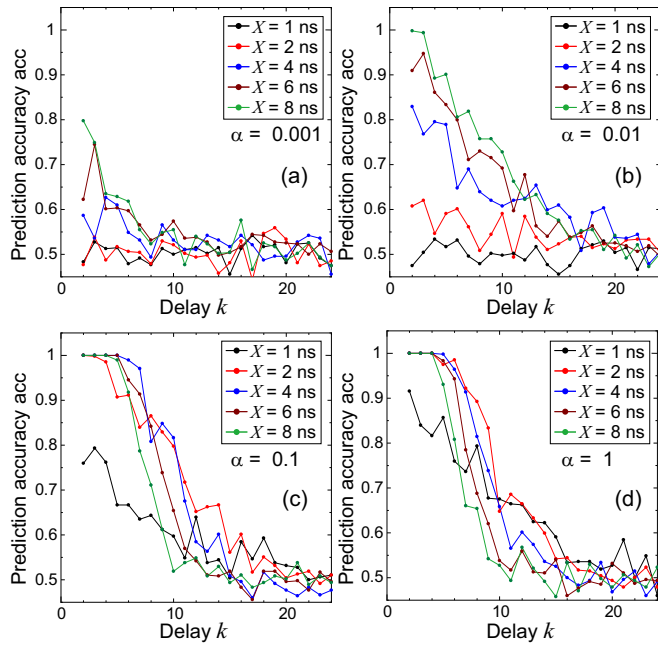
FIG. 13. Prediction accuracy plotted against delay k, which were estimated by solving XOR problems using 18-dimensional $\mathbf{y}(n)$ with $d_k(n) = \mathrm{XOR}(u(n-1), u(n-k))$ for delay $k = 2, \cdots, 24$: α = (a) 0.001, (b) 0.01, (c) 0.1, and (d) 1, where grey, orange, pale-blue, and pale-green dots represent the results with $X = 4$ ns for areas 1, 2, 3, and 4, respectively, and colored lines connecting nearest-neighbor dots are guides.

highly improved, as shown in Fig. 12(c). The curves for $X = 2, 4, 6$, and 8 ns have acc = 1 for $k \geqslant 2$, which are confirmed by the plateau in the low k range. The curves for $X = 4, 6$, and 8 ns have similar decreasing slopes. When α is increased from 0.1 to 1, the feature of the curve for $X = 2$ ns is slightly improved, whereas the features of the curves for $X = 4, 6$, and 8 ns are almost unchanged, as shown in Fig. 12(d). Table I lists CP for all the α and X conditions, in which CP $\geqslant 4$ was obtained for 7 conditions and the highest CP = 5 was obtained for $\alpha = 0.1$ with $X = 4$ and 6 ns.

So far, the reservoir computing was performed using 72-dimensional $\mathbf{y}(n)$ from 72 detectors. Here, the detectors are divided into four groups named areas 1, 2, 3, and 4, as shown
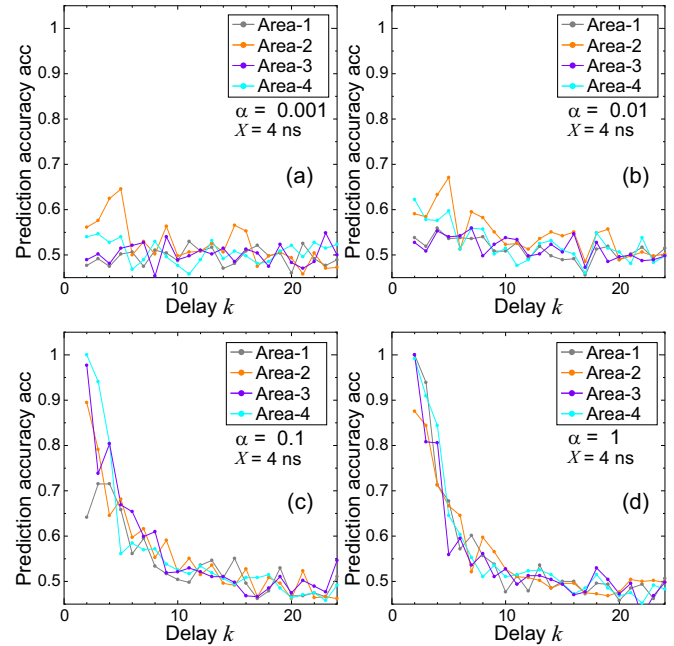
in Fig. 3(c), and then the temporal XOR problems were solved using 18 detectors in each area, when α was varied and X = 4 ns. In consequence, $\mathbf{y}(n)$ for each computation is an 18-dimensional vector. Figures 13(a)–13(d) show acc calculated for $\alpha = 0.001, 0.01, 0.1$, and 1, respectively, plotted against k, where grey, orange, pale-blue, and pale-green dots represent the results for areas 1, 2, 3, and 4, respectively. At each α value, the curves for areas 1, 2, 3, and 4 are similar to each other and have smaller values than the curve for $X = 4$ ns in Figs. 12(a)–12(d). The results clearly indicate that the $\mathbf{y}(n)$ dimension effective for reservoir computing increases with increasing number of the detector, i.e., higher dimensionality was efficiently obtained by the increase of the reservoir output

TABLE I. Summaries of the computing capabilities CP, MC$^{\mathrm{acc}}$, and MC obtained for various α and X values. CP for the temporal XOR ploblems is defined by the maximum value of delay k such that 100% prediction accuracy (acc = 1) is maintained when k is increased, where bars in CP columns represent N/A. MC$^{\mathrm{acc}}$ for the delay tasks is defined by the maximum value of delay k such that 100% prediction accuracy (acc = 1) is maintained when k is increased. MC for the delay tasks is defined by Eq. (13).

	$\alpha = 0.001$			$\alpha = 0.01$			$\alpha = 0.1$			$\alpha = 1$		
	XOR problem	Delay task		XOR problem	Delay task		XOR problem	Delay task		XOR problem	Delay task	
X	CP	MC$^{\mathrm{acc}}$	MC	CP	MC$^{\mathrm{acc}}$	MC	CP	MC$^{\mathrm{acc}}$	MC	CP	MC$^{\mathrm{acc}}$	MC
1 ns	—	0	14	—	0	36	—	15	31	—	18	28
2 ns	—	0	19	—	9	42	2	16	24	4	13	21
4 ns	—	1	20	—	15	32	5	12	16	4	11	15
6 ns	—	3	18	—	15	26	5	10	13	4	10	13
8 ns	—	5	18	—	13	23	4	9	10	4	8	11

signals with various wave forms. This finding also leads to a conclusion that the periodic arrangement of detectors in Fig. 3(c) works well for the realization of high dimensionality in $\mathbf{y}(n)$, which is advantageous for feasible implementation of the device.

Here, we discuss the effectiveness of the damping regions at the left and right boundaries. If spin waves are significantly reflected from these boundaries, they also contribute to the output signals at the detectors, particularly, in areas 1 and 4. This is because spin waves propagating from the input exciter can interfere with reflected spin waves that have memories in the further past. In this situation, acc values for areas 1 and 4 can differ from those for areas 2 and 3. On the other hand, the curves in Figs. 13(a)–13(d) do not have a clear area dependence at each α value. Thus, the results indicate that spin waves reflected from the left and right boundaries were satisfactorily suppressed.

Further analyses on the wave forms for $\alpha = 0.1$ and $X = 4$ ns were performed with various combinations of the detectors to mainly evaluate whether the wave forms near the left and right boundaries contribute to the computing capability (Appendix B). This is because these wave forms seemingly do not reflect the input information, as seen in that at detector 12 in Fig. 9. The results clearly support the fact that the wave form obtained at each detector has a similar contribution to the computing capability.

B. Delay tasks and memory capacity

Memory capacity was estimated by solving delay tasks with the target output:

$$d_k(n) = u(n - k) \quad \text{for delay } k = 1, \ldots, 72. \quad (12)$$

In this paper, we evaluated memory capacity by two standards. The first one is MC that is defined by the widely used formula as follows [40]:

$$\text{MC} = \sum_{k=1}^{72} \gamma^2(k)$$

$$= \sum_{k=1}^{72} \frac{\text{Cov}^2(u(n-k), b_k^*(n))}{\text{Var}(u(n))\text{Var}(b_k^*(n))}, \quad (13)$$

where γ, Cov, and Var denote the coefficient of determination, covariance, and variance, respectively, and $b_k^*(n)$ is the binary output at n in the delay task with k, respectively. The main purpose of using this index is to study a fundamental feature of the reservoir device by comparing estimated values with those in other works. On the other hand, the second one is MC$^{\text{acc}}$ that is defined by the maximum value of k such that 100% prediction accuracy (acc = 1) is maintained when k is increased. The purpose of using this index is to clarify the conditions for applications with high-accuracy and long-step memory, such as forecasting nonlinear autoregressive moving average time series [42]. It should be noted that MC$^{\text{acc}}$ has a value that is equal to or less than MC.

Figures 14(a)–14(d) show $\gamma^2(k)$ calculated for $\alpha = 0.001$, 0.01, 0.1, and 1, respectively, plotted against delay k, where black, red, blue, brown, and green dots represent the results for $X = 1, 2, 4, 6$, and 8 ns, respectively. When $\alpha = 0.001$
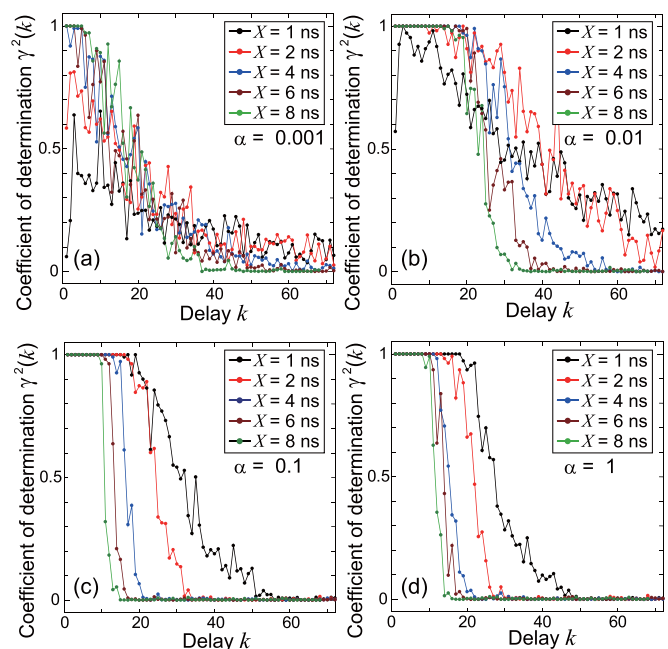

FIG. 14. Coefficient of determination $\gamma^2(k)$ plotted against delay k, which were estimated by solving delay tasks using 72-dimensional $\mathbf{y}(n)$ with the target output $d(n) = u(n - k)$ for delay $k = 1, \cdots, 72$: $\alpha = $ (a) 0.001, (b) 0.01, (c) 0.1, and (d) 1, where black, red, blue, brown, and green dots represent the results for $X = 1, 2, 4, 6$, and 8 ns, respectively, and colored lines connecting nearest-neighbor dots are guides.

in Fig. 14(a), the curves for all the X values monotonically decrease with increasing k and their shapes are similar to each other. The estimated values of MC$^{\text{acc}}$ for $X = 1, 2, 4, 6$, and 8 ns are 0, 0, 1, 3, and 5, respectively. When $\alpha = 0.01$ in Fig. 14(b), the curves show different features depending on X. As k is increased from 0, the curves except $X = 1$ ns show $\gamma^2(k) = 1$ at first and then they monotonically decrease. It was found that MC$^{\text{acc}}$ are not related to the fact that the decreasing slope becomes more steep with increasing X: MC$^{\text{acc}}$ for $X = 1, 2, 4, 6$, and 8 ns are 0, 9, 15, 15, and 13, respectively. When $\alpha = 0.1$ in Fig. 14(c), all the curves are clearly different. As k is increased from 0, all the curves show $\gamma^2(k) = 1$ at first, whereas the k position where the curve bends toward 0 becomes lower with increasing X. Owing to this feature, MC$^{\text{acc}}$ approximately decreases with increasing X: The estimated values of MC$^{\text{acc}}$ for $X = 1, 2, 4, 6$, and 8 ns are 15, 16, 12, 10, and 9, respectively. When $\alpha = 1$ in Fig. 14(d), the overall feature is similar to that in Fig. 14(c), whereas MC$^{\text{acc}}$ for $X = 1$ ns is slightly improved: The estimated values of MC$^{\text{acc}}$ for $X = 1, 2, 4, 6$, and 8 ns are 18, 13, 11, 10, and 8, respectively. On the other hand, further analyses on the wave forms for $\alpha = 0.1$ and $X = 4$ ns were performed with various combinations of the detectors, as the temporal XOR problems in Appendix B. It was also confirmed that the wave form obtained at each detector has a similar contribution to the computing capability.

The results for MC$^{\text{acc}}$ and MC are summarized in Table I. When $\alpha = 0.001$, MC is considerably larger than MC$^{\text{acc}}$ in all the X cases, which arise from the gentle decreases in $\gamma^2(k)$

with increasing k in Fig. 14(a). When $\alpha = 0.01$, MC is also considerably larger than MC^{acc} in the cases of $X = 1$ and 2 ns. This is because the decreasing slopes for $X = 1$ and 2 ns are more gradual than those for other X values, as seen in Fig. 14(b). Thus, under the α condition that the decreasing slope depends significantly on X, larger MC^{acc} does not always lead to larger MC, i.e., MC^{acc} is not correlated with MC. For $\alpha = 0.1$ and 1, where the decreasing slopes for all the X values are similar to each other, the estimated values of MC^{acc} are reasonably correlated with MC and, in a wide range of X, $\text{MC}^{\text{acc}} > 10$ and $\text{MC} > 13$ were achieved. These MC values are comparable to those obtained in the reservoir system with a Mackey-Glass nonlinearity and a feedback loop [43] and, to the best of our knowledge, the maximum MC value of 21 is the highest ever in the on-chip electronic systems [16,18–20].

C. Parameter conditions for high computing capabilities

The first purpose in this section is to find sets of the physical and input parameters, α and X, that achieve the high computing capabilities for both two tasks. From Table I, the following (α, X) conditions have relatively high computing capabilities: (0.1, 4 ns), (0.1, 6 ns), (1, 2 ns), and (1, 4 ns). The second purpose is to reveal the origins of why these conditions show the high computing capabilities, which will be discussed with the physical properties of the spin waves in the next section.

VII. DISCUSSION

Toward deep understanding of physical reservoir computing, it is worthwhile to discuss the relationship between the reservoir-computing capabilities CP and MC^{acc}, with

suggestions on the physical properties of the spin waves behind the reservoir computing. The preceding sections have mainly studied how the physical properties of spin waves and the reservoir-computing capabilities are changed by the two parameters: the device parameter α and the input signal parameter X.

In the evaluation of the capability for the temporal XOR problems, the α and X conditions are narrow for $\text{CP} \geqslant 2$: $X = 2, 4, 6,$ and 8 ns with $\alpha = 0.1$ and $X = 2, 4, 6,$ and 8 ns with $\alpha = 1$. Thus, α is the primal parameter for solving the temporal XOR problems. Since the decreases of α and X result in the enhancement of nonlinear phenomena characterized by the sub-f_{IN} and interhamonic components, excessive nonlinear phenomena degrade CP. It may be considered that richer nonlinear phenomena are preferable for linearly-inseparable problems as the temporal XOR problems, but the results indicate that there are effective and obstructive nonlinear phenomena for reservoir computing. In the evaluation of the memory capacities, the MC^{acc} values higher than 10 were obtained under the following conditions: $X = 4, 6,$ and 8 ns with $\alpha = 0.01$, $X = 1, 2, 4,$ and 6 ns with $\alpha = 0.1$, and $X = 1, 2, 4,$ and 6 ns with $\alpha = 1$. Thus, the X and α conditions are broad even though the DFT spectra are clearly changed with the X and α values in those ranges. Considering CP and MC^{acc}, the device and input conditions for high-accuracy computing is more restrictive for the temporal XOR problems.

On the other hand, as pointed out previously, whereas the post signal processing in the readout part extracts the characteristics only in the frequency range lower than $f_c = 1.3$ GHz, the computation capabilities strongly depend on α and X. This is clear evidence that the characteristics in the high and low frequency ranges dynamically interact with each other through nonlinear phenomena since the influence of the
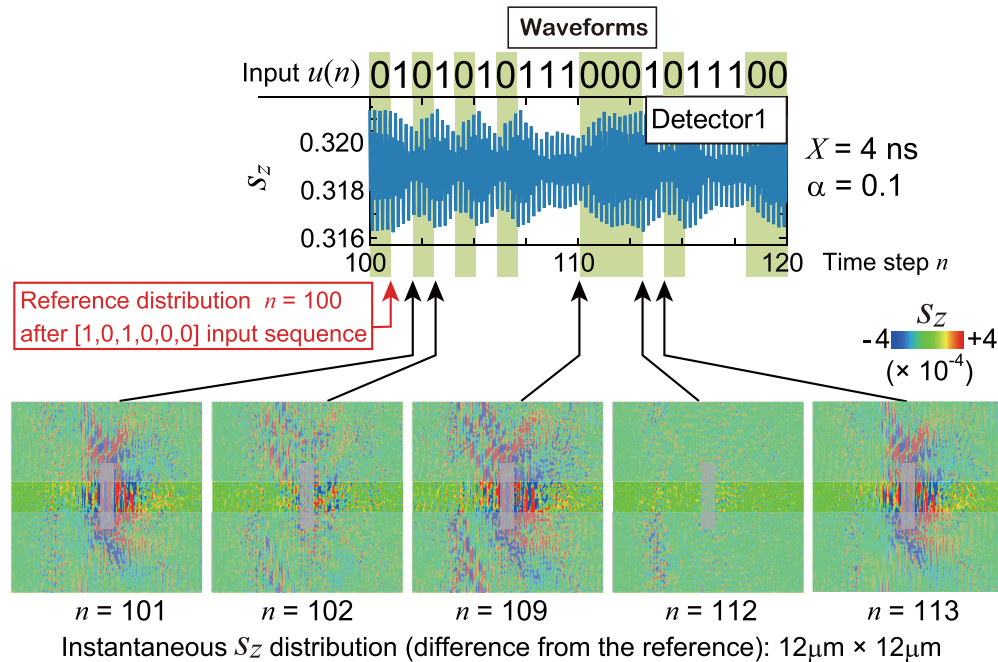

FIG. 15. Time-domain wave form at detector 1 (the upper side) and instantaneous top-view sz distributions in space (the bottom side), when $\alpha = 0.1$ and $X = 4$ ns. The s_z distributions with 12×12 μm^2 in area are the differences from the reference s_z distribution at the end of $n = 100$ (after [1, 0, 1, 0, 0, 0] input sequence), where bright and dense-gray areas are the detector and exciter areas, respectively, light-gray areas are the other areas, and the direction and strength of s_z is defined by a color bar.

change in α and X was mostly observed in the frequency range higher than f_c. Since nonlinear phenomena were characterized only by the DFT spectra that do not directly capture dynamic behavior, further analyses would be necessary. Nonetheless, the above finding is noteworthy for understanding of physical reservoir computing.

When the discussion is focused on the cases of $\alpha = 0.1$ and 1, MC^{acc} almost monotonically decreases and CP appears with increasing X at each α, as confirmed in Table I. This may indicate a tradeoff between MC^{acc} and CP, as pointed out for the echo state network [41]. In the present study, since the CP values are not so high, the change in CP with X is not significant. This probably arises from the fact that the $\mathbf{y}(n)$ dimension is 72. Considering the result that the CP value becomes higher with increasing number of the output electrodes in Figs. 12 and 13, it is expected that much higher CP values can be obtained with higher-dimensional $\mathbf{y}(n)$ that is generated by a reservoir device with larger numbers of output electrodes. This achievement may lead to a significant change in CP with X, which can provide further insight into the relation between MC^{acc} and CP.

VIII. SUMMARY

In the first half of the numerical experiments, it was shown how the physical properties of the spin waves were changed by the two parameters: the device parameter α and the input signal parameter X. From the characterization by the DFT spectra, it was found that nonlinear phenomena were obtained in all the α and X conditions examined here, which is mainly attributable to the stripe domain structure and unstable spin waves under the low external magnetic field. As α and X are decreased, nonlinear phenomena more frequently occur, which was mostly characterized by the increase of the sub-f_{IN} and interhamonic components in the DFT spectra. In the latter half of the numerical experiments, the reservoir computing was performed to evaluate CP by solving the temporal XOR problems as well as MC^{acc} and MC by solving the delay tasks. In the temporal XOR problems, $\text{CP} \geqslant 4$ was obtained under the following conditions: $X = 4$, 6, and 8 ns with $\alpha = 0.1$ and $X = 2$, 4, 6, and 8 ns with $\alpha = 1$. Since the sub-f_{IN} and interhamonic components for $\alpha = 0.1$ and 1 are smaller than those for $\alpha = 0.001$ and 0.01, moderate suppression of nonlinear physical phenomena was found to be a key to such high CP values. Among the conditions for $\text{CP} \geqslant 4$, MC^{acc} and MC are relatively high and decrease with increasing X. The maximum value of MC^{acc} is 13 at $\alpha = 1$ and $X = 2$ ns condition and the value of MC is 21 at the same condition. Therefore, this study demonstrated that the spin-wave-based reservoir computing system simultaneously achieves the relatively high CP, MC^{acc}, and MC values among electronic reservoir-computing hardware systems.

Since the spin-wave-based reservoir device utilizes the signal transmission through a continuous magnetic garnet film and the signal excitation/detection by input/output electrodes on the top, it is expected that the computing capabilities of the reservoir system can be further improved and various functionalities for edge-computing applications can be realized just by changing the number, selection, and the arrangement of input/output electrodes, an external magnetic field, and the

magnetic domain structure. Such scalability also can offer a rich opportunity for further understanding of the underlying nonlinear physics.

APPENDIX A: SPATIAL DISTRIBUTIONS OF s_z AT DIFFERENT n

Figure 15 shows the time-domain wave form at detector 1 and instantaneous top-view s_z distributions in space, when $\alpha = 0.1$ and $X = 4$ ns. The s_z distributions with 12×12 μm^2 in area are the differences from the reference s_z distribution at the end of $n = 100$ (after [1, 0, 1, 0, 0, 0] input sequence), where bright and dense-gray areas are the detector and exciter areas, respectively, and light-gray areas are the other areas, as
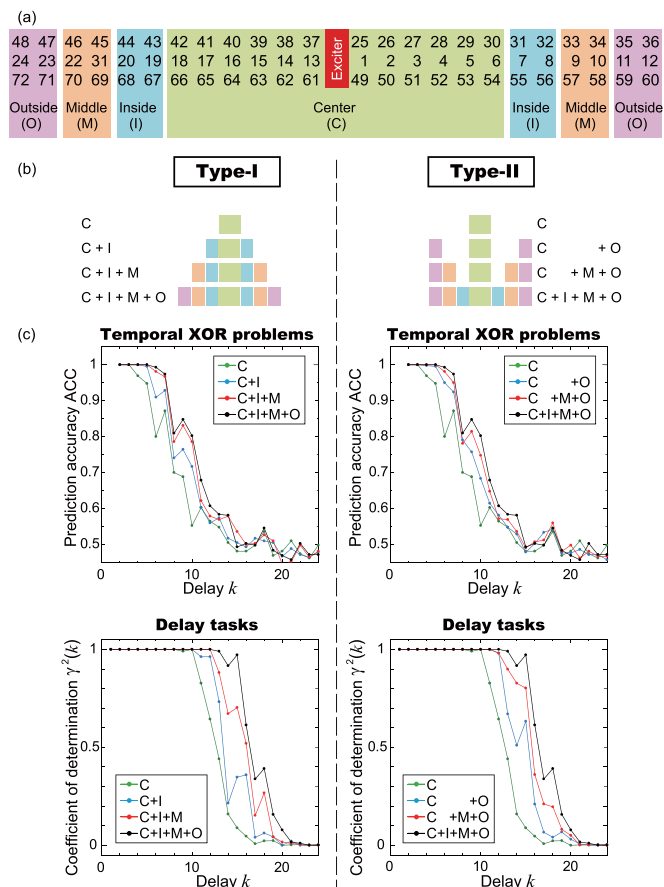

FIG. 16. (a) Detector numbers and group names, where center, inside, middle, and outside groups are referred to as C, I, M, and O, respectively, in (b) and (c). (b) Two types of methods for changing detector numbers): Type I increases the groups toward the outside boundaries (from I to O), whereas type II increases the groups toward the inside exciter (from O to I). (c) Computing capabilities for temporal XOR problems and delay tasks plotted against delay step k, where each color of line/dots shown in the inset corresponds to the combination of the groups in (b).

TABLE II. Summary of computing capabilities CP, MC^{acc}, and MC obtained with various group sets, where types I and II are the methods for changing detector groups and C, I, M, and O denote center, inside, Mmiddle, and outside groups in Fig. 16(a), respectively.

	Type I				Type II		
Combination	CP	MC^{acc}	MC	CP	MC^{acc}	MC	Combination
C	3	8	12	3	8	12	C
C + I	4	10	14	4	11	14	C + O
C + I + M	5	12	15	5	11	15	C + I + O
C + I + M + O	5	12	16	5	12	16	C + I + M + O

confirmed from Fig. 3. Unlike spherical spin waves propagating through a uniform magnetic domain in our previous study [22], the s_z distributions are complex due to the stripe domain structure. This fact is preferable for reservoir computing since it results in the various time-domain waveforms at the detectors, which leads to a high dimensionality in the reservoir output signals. From the s_z distributions in the detector areas, the following characteristics were found, which is consistent with $MC^{acc} = 13$.

(1) The s_z distributions at $n = 100$ and 112 were obtained after [1, 0, 1, 0, 0, 0] and [1, 1, 1, 0, 0, 0] input sequences, respectively. Since some red and blue distributions are seen in the s_z distribution at $n = 112$, these two distributions are different from each other. This result reflects the difference in the input sequence (0 or 1 at the past 5 time steps).

(2) The s_z distributions at $n = 101$ and 113 were obtained after [1, 0, 1, 0, 0, 0, 1] and [1, 1, 1, 0, 0, 0, 1] input sequences, respectively. From detailed observations, these s_z distributions are different from each other. This result reflects the difference in the input sequence (0 or 1 at the past six time steps).

APPENDIX B: CONTRIBUTIONS OF THE WAVE FORMS NEAR THE LEFT AND RIGHT BOUNDARIES TO THE COMPUTING CAPABILITY

There is a trend that the influence of the input sequence on the wave form becomes more ambiguous with increasing distance from the exciter, leading to the fact that the wave form at detector 12 in Fig. 9 seemingly does not contain the input information. To clearly analyze whether the wave forms near the outside boundaries contribute to the computing capabilities, numerical experiments were additionally performed.

The contribution of one detector to computing capabilities is not so large and it probably does not exhibit a linear increase with increasing number of the detectors. Based on these considerations, some numbers of the detectors were grouped, as shown in Fig. 16(a), where center, inside, middle, and outside groups will be referred to as C, I, M, and O, respectively. Then, while C group is commonly used, two types of methods are explored to change the detector numbers, as shown in Fig. 16(b): Type I increases the groups toward the outside boundaries (from I to O), whereas type II increases the groups toward the inside exciter (from O to I). If computation capabilities are improved with a similar trend in types I and II, the contribution of each group is comparable, i.e., computation capabilities are determined mostly by the numbers of the detectors, irrespective of the spatial position. In the following numerical experiment, we use the wave forms obtained for $\alpha = 0.1$ and $X = 4$ ns.

Figure 16(c) shows computing capabilities for the temporal XOR problems and delay tasks, where results obtained in types I and II are shown on the left- and right-hand sides, respectively, and each color of line and/or dots shown in the inset corresponds to the combination of the groups in (b). In all the graphs, the downward slope shifts toward the right and the flat range at 1 basically increases with increasing detector numbers, which indicates the improvement of the computing capabilities. When the computing capabilities for the same task in types I and II are compared with each other, they similarly increase with increasing number of the detectors, as summarized in Table II. Thus, the detectors in O group (including detector 12) actually improves the computing capabilities and its contribution is comparable to those in I and M groups.

[1] H. Jaeger, The "echo state" approach to analysing and training recurrent neural networks-with an erratum note, German National Research Center for Information Technology GMD Technical Report No. 148, 2001, p. 13.

[2] W. Maass, T. Natschläger, and H. Markram, Real-time computing without stable states: A new framework for neural computation based on perturbations, Neural Comput. **14**, 2531 (2002).

[3] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, Edge computing: Vision and challenges, IEEE Internet of Things J. **3**, 637 (2016).

[4] N. Abbas, Y. Zhang, A. Taherkordi, and T. Skeie, Mobile edge computing: A survey, IEEE Internet of Things J. **5**, 450 (2017).

[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 Networks **115**, 100 (2019).

[6] L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutiérrez, L. Pesquera, C. R. Mirasso, and I. Fischer, Photonic information processing beyond turing: An optoelectronic implementation of reservoir computing, Opt. Express **20**, 3241 (2012).

[7] Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, Optoelectronic reservoir computing, Sci. Rep. **2**, 287 (2012).

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

[9] D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, Parallel photonic information processing at gigabyte per second data rates using transient states, Nat. Commun. **4**, 1364 (2013).

[10] 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).

[11] C. Du, F. Cai, M. A. Zidan, W. Ma, S. H. Lee, and W. D. Lu, Reservoir computing using dynamic memristors for temporal information processing, Nat. Commun. **8**, 2204 (2017).

[12] J. Moon, W. Ma, J. H. Shin, F. Cai, C. Du, S. H. Lee, and W. D. Lu, Temporal data classification and forecasting using a memristor-based reservoir computing system, Nat. Electron. **2**, 480 (2019).

[13] K. Nakajima, H. Hauser, T. Li, and R. Pfeifer, Information processing via physical soft body, Sci. Rep. **5**, 10487 (2015).

[14] C. Fernando and S. Sojakka, Pattern recognition in a bucket, in Advances in Artificial Life, *European Conference on Artificial Life*, edited by W. Banzhaf, J. Ziegler, T. Christaller, P. Dittrich, and J. T. Kim (Springer, Berlin, Heidelberg, 2003), Vol. 2801, pp. 588–597.

[15] G. Dion, S. Mejaouri, and J. Sylvestre, Reservoir computing with a single delay-coupled non-linear mechanical oscillator, J. Appl. Phys. **124**, 152132 (2018).

[16] E. Nako, K. Toprasertpong, R. Nakane, Z. Wang, Y. Miyatake, M. Takenaka, and S. Takagi, Proposal and experimental demonstration of reservoir computing using $Hf_{0.5} Zr_{0.5} O_2$/Si FeFETs for neuromorphic applications, in *2020 IEEE Symposium on VLSI Technology* (IEEE, 2020), pp. 1–2.

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

[18] 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).

[19] 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. Research **2**, 023389 (2020).

[20] 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).

[21] H. Nomura, T. Furuta, K. Tsujimoto, Y. Kuwabiraki, F. Peper, E. Tamura, S. Miwa, M. Goto, R. Nakatani, and Y. Suzuki, Reservoir computing with dipole-coupled nanomagnets, Jpn. J. Appl. Phys. **58**, 070901 (2019).

[22] R. Nakane, G. Tanaka, and A. Hirose, Reservoir computing with spin waves excited in a garnet film, IEEE access **6**, 4462 (2018).

[23] R. Nakane, G. Tanaka, and A. Hirose, In a spin-wave reservoir for machine learning, in *2019 International Joint Conference on Neural Networks (IJCNN)* (IEEE, 2019), pp. 1–9.

[24] A. Mahmoud, F. Ciubotaru, F. Vanderveken, A. V. Chumak, S. Hamdioui, C. Adelmann, and S. Cotofana, Introduction to spin wave computing, J. Appl. Phys. **128**, 161101 (2020).

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

[26] B. Rana and Y. Otani, Towards magnonic devices based on voltage-controlled magnetic anisotropy, Commun. Phys. **2**, 90 (2019).

[27] A. Hubert and R. Schäfer, *Magnetic Domains: The Analysis of Magnetic Microstructures* (Springer-Verlag, Berlin, Heidelberg, 1998), Chap. 5.6.1.

[28] D. D. Stancil and A. Prabhakar, *Spin Waves* (Springer, US, 2009), Chap. 9.

[29] D. Craik and P. Cooper, Bias magnet design for bubble memory devices, IEEE Trans. Magn. **14**, 306 (1978).

[30] 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).

[31] E. E. Anderson, Molecular field model and the magnetization of YIG, Phys. Rev. **134**, A1581 (1964).

[32] S. Klingler, A. V. Chumak, T. Mewes, B. Khodadadi, C. Mewes, C. Dubs, O. Surzhenko, B. Hillebrands, and A. Conca, Measurements of the exchange stiffness of YIG films using broadband ferromagnetic resonance techniques, J. Phys. D **48**, 015001 (2014).

[33] S. A. Manuilov, S. Khartsev, and A. M. Grishin, Pulsed laser deposited $Y_3Fe_5O_{12}$ films: Nature of magnetic anisotropy I, J. Appl. Phys. **106**, 123917 (2009).

[34] C. Dubs, O. Surzhenko, R. Thomas, J. Osten, T. Schneider, K. Lenz, J. Grenzer, R. Hübner, and E. Wendler, Low damping and microstructural perfection of sub-40nm-thin yttrium iron garnet films grown by liquid phase epitaxy, Phys. Rev. Mater. **4**, 024416 (2020).

[35] P. Pirro, T. Brächer, A. Chumak, B. Lägel, C. Dubs, O. Surzhenko, P. Görnert, B. Leven, and B. Hillebrands, Spin-wave excitation and propagation in microstructured waveguides of yttrium iron garnet/Pt bilayers, Appl. Phys. Lett. **104**, 012402 (2014).

[36] J. Ding, T. Liu, H. Chang, and M. Wu, Sputtering growth of low-damping yttrium-iron-garnet thin films, IEEE Magn. Lett. **11**, 1 (2020).

[37] G. Schmidt, C. Hauser, P. Trempler, M. Paleschke, and E. T. Papaioannou, Ultra thin films of yttrium iron garnet with very low damping: A review, Phys. Status Solidi (b) **257**, 1900644 (2020).

[38] A. Testa, M. Akram, R. Burch, G. Carpinelli, G. Chang, V. Dinavahi, C. Hatziadoniu, W. Grady, E. Gunther, M. Halpin *et al.*, Interharmonics: Theory and modeling, IEEE Trans. Power Delivery **22**, 2335 (2007).

[39] H. Jaeger, Long short-term memory in echo state networks: Details of a simulation study, Technical Report No. 27, Jacobs University Bremen, 2012.

[40] H. Jaeger, Short term memory in echo state networks, GMD-German National Research Institute for Computer Science GMD Report No. 152, 2002.

[41] M. Inubushi and K. Yoshimura, Reservoir computing beyond memory-nonlinearity trade-off, Sci. Rep. **7**, 10199 (2017).

[42] A. F. Atiya and A. G. Parlos, New results on recurrent network training: Unifying the algorithms and accelerating convergence, IEEE transactions on neural networks **11**, 697 (2000).

[43] L. Appeltant, Reservoir computing based on delay-dynamical systems, Ph.D. thesis, Vrije Universiteit Brussel/Universitat de les Illes Balears, 2012, Chap. 5.2.