scientific reports

OPEN

Associative memory by virtual oscillator network based on single spin-torque oscillator

Yusuke Imai & Tomohiro Taniguchi ✉

A coupled oscillator network may be able to perform an energy-efficient associative memory operation. However, its realization has been difficult because inhomogeneities unavoidably arise among the oscillators during fabrication and lead to an unreliable operation. This issue could be resolved if the oscillator network were able to be formed from a single oscillator. Here, we performed numerical simulations and theoretical analyses on an associative memory operation that uses a virtual oscillator network based on a spin-torque oscillator. The virtual network combines the concept of coupled oscillators with that of feedforward neural networks. Numerical experiments demonstrate successful associations of 60-pixel patterns with various memorized patterns. Moreover, the origin of the associative memory is shown to be forced synchronization driven by feedforward input, where phase differences among oscillators are fixed and correspond to the colors of the pixels in the pattern.

The human brain has a sophisticated function called associative memory[1], whereby it can remember a pattern when shown a portion of that pattern. This function has been modeled in various ways with the goal of achieving a better understanding of brain activity and realizing energy-efficient bio-inspired computing. Since the development of an autocorrelation model in the 1970s[2–4], several theoretical models, such as the Hopfield model[5], have been developed that draw their inspiration from the characteristics of neural activity[6–15]. These models have also been implemented in experimental devices. For example, the associative memory operation was recently performed in a spintronic memory consisting of a nanometer-scale ferromagnetic multilayer[16]. In addition to these efforts embodying neuronal dynamics, it has been proposed that synchronized phenomena in coupled oscillator networks can be used to perform the associative memory operation[17–21]. For example, a detailed analysis was conducted on an *LC*-circuit oscillator network performing the operation[21]. A network of spintronic oscillators, called spin-torque oscillators (STOs), has also been shown to perform an associative memory operation[22].

There are two major issues with using an oscillator network for the associative memory operation. One is unstable operation due to inhomogeneity in the oscillator's parameters. For example, variations in frequency among the oscillators are unavoidable in experimental realizations; they prevent a synchronization between the oscillators and decrease the accuracy of the associative memory[21]. The other issue is that the required number of oscillators grows with the amount of input data. There are numerous challenges in fabricating a large number of oscillators and getting them to interact with each other. These issues might be resolved if we can construct an oscillator network virtually by using a single physical oscillator[23]. Such a network would have no inhomogeneities in its parameters as only one oscillator would have to be fabricated. However, there are questions on how such a network could be realized and how it could show synchronization phenomena.

In this work, we demonstrate an associative memory operation by a virtual oscillator network through numerical simulations and theoretical analyses. First, we provide a detailed description of the virtual oscillator network consisting of a single physical oscillator. In particular, we discuss the principles involved, i.e., those of the coupled oscillator networks and feedforward neural networks. Next, we show that a virtual oscillator network consisting of a single STO can recognize several different 60-pixel patterns by numerically simulating the motion of the STO. We reveal that the feedforward input in the virtual network forces the virtual oscillators to synchronize and that this phenomenon results in the associative memory operation.

National Institute of Advanced Industrial Science and Technology (AIST), Research Center for Emerging Computing Technologies, Tsukuba, Ibaraki 305-8568, Japan. ✉email: tomohiro-taniguchi@aist.go.jp

Results

Associative memory operation of this study

The associative memory operation studied here is to associate a pattern, called the pattern to be recognized, with a pattern in a stored set of patterns, called memorized patterns. For example, suppose that the three patterns, "0", "1", and "2", shown in Fig. 1a are memorized, and the one shown in Fig. 1b is the pattern to be recognized: we can see that the pattern to be recognized is similar to the memorized pattern "1". Throughout this paper, we will suppose the memorized patterns use 10(rows)× 6(columns)= 60-pixels patterns for memorized patterns and patterns to be recognized.

In the following subsections, we describe the concept of our virtual oscillator network after briefly reviewing a conventional oscillator network for comparison. Then, we demonstrate through numerical simulations that the virtual oscillator network can perform the associative memory operation.

Associative memory operation by conventional oscillator network

The associative memory operation by a conventional coupled oscillator network consists of two steps[21]. The first step is to give a correspondence between the phases of the oscillators and the colors of the pattern to be recognized. We prepare N oscillators corresponding to the pixels of the pattern to be recognized, where N is the number of oscillators (pixels). We introduce phases ψ_i ($i = 1, \cdots, N$) and phase differences $\Delta \psi_i = \psi_i - \psi_1$. The color of the ith pixel is determined by $\cos \Delta \psi_i$, which is white (black) when $\Delta \psi_i = 0 \, (\pi)$. According to this definition, the color of the first pixel is always white (see also the "Methods" for the definitions of color). Initially, there are no interactions between the oscillators. Thus, their phases are arbitrary, and the colors in the pattern are random, as schematically shown on the left of Fig. 2a. When interactions between the oscillators are introduced and the interaction strengths are appropriately determined by the Hebbian rule, all the phase differences become 0 or π in correspondence with the white and black pixels of the pattern to be recognized, as shown in the middle of Fig. 2a (see also "Methods" for model of the conventional oscillator network). Here, the Hebbian rule means that the interaction strength between the jth and ith oscillator is proportional to the weight,

$$w_{ij}^{(1)} = \xi_i^{\mathrm{R}} \xi_j^{\mathrm{R}}, \tag{1}$$

where $\xi_i^{\mathrm{R}} = +(-)1$ when the color of the pattern to be recognized at the ith pixel is white (black). Thus, $w_{ij}^{(1)} = +(-)1$ when the colors of the ith and jth are the same (opposite).

The second step is to replace the weights by the following ones, which can be regarded as an average of the weights among the memorized patterns,

$$w_{ij}^{(2)} = \frac{1}{N_{\mathrm{m}}} \sum_{m=1}^{N_{\mathrm{m}}} \xi_i^m \xi_j^m, \tag{2}$$

where N_{m} is the number of memorized patterns. The symbol $m = 1, 2, \cdots, N_{\mathrm{m}}$ is used to distinguish the memorized patterns. For example, the memorized patterns "0", "1", and "2" in Fig. 2a are labelled $m = 1, 2,$ and 3. The parameter ξ_i^m is $+(-)1$ when the color of the ith pixel in the mth memorized pattern is white (black). Then, the oscillator phases change to those of the memorized pattern most resembling the pattern to be recognized, and the association is achieved, as shown in the right in Fig. 2a.

Description of associative memory operation by virtual oscillator network

The associative memory operation by a virtual oscillator network consists of three steps.

First, we measure an oscillation of a single oscillator and divide it into N parts, as schematically shown on the first line of Fig. 2b. The ith part of the measured data is regarded as the output from the ith oscillator in a virtual network. In this step, the phase of each part is arbitrary, and therefore, the pattern arising from it is random. The measured data should be stored in a computer in order for it to be used in the next step.

Second, we excite another oscillation and divide the measured data into N parts again. At the initial time of each part, the phase, as well as the pattern determined from it, is arbitrary, as shown in the middle of Fig. 2b. This time, however, we apply an external force to the oscillator that is proportional to a linear combination of
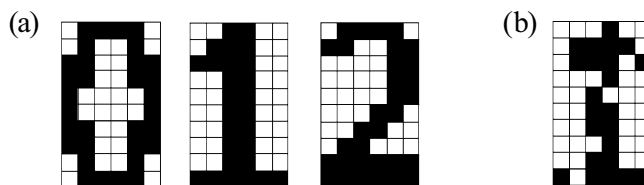
(a)

(b)

Figure 1. Examples of memorized patterns and a pattern to be recognized. (**a**) Three ($N_{\mathrm{m}} = 3$) memorized patterns, "0", "1", and "2". (**b**) The pattern to be recognized resembles memorized pattern "1". The oscillator network tries to associate the pattern to be recognized with the pattern "1". In an associative memory operation performed by a system consisting of N oscillators, the color of the ith ($i = 1, 2, \cdots, N$) pixel is determined by the phase ψ_i of the corresponding oscillator. The color is white (black) when the phase difference, $\Delta \psi_i = \psi_i - \psi_1$, is $0 \, (\pi)$. The color is on a gray scale when the phase difference is $0 < \Delta \psi_i < \pi$.
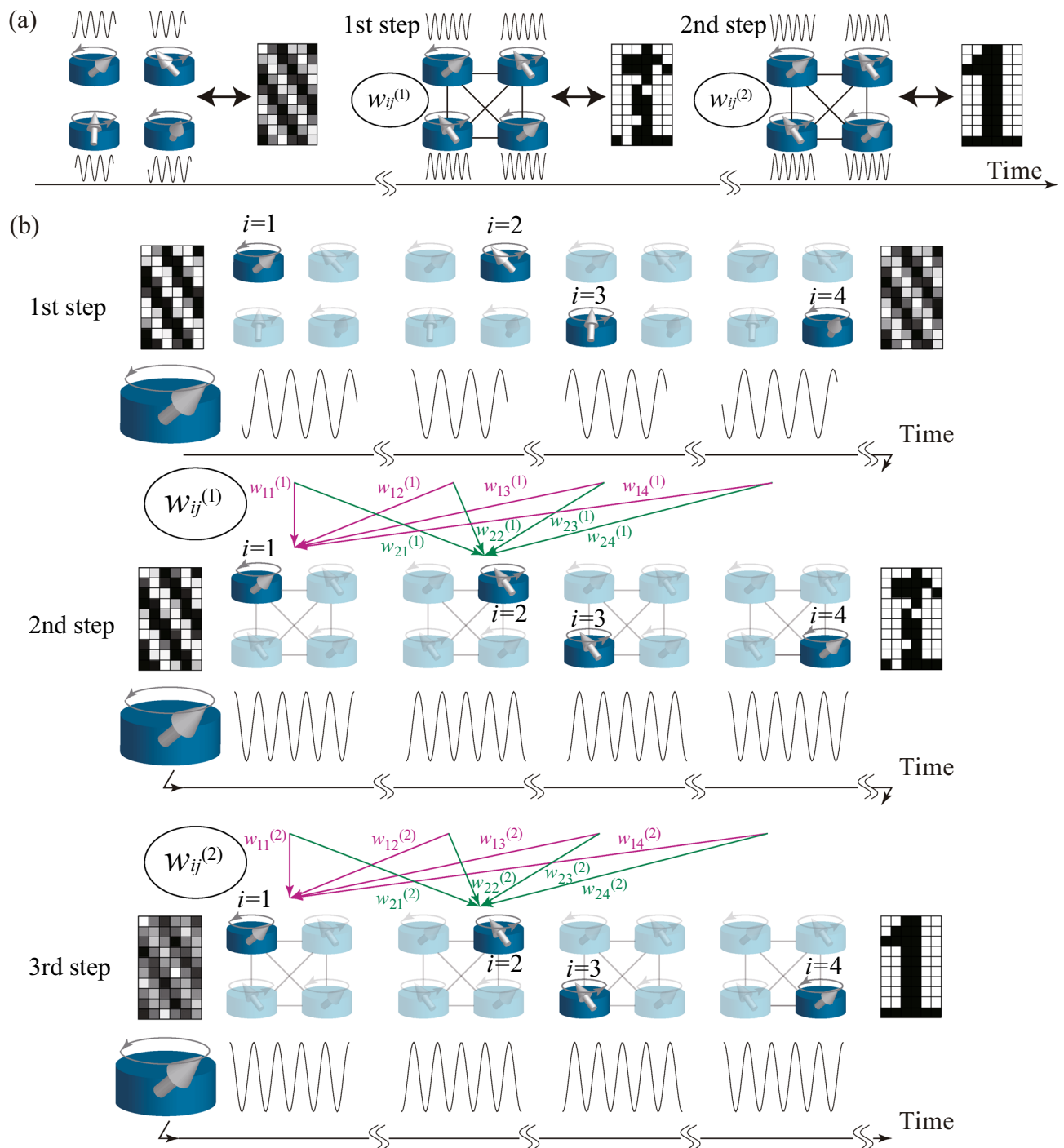
Figure 2. Schematic illustration of conventional and virtual oscillator networks. (**a**) In the conventional oscillator network, the oscillators are initially uncoupled (left). Therefore, the phase of each oscillator is arbitrary. When the oscillators interact with appropriate weights $[w_{ij}^{(1)}]$, the phases saturate to values corresponding to the pattern to be recognized (middle). When the weight changes $[w_{ij}^{(2)}]$, the phases change so that the corresponding pattern resembles one of memorized patterns (right). (**b**) In a virtual oscillator network, we drive an oscillation of a single oscillator and divide its output into N parts. The ith part is regarded as an output from the ith virtual oscillator. First [top of (**b**)], we measure the N outputs. The corresponding pattern in this step is arbitrary because there is no correlation among the oscillators. Second [middle of (**b**)], an external force is added to the oscillator. This force is a linear combination of the outputs in the first step with appropriated weights $[w_{ij}^{(1)}]$. The phase of each part eventually saturates to a value corresponding to the pixel color in the pattern to be recognized. Third [bottom of (**b**)], the second step is repeated while the force is a linear combination of the outputs in the second step with weights $w_{ij}^{(2)}$. Eventually, the phases saturate to the values corresponding to the memorized pattern most resembling the pattern to be recognized.

the measured data in the first step with weights (1). For example, in this study, the external force comes from a torque excited by an external magnetic field, which applied during the ith part of the oscillation is given by

$$H_i^{(1)} = \mathscr{H} \sum_{j=1}^{N} w_{ij}^{(1)} y_j^{(1)}, \tag{3}$$

where $\mathscr{H}$ denotes the amplitude and $y_j^{(1)}$ is the output from the jth oscillator measured in the first step [see also "Methods" for the detailed definition of $y_j^{(1)}$ in the numerical simulations]. Therefore, Eq. (3) is an oscillating function with the frequency of the oscillator. Because of the application of the magnetic field, the phase in each part eventually saturates to a certain value, and the pattern to be recognized is output, as shown in the middle of Fig. 2b. Note that the output signal of this process should be stored in a computer.

Third, we perform a measurement similar to the one in the second step but the magnetic field applied during the ith part is replaced by

$$H_i^{(2)} = \mathscr{H}' \sum_{j=1}^{N} w_{ij}^{(2)} y_j^{(2)}, \tag{4}$$

where $\mathscr{H}'$ denotes the amplitude, while $y_j^{(2)}$ is the output from the jth oscillator measured at the second step (see also "Methods" pertaining to the numerical simulations). The weights $w_{ij}^{(2)}$ are given by Eq. (2). The phase at the end of each part saturates to a value corresponding to the memorized pattern most resembling the pattern to be recognized, as shown in the bottom of Fig. 2b; i.e., the associative memory operation is completed.

There are several differences between the conventional and virtual oscillator networks (see also "Methods" for the models). For example, the oscillators in the conventional oscillator network interact instantaneously, and their phase differences saturate to values corresponding to pixel colors as a result of mutual synchronization. On the other hand, the oscillators in the virtual oscillator network do not interact each other instantaneously. As can be seen in Eqs. (3) and (4), the oscillator outputs from the previous steps are used in the magnetic field in the current step. From perspective, the virtual oscillator network is similar to a feedforward neural network because the information on the oscillator phases in one step is sent to the oscillation in the next step. At the same time, we should note that the weights in the virtual oscillator network are fixed, as in the case of the conventional oscillator network. This is in contrast with a feedforward neural network used in deep learning, in which weights are updated by backpropagation. Thus, the virtual oscillator network can be regarded as a hybrid combination of a coupled oscillator network and a feedforward neural network. In the discussion below, we will reveal that the feedforward inputs cause forced synchronization among the divided parts and result in the associative memory operation. Before that, however, we must demonstrate that this virtual oscillator network can actually perform the associative memory operation.

Equation of motion of oscillator

As an oscillator in the virtual oscillator network, we use a vortex STO, which has various advantages for practical applications and has been frequently used in spintronics experiments on bio-inspired computing[24–27]. An STO consists of a ferromagnetic/nonmagnetic multilayer on the nanometer scale, as schematically shown in Fig. 3a. A vortex of magnetic moments appears when a diameter and thickness of a cylinder-shape ferromagnet are on the order of 100 and 1 nm, respectively. When an electric current and/or magnetic field are applied to the STO, magnetic moments show precessions around their equilibrium direction. According to a recent experiment on chaos excitation in an STO[28], we assume that a force added to the virtual oscillator network corresponds to a torque excited by magnetic field, as mentioned above. It has been shown both experimentally and theoretically that the dynamics in a vortex STO are well described by the Thiele equation[29–36], which is the equation of motion for a center of the vortex structure, called the vortex core (see also "Methods" for Thiele equation):

$$-G\mathbf{e}_z \times \dot{\mathbf{X}} - |D|\left(1 + \xi s^2\right)\dot{\mathbf{X}} - \kappa\left(1 + \zeta s^2\right)\mathbf{X} + a_J J p_z \mathbf{e}_z \times \mathbf{X} + c a_J J R_0 p_x \mathbf{e}_x + c\mu^* \mathbf{e}_z \times \mathbf{H} = \mathbf{0}, \tag{5}$$

where $\mathbf{X} = (X, Y, 0)$ represents the position of the vortex core in the xy plane. While the physical meanings and the values of many parameters are explained in "Methods", two quantities should be explained here. The first is the current density J, which causes a limit-cycle oscillation of the vortex core. The other is the external magnetic field $\mathbf{H}$, which is used to excite a torque. It is useful to notice that Eq. (5) can be approximated as (see also "Methods" for the analytical solution of the Thiele equation)

$$\dot{s} = as - bs^3 - \frac{c\mu^*}{GR}H_y \cos\psi, \tag{6}$$

$$\dot{\psi} = \frac{\kappa}{G}\left(1 + \zeta s^2\right) + \frac{c\mu^*}{GRs}H_y \sin\psi, \tag{7}$$

where $s = |\mathbf{X}|/R$ ($0 \leq s \leq 1$) is the distance of the vortex core from the center of the ferromagnet normalized by the disk radius R, while $\psi = \tan^{-1}(Y/X)$ is the phase. Here, $a = (|D|\kappa/G^2)[(J/J_c) - 1]$ and $b = (|D|\kappa/G^2)(\xi + \zeta)$, where $J_c = |D|\kappa/(Ga_J p_z)$. The magnetic field $\mathbf{H}$ is assumed to have only a y component H_y. Note that Eqs. (6) and (7) are similar to the equation of motion of the Stuart-Landau oscillator[37]. Therefore, the vortex core shows a limit-cycle oscillation around the disk center in the xy plane with an oscillating
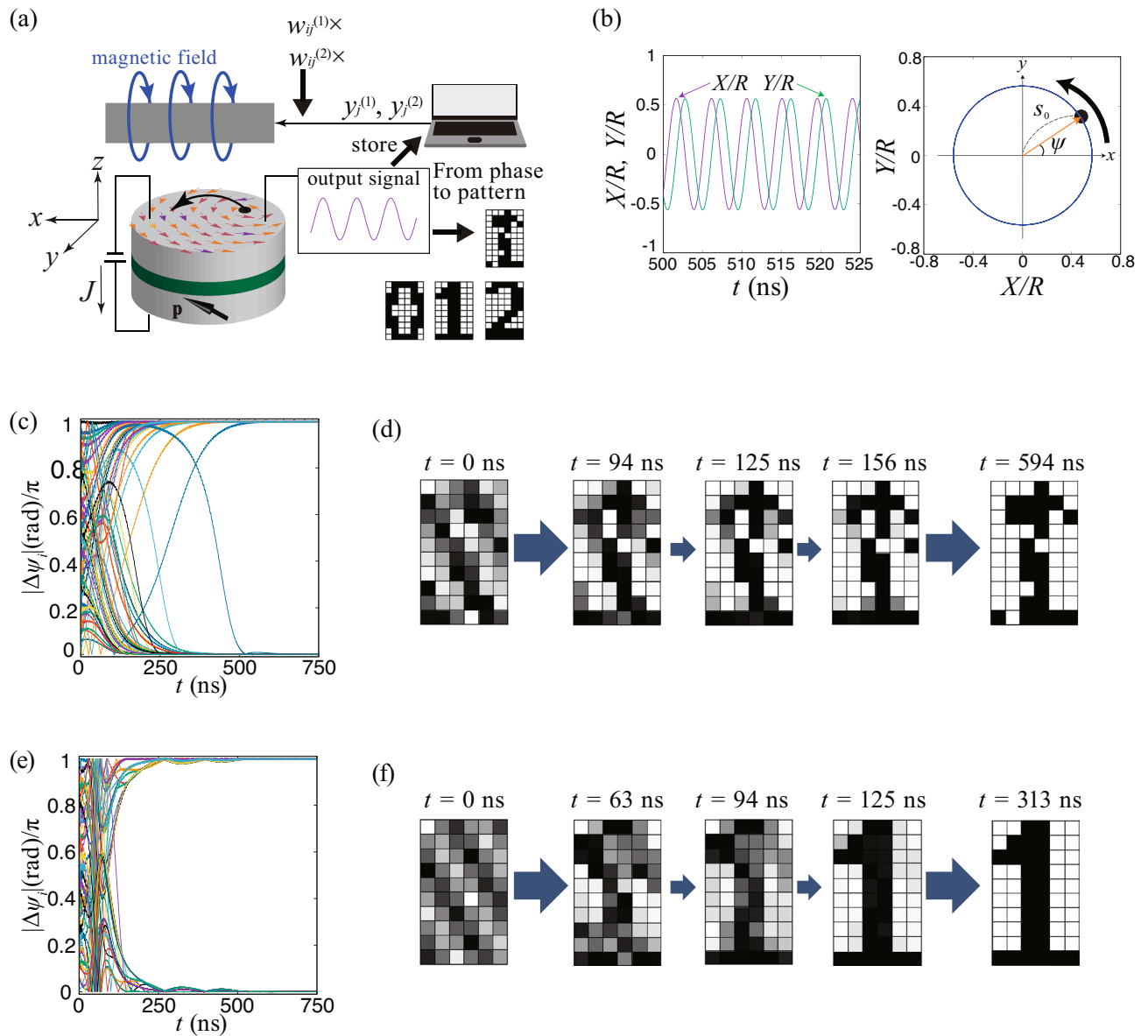
Figure 3. Description of STO and demonstration of associative memory by a virtual oscillator network. (**a**) Schematic illustration of vortex spin torque oscillator and (**b**) vortex-core dynamics driven by electric current. The STO has a cylindrical shape, and the z axis is orthogonal to the circular plane. Magnetic moments, shown as colored arrows in top ferromagnet, form a circular structure. The black dot around which the moments turn is the vortex core. Electric current is injected into the STO; positive current flows from bottom to top in the figure. When the electric current density J exceeds a threshold value, the vortex core oscillates around the disk center. The output signals from the STO during the first (second) step in Fig. 2b are stored, and their linear combination with weights $w_{ij}^{(1)}$ [$w_{ij}^{(2)}$] defined from the pattern to be recognized (memorized patterns) is used as magnetic field during the second (third) step. For simplicity, the dynamics in the absence of the magnetic field is shown. The components of the vortex-core's position, X/R, and Y/R, oscillate around the disk center, and a trajectory is approximately a circle. The distance of the vortex-core's position from the disk center, s, is approximately constant value, s_0. The phase measured from the x axis is denoted as ψ. (**c**) Time evolutions of the 59 phase differences, $\Delta\psi_i$ ($i = 2, 3, \cdots, 60$) and (**d**) snapshots of generating a pattern to be recognized on 60-pixels. (**e**) Time evolutions of the phase difference and (**f**) snapshots of the corresponding pattern for association from memorized patterns.

amplitude $s_0 = \sqrt{a/b}$ when J exceeds a threshold value J_c, while the terms related to H_y act as a perturbation. The connection to such a fundamental nonlinear oscillator model indicates that our results are also valid for various oscillators in nature and engineering. Figure 3b shows an example of nonperturbative vortex dynamics, showing an approximately circular oscillation of the vortex core around the disk center. The phase difference of the oscillation was used to define the colors in the patterns in the associative memory operation. Readers should

note that the plots in Fig. 3b, as well as the results of the numerical simulations shown below, were obtained by solving Eq. (5), while the approximate equations, Eqs. (6) and (7), are used in the model analyses described below.

Demonstration of associative memory

Figure 3c shows the time evolution of the phase difference, $\Delta\psi_i$, obtained by solving Eq. (5) with Eq. (3) substituting for H_y. Note that this solution corresponds to the second step in Fig. 2b. The phase differences saturate to 0 or π within a few hundred nanoseconds. Snapshots of patterns corresponding to this time evolution of the phases are shown in Fig. 3d. The patterns eventually settle to the one to be recognized. Figure 2b shows the result of solving Eq. (5) with Eq. (4) substituting for H_y. Here, Eq. (2) in Eq. (4) is for the three memorized patterns in Fig. 1a. Figure 3e and f show the time evolution of the phase differences and snapshots of the corresponding patterns. Remind that the information of the phases corresponding to the colors of the pixels in the pattern to be recognized is included in the magnetic field in Eq. (4) through $y_j^{(2)}$. Consequently, even though the initial pattern is random, the oscillator phases finally saturate to values corresponding to one of the memorized patterns [Fig. 3f].

The associative memory operation becomes more difficult when there are similar memorized patterns. To clarify this point, let us examine what happens when the number of the memorized patterns is increased, as shown in Fig. 4a from the three in Fig. 1a. The added patterns do not affect the second step in Fig. 2b. For the association corresponding to the third step in Fig. 2b, the magnetic field, defined by Eq. (4), is changed by these new memorized patterns. As a result, the final pattern output resembles none of the memorized ones [Fig. 4b].

This failure of the associative memory operation is due to two reasons. The first is that the pattern "7" is similar to the pattern "1", which should be the one associated. When "7" is excluded from the memorized patterns, the association succeeds, as shown in Fig. 4c. The second reason is that the number of memorized patterns is large. As shown in Fig. 4d, the association succeeds when the memorized patterns include only "1" and "7", the association is succeeded. Therefore, we conclude that an association may fail when the memorized patterns include similar patterns and the number of memorized patterns is large.

To quantify the similarity between patterns A and B, we introduce the degree of overlap:

$$\mathcal{O}\ (\boldsymbol{\xi}^A, \boldsymbol{\xi}^B) \equiv \frac{1}{N} \left| \sum_{i=1}^{N} \xi_i^A \xi_i^B \right|, \tag{8}$$

where $\boldsymbol{\xi}^A = (\xi_1^A, \cdots, \xi_N^A)$ is defined from the color of the ith pixel of pattern A [$\xi_i^A = +(-)1$ when the ith pixel is white (black)]. The overlap becomes 1 when the two patterns are completely identical or their black and white colors are all exchanged (see also "Methods" for the definitions of color and overlap). For example, in the example shown in Figs. 1 and 3, the degree of overlap between the pattern to be recognized and the memorized pattern "0" is $\mathcal{O}\ (\boldsymbol{\xi}^R, \boldsymbol{\xi}^1) = 18/60 = 0.30$. It is $\mathcal{O}\ (\boldsymbol{\xi}^R, \boldsymbol{\xi}^2) = 44/60 \simeq 0.73$ for pattern "1", and $\mathcal{O}\ (\boldsymbol{\xi}^R, \boldsymbol{\xi}^3) = 6/60 = 0.10$ for pattern "2" (the memorized patterns are labelled as $m = 1, 2, 3, \cdots$ while the examples of memorized patterns in this work are "0", "1", "2", etc.; thus, the label m and the corresponding number are off by one). Since the degree of overlap of the pattern to be recognized and "1" is large in the examples in Figs. 1 and 3, pattern "1" should be associated in this case. On the other hand, in the example shown in Fig. 4, the overlap between the pattern to be recognized [Fig. 1(b)] and "7" is also relatively large, i.e., $\mathcal{O}\ (\boldsymbol{\xi}^R, \boldsymbol{\xi}^8) = 32/60 \simeq 0.53$. In addition, the overlap between the memorized patterns "1" and "7", $\mathcal{O}\ (\boldsymbol{\xi}^2, \boldsymbol{\xi}^8) = 28/60 \simeq 0.47$, is also relatively large compared with those between the other patterns; for example, the overlap between "1" and "8" is $\mathcal{O}\ (\boldsymbol{\xi}^2, \boldsymbol{\xi}^9) = 2/60 \simeq 0.03$ (see also Supplementary Information, where the overlaps of the ten memorized patterns are summarized). Accordingly, when the memorized patterns include "1" and "7", the virtual oscillator network cannot associate a correct pattern, and the final pattern produced corresponds to none of the memorized ones. Similarly, when the number of memorized patterns is large, there might be patterns having large overlaps and the association fails.

In summary, we have shown that the virtual oscillator network based on the algorithm in Fig. 2b can perform the associative memory operation. Its accuracy, however, is low when the memorized patterns include some patterns having large overlaps and there is a large number of memorized patterns. Note that the maximum number of patterns that can be memorized by neural network is approximately $N/(2 \log N)$[8]. It would be of interest if such a formula can be derived for virtual oscillator networks in future.

We examined the associative memory operation for various cases, i.e., for different patterns to be recognized, and studied the rate of the accurate association; see Supplementary Information.

Discussion

Here we discuss the principles of the associative memory operation analytically by using Eqs. (6) and (7). As mentioned above, the operation consists of three steps, and in each step, the oscillator output is divided into N parts. In what follows, we denote the phase of the vortex core during the ith part of the kth step as $\psi_i^{(k)}$. We also assume that the oscillation amplitude s_0 is approximately constant because the current density is fixed. Therefore, the oscillation frequency, $f = \Omega/(2\pi) = [\kappa/(2\pi G)](1 + \zeta s_0^2)$, is also approximately constant (see also "Methods" for the analytical solution of the Thiele equation).

The phase in the second step obeys,

$$\dot{\psi}_i^{(2)} = \Omega + \frac{c\mu^*}{GRs_0} \mathcal{H} \sum_{\ell=1}^{N} \xi_i^R \xi_\ell^R y_\ell^{(1)} \sin \psi_i^{(2)}. \tag{9}$$

Thus, the phase difference between the ith and jth parts obeys,

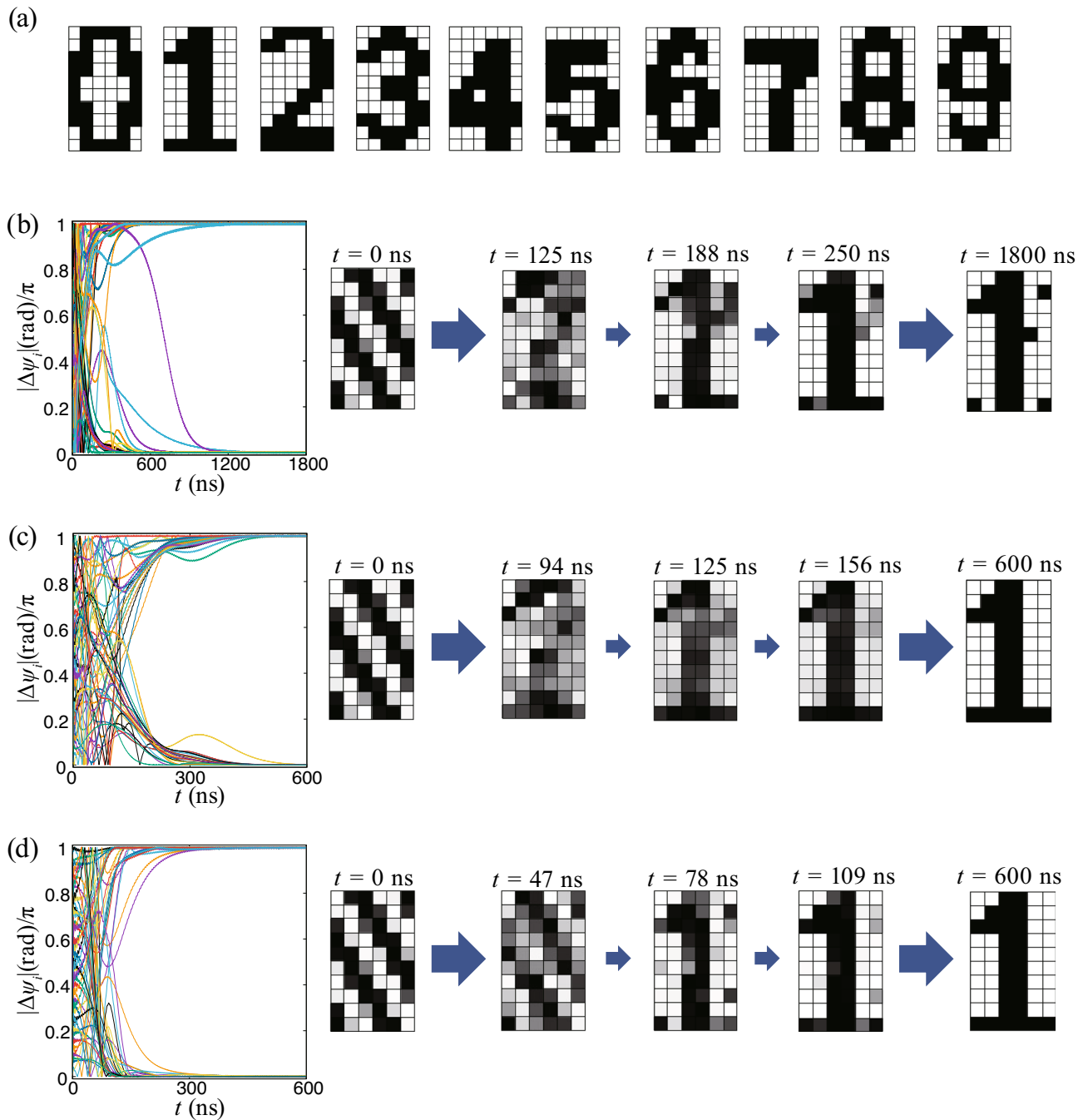**Figure 4.** Problem of associative memory operation when the similarity between the memorized patterns is high and the number of patterns is large. (**a**) Ten ($N_m = 10$) memorized patterns, "0", "1"; $\cdots$,"9". (**b**) Time evolution of the phase difference during the association and snapshots of the corresponding pattern. In this case, the memorized patterns include both "1" and "7". Because of their similarity, the pattern does not finally saturate to "1". (**c**) When "7" is removed from the memorized patterns ($N_m = 9$), the association is successful, even though there are nine remaining memorized patterns. (**d**) The association is successful when the memorized patterns include only "1" and "7".

$$\dot{\psi}_i^{(2)} - \dot{\psi}_j^{(2)} = \frac{c\mu^*}{GRs_0} \mathscr{H}\left(\sum_{\ell=1}^{N} \xi_\ell^R y_\ell^{(1)}\right)\left(\xi_i^R \sin\psi_i^{(2)} - \xi_j^R \sin\psi_j^{(2)}\right). \tag{10}$$

The steady state condition on the phase difference leads to

$$\xi_i^R \sin\psi_i^{(2)} - \xi_j^R \sin\psi_j^{(2)} = 0. \tag{11}$$

Note that $\xi_i^{\mathrm{R}} = +(-)1$ when the color at the ith pixel of the pattern to be recognized is white (black). Therefore, $\psi_i^{(1)}$ and $\psi_j^{(1)}$ will be in-phase $\psi_i^{(1)} = \psi_j^{(1)}$ [anti-phase $\psi_i^{(1)} = \psi_j^{(1)} \pm \pi$] when the colors of the ith and jth pixels are the same (opposite). As a result, the phase differences in the second step saturate to 0 or π corresponding to the white or black in the pattern to be recognized. Note that this synchronization is caused by a feedforward input from the first step, which corresponds to the second term on the right-hand side in Eq. (9). Here, the term $\sum_{\ell=1}^{N} \xi_\ell^{\mathrm{R}} y_\ell^{(1)}$ in Eq. (9) is the sum of the N oscillator outputs $y_\ell^{(1)}$ in the first step, multiplied by the factor ξ_ℓ^{R} determining the pixel color of the pattern to be recognized, and is common for all i of Eq. (9). Equation (9) also includes a factor ξ_i^{R}, which determines the sign of the input. Regarding these facts, the feedforward input has only two values, depending on the value of ξ_i^{R}. The phase synchronization among the N parts in the second step is the result of forced synchronization with respect to this feedforward input, and the phase difference has only two values, 0 or π, depending on the value of ξ_i^{R}. This mechanism is in contrast with that of the previous work[21], where a mutual synchronization is the origin of the associative memory operation. Also, the method is different from the previous works[38,39]. In Ref.[38], a forced synchronization of frequency with respect to an external signal was studied, while the input signal in the present work is generated by the oscillator output itself and the phase synchronization plays the central role in the associative memory operation. In Ref.[39], a delayed-feedback was used to generate input signal, while the input signal in the present work is generated by multiplying appropriated weight to perform the associative memory operation.

We also note that, when $y_\ell^{(1)}$ is a simple trigonometric function, its linear combination, $\sum_{\ell=1}^{N} \xi_\ell^{\mathrm{R}} y_\ell^{(1)}$, is also a trigonometric function with the same frequency and a different phase. According to the above discussion, the phase of the term $\sum_{\ell=1}^{N} \xi_\ell^{\mathrm{R}} y_\ell^{(1)}$ does not play any role to excite forced synchronization among the N parts. Thus, the term $\sum_{\ell=1}^{N} \xi_\ell^{\mathrm{R}} y_\ell^{(1)}$ could be replaced by, for example, $y_1^{(1)}$. In this case, it is unnecessary to measure other $(N-1)$ $y_\ell^{(1)}$ ($\ell = 2, 3, \cdots, N$) in the first step in Fig. 2b, although we solved the equation of motion for N virtual oscillators to clarify similarities and differences between the second and third step. When $(N-1)$ parts in the first step are omitted for simplicity, the power consumption to drive the oscillator in the virtual oscillator network is proportional to $2N+1$, where $2N$ comes from the second and third steps in Fig. 2b. On the other hand, the power consumption in the conventional oscillator network is proportional to $2N$ because N oscillators are driven two times, as implied in Fig. 2a. For a large N, the power consumption of two oscillator networks are comparable. The time required for the operation increases linearly as N increases, which is not suitable for practical applications, although the same might be true for a conventional oscillator network because the relaxation time of the phase will depend on the number of the oscillators. the time of a conventional (coupled) oscillator network might also increase as N increases. However, the virtual oscillator network has an advantage from a viewpoint of reliability, as discussed below.

Next, we focus on the third step, where the phase during the ith part obeys

$$\dot{\psi}_i^{(3)} = \Omega + \frac{c\mu^*}{GRs_0} \mathscr{H}' \frac{1}{N_{\mathrm{m}}} \sum_{m=1}^{N_{\mathrm{m}}} \sum_{\ell=1}^{N} \xi_i^m \xi_\ell^m y_\ell^{(2)} \sin \psi_i^{(3)}. \tag{12}$$

Since the oscillators in the second step are in the synchronized state, the output $y_\ell^{(2)}$ can be expressed as $y_\ell^{(2)} = \xi_\ell^{\mathrm{R}} \xi_1^{\mathrm{R}} y_1^{(2)}$, where $y_1^{(2)}$ is the output of the first part in the second step. We substitute this relation into Eq. (12) and assume that

$$\sum_{\ell=1}^{N} \xi_\ell^m \xi_\ell^{\mathrm{R}} \simeq \delta_{m,\mathscr{A}} \sum_{\ell=1}^{N} \xi_\ell^m \xi_\ell^{\mathrm{R}}, \tag{13}$$

where the symbol $\mathscr{A}$ corresponds to a pattern in the memorized patterns that resembles the pattern to be recognized. The assumption (13) means that only a pattern having a large degree of overlap with the pattern to be recognized contributes to the feedforward input. The other memorized patterns, which are greatly different from the pattern to be recognized, do not contribute to the feedforward input because of their small overlap. When the assumption is satisfied, Eq. (12) becomes

$$\dot{\psi}_i^{(3)} = \Omega + \frac{c\mu^*}{GRs_0} \mathscr{H}' \frac{1}{N_{\mathrm{m}}} y_1^{(2)} \xi_1^{\mathrm{R}} \left(\sum_{\ell=1}^{N} \xi_\ell^{\mathscr{A}} \xi_\ell^{\mathrm{R}} \right) \xi_i^{\mathscr{A}} \sin \psi_i^{(3)}. \tag{14}$$

Equation (14) is similar to Eq. (9), and therefore, the steady-state condition of the phase difference between the ith and jth parts in the third step is given by

$$\xi_i^{\mathscr{A}} \sin \psi_i^{(3)} - \xi_j^{\mathscr{A}} \sin \psi_j^{(3)} = 0. \tag{15}$$

Equation (15) means that in-phase or anti-phase synchronization between the N parts occurs, and the phase differences in the third step saturate to 0 or π corresponding to the white or black colors in a memorized pattern most resembling the one to be recognized.

The operation principle is based on Eq. (13). Equation (13) is satisfied if there is only one pattern that has a large degree of overlap with the pattern to be recognized. On the other hand, if there are other patterns having large overlaps with the pattern to be recognized, Eq. (13) is not satisfied. In this case, Eq. (15) is not necessarily

satisfied, and the colors in the steady state in the third step might be different from the pattern most resembling the one to be recognized or they might be gray (neither black nor white); see also Supplementary Information.

Our analysis also assumed that the oscillation frequencies of the N parts are the same. This assumption is a natural one because each part is obtained from a single oscillator. Technically speaking, the oscillation frequency in each part is varied by changing the magnitude of the electric current. If the oscillation frequencies of the ith and jth parts, denoted as $\Omega_i/(2\pi)$ and $\Omega_j/(2\pi)$, are different, the right-hand side of Eq. (10) has an additional term $\Omega_i - \Omega_j$. In such a case, the phase difference is not well defined because ψ_i and ψ_j oscillate with different frequencies. Even if we introduce an instantaneous phase by, for example, making a Hilbert transformation, as was done in experiments[40], the phase difference still does not necessarily saturate to 0 or π. In such a case, the associative memory operation fails. Therefore, there is no reason to change the oscillation frequency in each part. This fact also indicates an advantage to using the virtual oscillator network. In the conventional oscillator network, variations in the oscillation frequency naturally appear because inhomogeneities in the parameters of the oscillators are unavoidable, and such variations lead to the failure of the associative memory operation[21]. The virtual oscillator network does not have such variation and thus would be a more reliable associative memory. A weak point of the present proposal is, on the other hand, that the method requires a computer to store the output signal in each step, which is not preferable for practical applications. We would like to keep this issue as a future work.

In conclusion, we described the concept of the associative memory operation by a virtual oscillator network and performed numerical simulations. The operation consists of three steps, where the output of one step is sent to the next step with weights defined by the Hebbian rule. In this sense, the virtual oscillator network can be regarded as a hybrid combination of a coupled oscillator network and a feedforward neural network. The network successfully associated black-and-white patterns with a few memorized patterns. However, it failed to make an association when the number of memorized patterns was large (ten compared to three) and some of the memorized patterns resembled each other. We also developed a theoretical analysis and clarified that the origin of the associative memory operation is forced synchronization driven by feedforward input. Either in-phase or anti-phase synchronization was excited among the oscillators and provides appropriate correspondence between the oscillator phases and the colors in the patterns. The virtual oscillator network is more reliable than a conventional oscillator network, which is affected by unavoidable inhomogeneities among the oscillators.

Methods
Definitions of color and overlap
By convention, the first pixel (the pixel in the top left-hand corner of a pattern) is always white. The pattern should be regarded as the same even when all of the black and white pixels are swapped for each other, Mathematically, this means that $\sum_{i=1}^{N} \xi_i^A \xi_i^B = N$ when the patterns A and B are completely the same, and $\sum_{i=1}^{N} \xi_i^A \xi_i^B = -N$ when patterns A and B represent the same pattern but their black and white colors are completely swapped. According to this definition of the same figure, the maximum number of the difference between two patterns is $N/2$; in this case, the degree of overlap is zero (see also the discussion on noise in Supplementary Information).

Models of conventional and virtual oscillator networks
The conventional oscillator network for the associative memory operation[21] is based on the Kuramoto model[37]. The Kuramoto model describes the oscillator dynamics with a generalized phase, θ. Moreover, the oscillators interact instantaneously, and the phase of the ith oscillator obeys

$$\dot{\theta}_i = \omega + \mathscr{D} \sum_{j=1}^{N} w_{ij} \sin\left(\theta_i - \theta_j\right), \tag{16}$$

where $\omega/(2\pi)$ is the oscillation frequency while $\mathscr{D}$ is the interaction strength. For simplicity, we will assume that all oscillators share the same values of ω and $\mathscr{D}$. The weight w_{ij} is given by Eq. (1) or (2) depending on the step of the procedure. In the *LC*-circuit model[21], $\mathscr{D} w_{ij}$ is proportional to the transconductance. The phase difference between the ith and jth oscillators obeys

$$\dot{\theta}_i - \dot{\theta}_j = \mathscr{D} \left[\sum_{\ell=1}^{N} w_{i\ell} \sin\left(\theta_i - \theta_\ell\right) - \sum_{\ell=1}^{N} w_{j\ell} \sin\left(\theta_j - \theta_\ell\right) \right]. \tag{17}$$

In a limiting case of only two oscillators ($N = 2$), the phase difference obeys

$$\dot{\theta}_1 - \dot{\theta}_2 = 2\mathscr{D} w_{12} \sin\left(\theta_1 - \theta_2\right), \tag{18}$$

and the in-phase (anti-phase) synchronization of θ_1 and θ_2 is a stable fixed point when $\mathscr{D} w_{12}$ is negative (positive). The phase differences of $\theta_i - \theta_j = 0, \pi$ are always fixed points even when there are a large number of oscillators ($N \geq 3$). Accordingly, the phase differences in the conventional oscillator network saturate to the in-phase or anti-phase state, which thereby enables the associative memory operation.

In the presence of frequency variations, the right-hand side of Eq. (17) has an additional term $\omega_i - \omega_j$. In this case, the phase difference is not stabilized, and this instability leads to an inaccurate associative memory operation[21].

The Thiele equation is slightly different from the Kuramoto model in the following ways. First, the Thiele equation uses the phase ψ, which describes the vortex core's position in the xy plane, instead of a generalized phase. This is because the quantity measured in experiments is the vortex core's position, and the phase

synchronization studied in the experiments[40] corresponds to that of ψ, not a generalized phase θ. Note that we can introduce a generalized phase analytically as $\theta = \psi + [\zeta\kappa/(Gb)]\ln(s/s_0)$ with a phase sensitivity function $\mathbf{Z} = (-\sin\theta + [\zeta\kappa/(Gb)]\cos\theta, \cos\theta + [\zeta\kappa/(Gb)]\sin\theta, 0)/s_0$. The analysis is mostly unchanged with the generalized phase, so we decided to use ψ for simplicity. Second, the equation of motion for the phase difference, Eq. (10), includes a term $\sin\psi_i - \sin\psi_j$ whereas the Kuramoto model often uses an interacting term proportional to $\sin(\theta_i - \theta_j)$. More generally, the interaction term in the Kuramoto model can be assumed to be a function of the phase difference, $\theta_i - \theta_j$ after applying an averaging technique with respect to a fast variable (see Ref.[37] for details). The difference between the two models might however be insignificant; notice that, by using formulas, $\sin x - \sin y = 2\cos[(x+y)/2]\sin[(x-y)/2]$ and $\sin x + \sin y = 2\sin[(x+y)/2]\cos[(x-y)/2]$ and applying the averaging technique, the interaction term in our model can be approximated as a function of $\theta_i - \theta_j$. Third, as mentioned above, the input term in the virtual oscillator network consists of the oscillator output from the previous step, while the interaction in the Kuramoto model is instantaneous. Because of these differences, the associative memory operation by the virtual oscillator network is significantly different from those of conventional coupled oscillator networks on which previous experiments and the theoretical analyses have been conducted.

Parameters in the Thiele equation

Spin torque oscillators (STOs) mainly consist of a ferromagnetic metal/insulating layer/ferromagnetic metal trilayer. The first ferromagnetic layer of the trilayer is called the free layer and is where the magnetic vortex forms. The second ferromagnetic layer having a uniform magnetization is called the reference layer. When electric current is injected into STOs, spin-transfer torque[41–43] is excited on the magnetic moments in the free layer and drives their dynamics[35,36]. The output signal from the STOs depends on the relative angle between the magnetizations in the free and reference layers.

The definitions and physical meanings of the parameters in Eq. (5) are as follows. The parameters $G = 2\pi pML/\gamma$ and $D = -(2\pi\alpha ML/\gamma)[1 - (1/2)\ln(R_0/R)]$ consist of the polarity $p(=\pm 1)$ of the vortex core, the saturation magnetization M, the thickness L of the ferromagnet, the gyromagnetic ratio γ, the Gilbert damping constant α, and the vortex radius R_0. The chirality $c(\pm 1)$ of the vortex core also appears in Eq. (5). The parameters κ and ζ relate to a magnetic potential energy defined as $W = (\kappa/2)[1 + (\zeta/2)s^2]|\mathbf{X}|^2$. The dimensionless parameter ξ is introduced to describe the nonlinear damping in a highly excited state[35]. The parameter κ relates to the material parameters as $\kappa = (10/9)4\pi M^2L^2/R$[35]. The parameter $a_J = \pi\hbar P/(2e)$ includes the reduced Planck constant $\hbar$, spin polarization P of the electric current, and the elementary charge $e(>0)$. The vector $\mathbf{p} = (p_x, 0, p_z)$ is the unit vector pointing in the magnetization direction in the reference layer. Here, we assume that $\mathbf{p}$ lies in the xz plane, by convention. As a result, the output signal from the vortex STO is proportional to the y component of the vortex core's position. The parameter μ^* is πMLR.

The material parameters used in this study were taken from typical experiments and simulations[35,36,44]: $M = 1300$ emu/cm^3, $\gamma = 1.764 \times 10^7$ rad/(Oe s), $\alpha = 0.01$, $L = 5$ nm, $R = 187.5$ nm, $R_0 = 10$ nm, $P = 0.7$, $\xi = 2.0$, and $\zeta = 0.1$. The polarity and chirality were assumed to be $p = +1$ and $c = +1$, for simplicity. The magnetization direction in the reference layer was $\mathbf{p} = (\sin 60°, 0, \cos 60°)$. An electric current I of 1 mA corresponded to a current density J of 0.9 MA/cm^2. The electric current in the numerical simulations was set to 4.0 mA.

We do not include field-like torque in the Thiele equation, which is expressed as $-cb_JJRp_x\mathbf{e}_y$ in Eq. (5); see, for example, Ref.[45]. This is because its magnitude was not visible in an experiment using CoFeB/MgO based STO[23]. One might consider to inject the input through the field-like torque, instead of the torque due to the external magnetic field as we have done. However, the modulation of the field-like torque requires that of electric current, which leads to the modulation of the frequency of the STO. Since the advantage of our proposal is that the frequency is unique during the operation, we do not prefer to use the field-like torque for injecting the input.

Analytical solution of the Thiele equation

The Gilbert damping constant α is often small, in such cases, $|D|/G \simeq \alpha \ll 1$. Also, the radius R_0 of the vortex core is much shorter than the disk radius, R. Therefore, by neglecting terms related to R_0 and higher-order terms of α, we can approximate Eq. (5) as Eqs. (6) and (7) in terms of $s = |\mathbf{X}|/R$ and $\psi = \tan^{-1}(Y/X)$. The approximated Thiele equation without magnetic field is

$$\dot{s} = as - bs^3, \tag{19}$$

$$\dot{\psi} = \frac{\kappa}{G}\left(1 + \zeta s^2\right). \tag{20}$$

These equations are identical to the Stuart-Landau equation[37], which was introduced by Landau to describe the evolution of turbulence phenomenologically and was derived from hydrodynamics by Stuart. This equation provides one of the simplest example of Hopf bifurcation. A stable solution of s is $s_0 = \sqrt{a/b}$ (0) for $a > (<)0$, or equivalently, $J/J_c > (<)1$. When $J/J_c > 1$, i.e., the current density J exceeds a threshold value J_c, the vortex core oscillates around the disk center with an oscillation amplitude s_0 and the frequency $f = [\kappa/(2\pi G)](1 + \zeta s_0^2)$. Note that the oscillation frequency is proportional to the current density J through the term $s_0^2 = a/b$ ($a \propto J$), which has been confirmed by both experiments and simulations[35,36]. Even in the presence of the magnetic field, the oscillation frequency remains f, if the input strength is weak.

The solution of s obtained from the exact Thiele equation, Eq. (5), shows a small oscillation around s_0[46]. This means that the trajectory of a limit-cycle oscillation is approximately circular but also has a small amplitude modulation. This small modulation is caused by the term $ca_JJR_0p_x\mathbf{e}_x$ in Eq. (5), which breaks the axial symmetry

of the dynamics around the z-axis. The deviation of s from s_0 is, however, negligible, and the oscillation trajectory is approximately circular, as shown in Fig. 3b. Therefore, it is reasonable to omit the term from Eqs. (6) and (7). Note that this term arises from the in-plane component p_x of the magnetization in the reference layer. p_x plays a role in experiments for the following reason. Recall that the output signal measured in experiments depends on the relative angle of the magnetizations in the free and reference layers. Since the vortex core is located in the xy plane, a finite p_x is necessary to detect its position. On the other hand, the z component p_z is also necessary because the spin-transfer torque originating from it excites the limit-cycle oscillation of the vortex core. In fact, the threshold current density $J_c = |D|\kappa/(Ga_J p_z)$ is inversely proportional to p_z; therefore, if p_z is zero, J_c becomes infinite and the oscillation cannot be excited. In experiments[28,40], the magnetization initially pointed in an in-plane direction, where $p_z = 0$. A finite p_z was induced by applying an external magnetic field in the z direction.

According to Eqs. (6) and (7), one might consider that the magnetic field changes the value of s from s_0 and modifies the oscillation frequency. Such a frequency shift is, however, negligibly small, which can be discussed accordingly. First, remind that the frequency of the magnetic field applied during the second step is the frequency of the vortex core without the magnetic field because it consists of the output during the first step. The fact that the phases in the second step are saturated to 0 or π, as shown in Fig. 3c, indicates that the forced phase synchronization occurs, and the frequency of the vortex core in the second step is the same with that in the first step. Second, let us roughly estimate the frequency shift by the application of the magnetic field. The change of s by the magnetic field will be maximized when the phase of the magnetic field H_y in Eq. (6) is the same with ψ. In this case, the magnitude of the last term in Eq. (6), averaged over a precession period $\tau = 1/f$, is about $[c\mu^*/(2GR)]H_y\tau \sim (\gamma/2)\mathscr{H}\tau$. The period τ is about 5 ns while $\mathscr{H}$ is on the order of 1 Oe; see next section. Accordingly, the shift Δs of s by the application of the magnetic field is less than 0.1 at maximum. As mentioned, the oscillation frequency is proportional to $1 + \zeta s^2$. Using $\zeta = 0.1$ and $s_0 \simeq 0.6$, estimated from Fig. 3b, the frequencies with and without Δs, which are proportional to $1 + \zeta s_0^2$ and $1 + \zeta(s_0 + \Delta s)^2$, respectively, differ only 1 % at maximum. Therefore, we consider that the frequency modulation by the application of the magnetic field is negligible.

One might be of interested in the applicability of the Thiele equation. While the original Thiele equation assumes a translation symmetry in an infinite space, a finite-size effect of nanostructured may restrict the applicability of the equation. Therefore, the Thiele equation had been applied to analyses on small-amplitude dynamics[47]. There have been, at the same time, several efforts to make the equation applicable to large-amplitude dynamics. For example, adding nonlinear frequency and damping terms is one approach[35,36], which is also used in the present work, where the additional terms are characterized by the dimensionless parameters ξ and ζ. Adding further higher-order nonlinear terms is also investigated recently[48–50]. It was also shown that the Thiele equation is applicable to analyze small-amplitude dynamics, and effort has been made to extrapolating it to a large-amplitude dynamics, such as vortex-core expulsion, although there are some limitations[51]. In the present study, we use the model developed in Refs.[35,36] due to the following reasons. First, the applicability of the model to wide ranges of parameters has been verified by comparison with experiments[35,36,44]. Second, adding higher-order nonlinear terms does not change main conclusion in this work. These terms might change, for example, current dependence of the oscillation frequency. In the present work, however, the frequency is kept constant, and thus, adding such terms do not play a central role in the associative memory operation. Third, the Thiele equation with the present approximation clarifies the connection between spintronics and other research fields such as nonlinear science and computer science. This is because the equation can be reduced to the Stuart-Landau equation, as mentioned above. The Stuart-Landau equation has a long history, as in the case of the Thiele equation, and has been frequently used in nonlinear science[37,52]. The present work indicates that the Stuart-Landau oscillator can be emulated in nanostructures and therefore, prompts communications between spintronics and other research fields. Therefore, although we understand that there have been great efforts[48–50,53] for the validity and applicability of the Thiele equation, we use the model developed in Refs.[35,36]. Note that the Oersted field generated in the current, discussed in these previous works, does not play a role in the associative memory operation because the current magnitude is kept constant during the operation. Also, since the external magnetic field induces forced synchronization, a frequency shift due to an external magnetic field studied in the previous work[48] does not exist in the present algorithm.

Details of the numerical simulations

The associative memory operation in the virtual oscillator network consists of three steps. The initial state of the vortex core in each step is prepared by adding a thermal activation to the Thiele equation and solving it in the absence of magnetic field, as is done in Ref.[45]. The torque due to the thermal activation gives an additional term, $-\eta_x \mathbf{e}_x - \eta_y \mathbf{e}_y$, to the left-hand side of Eq. (5), which obeys the fluctuation-dissipation theorem,

$$\langle \eta_i(t)\eta_j(t') \rangle = 2k_B T|D|\delta_{ij}\delta(t - t'), \tag{21}$$

where the temperature T is 300 K. The solution of the Thiele equation in each step is divided into $N = 60$ parts, where the time width of each part is denoted as $\tilde{t}$. In the experiment[23], a certain time period was inserted between these parts to remove their correlation. In contrast, our numerical simulations used parallel computations, wherein the initial state of each part was randomly prepared using the method described above. The value of $\tilde{t}$ was changed depending on the number of memorized patterns, as well as the number of noisy pixels in the pattern to be recognized. For example, $\tilde{t}$ is 750 ns in Fig. 3c. For all cases, $\tilde{t}$ was divided into $\tilde{n} = \tilde{t}/t_p$ parts, where $t_p = 0.125$ ns.

Now let us explain the meanings of $y_\ell^{(1)}$ and $y_\ell^{(2)}$ in Eqs. (3) and (4). Since they are defined in a similar manner, we will describe only $y_\ell^{(1)}$. When defining the magnetic field in Eq. (3), it is convenient to reset the time

origin for each part; i.e., each of the N parts runs from $t = 0$ to $t = \tilde{t}$. Remember that the output from the STO is proportional to the y component of the vortex core's position, Y. We denote the solution of the normalized y component, $y = Y/R$ ($0 \leq y \leq 1$), during the ℓth part in the first step as y_ℓ. Then, $y_\ell^{(1)}$ is made from y_ℓ as follows,

$$y_\ell^{(1)} = \sum_{n=0}^{\tilde{n}-1} y_\ell(nt_\mathrm{p})\big\{\Theta(t - nt_\mathrm{p}) - \Theta[t - (n+1)t_\mathrm{p}]\big\}, \tag{22}$$

where $\Theta(t)$ is a step function. Note that $\Theta(t - nt_\mathrm{p}) - \Theta[t - (n+1)t_\mathrm{p}]$ is 1 for $nt_\mathrm{p} \leq t < (n+1)t_\mathrm{p}$ and is zero for the other times; thus, it has a pulse shape. Equation (22) means that the input strength is constant for $nt_\mathrm{p} \leq t < (n+1)t_\mathrm{p}$ and is proportional to $y_\ell(t)$ at $t = nt_\mathrm{p}$. $\tilde{n}$ is the number of input pulses. There are two reasons to shape the output y into a pulse. The first one relates to numerical simulations. In this work, the Thiele equation was solved within a time increment of $\Delta t = 0.005$ ns, which is shorter than the pulse width t_p. It was, however, impractical to store the output at each Δt step because the amount would have been huge. Second, there is a technical limitation in real experiments on a measurable time step. The value we used, $t_\mathrm{p} = 0.125$ ns, is close to shortest possible time step in an experiment[23]. Because of these reasons, we define $y_\ell^{(1)}$ used in the magnetic field, Eq. (3), as a pulse input. At the same time, we emphasize that t_p is much shorter than an oscillation period of the vortex core, $1/f = 4.48$ ns ($f = 223$ MHz). In addition, the pulse-shaped $y_\ell^{(1)}$s are continuously injected. Therefore, the magnetic field can be approximately regarded as a continuously oscillating signal with respect to the STO.

The strength of the input $\mathscr{H}$ in the second step is 1.0 Oe, while that in the third step is $\mathscr{H}' = N_\mathrm{m} \times 0.2$ Oe. Here, we increase $\mathscr{H}'$ as the number N_m of memorized patterns increases. This is because the time necessary to reach a steady state becomes long as N_m increases; therefore, to perform the numerical simulations efficiently, the input strength should be made to increase with N_m.

Data availability
The datasets used and/or analyses during the current study available from the corresponding author on reasonable request.

References
1. Kohonen, T. *Self-Organization and Associative Memory* Vol. 8 (Springer Science and Business Media, 2012).
2. Nakano, K. Associatron-A model of associative memory. *IEEE Trans. Syst. Man Cybern.* **2**, 380–388 (1972).
3. Kohonen, T. Correlation matrix memories. *IEEE Trans. Comput.* **100**, 353–359 (1972).
4. Anderson, J. A. A simple neural network generating an interactive memory. *Math. Biosci.* **14**, 197–220 (1972).
5. Hopfield, J. J. Neural network and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 2554–2558 (1982).
6. Amari, S.-I. Neural theory of association and concept-formation. *Biol. Cybern.* **26**, 175–185 (1977).
7. van Hemmen, J. L. & Kühn, R. Nonlinear neural networks. *Phys. Rev. Lett.* **57**, 913 (1986).
8. McEliece, R. J., Posner, E., Rodemich, E. & Venkatesh, S. The capacity of the Hopfield associative memory. *IEEE Trans. Inf. Theory* **33**, 461–482 (1987).
9. Waugh, F. R., Marcus, C. M. & Westervelt, R. M. Fixed-point attractors in analog neural computation. *Phys. Rev. Lett.* **64**, 1986 (1990).
10. Morita, M. Associative memory with nonmonotone dynamics. *Neural Netw.* **6**, 115–126 (1993).
11. Yoshizawa, S., Morita, M. & Amari, S. Capacity of associative memory using a nonmonotonic neuron model. *Neural Netw.* **6**, 167–176 (1993).
12. Bollé, D. & Huyghebaert, J. Mixture states and storage of biased patterns in potts-glass neural networks. *Phys. Rev. E* **48**, 2250 (1993).
13. Krebs, P. R. & Theumann, W. K. Categorization in the symmetrically dilute hopfield network. *Phys. Rev. E* **60**, 4580 (1999).
14. McGraw, P. N. & Menzinger, M. Bistable gradient networks. i. Attractors and pattern retrieval at low loading in the thermodynamic limit. *Phys. Rev. E* **67**, 016118 (2003).
15. Zhao, H. Designing asymmetric neural networks with associative memory. *Phys. Rev. E* **70**, 066137 (2004).
16. Borders, W. A. *et al.* Analogue spin-orbit torque device for artificial-neural-network-based associative memory operation. *Appl. Phys. Express* **10**, 013007 (2017).
17. Hoppensteadt, F. C. & Izhikevich, E. M. Associative memory of weakly connected oscillators. In *Proceedings of International Conference on Neural Networks (ICNN'97)* Vol. 2 (eds Hoppensteadt, F. C. & Izhikevich, E. M.) 1135–1138 (IEEE, 1997).
18. Hoppensteadt, F. C. & Izhikevich, E. M. Oscillatory neurocomputers with dynamic connectivity. *Phys. Rev. Lett.* **82**, 2983 (1999).
19. Fernando, C., Michele, B. & Marco, G. Weakly connected oscillatory network models for associative and dynamic memories. *Int. J. Bifurc. Chaos* **17**, 4365–4379 (2007).
20. Mirchev, M., Basnarkov, L., Corinto, F. & Kocarev, L. Cooperative phenomena in networks of oscillators with non-identical interactions and dynamics. *IEEE Trans. Circuits Syst. I Regul. Pap.* **61**, 811–819 (2013).
21. Maffezzoni, P., Bahr, B., Zhang, Z. & Daniel, L. Oscillator array models for associative memory and pattern recognition. *IEEE Trans. Circuits Syst. I Regul. Pap.* **62**, 1591–1598 (2015).
22. Prasad, N., Mukim, P., Madhavan, A. & Stiles, M. D. Associative memories using complex-valued hopfield networks based on spin-torque oscillator arrays. *Neuromorph. Comput. Eng.* **2**, 034003 (2022).
23. Tsunegi, S., Taniguchi, T., Yakushiji, K., Yuasa, S. & Kubota, H. *Associative Memory Operation Using Spin Torque Oscillator* (The 83rd Japan Society of Applied Physics Autumn Meeting, 2022).
24. Locatelli, N., Cros, V. & Grollier, J. Spin-torque building blocks. *Nat. Mater.* **13**, 11 (2014).
25. Grollier, J., Querlioz, D. & Stiles, M. D. Spintronic nanodevices for bioinspired computing. *Proc. IEEE* **104**, 2024 (2016).
26. Torrejon, J. *et al.* Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428 (2017).
27. Grollier, J. *et al.* Neuromorphic spintronics. *Nat. Electron.* **3**, 360 (2020).
28. Kamimaki, A. *et al.* Chaos in spin-torque oscillator with feedback circuit. *Phys. Rev. Res.* **3**, 043216 (2021).

29. Thiele, A. A. Steady-state motion of magnetic domains. *Phys. Rev. Lett.* **30**, 230 (1973).
30. Guslienko, K. Y., Han, X. F., Keavney, D. J., Divan, R. & Bader, S. D. Magnetic vortex core dynamics in cylindrical ferromagnetic dots. *Phys. Rev. Lett.* **96**, 067205 (2006).
31. Guslienko, K. Y. Low-frequency vortex dynamic susceptibility and relaxation in mesoscopic ferromagnetic dots. *Appl. Phys. Lett.* **89**, 022510 (2006).
32. Ivanov, B. A. & Zaspel, C. E. Excitation of spin dynamics by spin-polarized current in vortex state magnetic disks. *Phys. Rev. Lett.* **99**, 247208 (2007).
33. Khvalkovskiy, A. V., Grollier, J., Dussaux, A., Zvezdin, K. A. & Cros, V. Vortex oscillations induced by spin-polarized current in a magnetic nanopillar: Analytical versus micromagnetic calculations. *Phys. Rev. B* **80**, 140401(R) (2009).
34. Guslienko, K. Y., Aranda, G. R. & Gonzalez, J. Spin torque and critical currents for magnetic vortex nano-oscillator in nanopillars. *J. Phys. Conf. Ser.* **292**, 012006 (2011).
35. Dussaux, A. *et al.* Field dependence of spin-transfer-induced vortex dynamics in the nonlinear regime. *Phys. Rev. B* **86**, 014402 (2012).
36. Grimaldi, E. *et al.* Response to noise of a vortex based spin transfer nano-oscillator. *Phys. Rev. B* **89**, 104404 (2014).
37. Kuramoto, Y. *Chemical Oscillations, Waves, and Turbulence* (Dover, 2003).
38. Romera, M. *et al.* Vowel recognition with four coupled spin-torque nano-oscillators. *Nature* **563**, 230 (2018).
39. Skowronski, W., Checinski, J., Zietek, S., Yakushiji, K. & Yuasa, S. Micorwave magnetic field modulation of spin torque oscillator based on perpendicular magnetic tunnel junctions. *Sci. Rep.* **9**, 19091 (2019).
40. Tsunegi, S. *et al.* Physical reservoir computing based on spin torque oscillator with forced synchronization. *Appl. Phys. Lett.* **114**, 164101 (2019).
41. Slonczewski, J. C. Current-driven excitation of magnetic multilayers. *J. Magn. Magn. Mater.* **159**, L1 (1996).
42. Berger, L. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B* **54**, 9353 (1996).
43. Slonczewski, J. C. Currents, torques, and polarization factors in magnetic tunnel junctions. *Phys. Rev. B* **71**, 024411 (2005).
44. Tsunegi, S. *et al.* Control of the stochastic response of magnetization dynamics in spin-torque oscillator through radio-frequency magnetic fields. *Sci. Rep.* **11**, 16285 (2021).
45. Imai, Y., Tsunegi, S., Nakajim, K. & Taniguchi, T. Noise-induced synchronization of spin-torque oscillators. *Phys. Rev. B* **105**, 224407 (2022).
46. Yamaguchi, T. *et al.* Periodic structure of memory function in spintronics reservoir with feedback current. *Phys. Rev. Res.* **2**, 023389 (2020).
47. Goto, M. *et al.* Electric spectroscopy of vortex states and dynamics in magnetic disks. *Phys. Rev. B* **84**, 064406 (2011).
48. Araujo, F. A., Chopin, C. & de Wergifosse, S. Ampere-Oersted field splitting of the nonlinear spin-torque vortex oscillator dynamics. *Sci. Rep.* **12**, 10605 (2022).
49. De Wergifosse, S., Chopin, C. & Araujo, F. A. arXiv:2206.13438. Preprint at https://arxiv.org/abs/2206.13438.
50. Araujo, F. A., Chopin, C. & de Wergifosse, S. arXiv:2206.13596. Preprint at https://arxiv.org/abs/2206.13596.
51. Jenkins, A. S. *et al.* Spin-torque resonant expulsion of the vortex core for an efficient radiofrequency detection scheme. *Nat. Nanotechnol.* **11**, 360 (2016).
52. Pikovsky, A., Rosenblum, M. & Kurths, J. *Synchronization: A Universal Concept in Nonlinear Sciences* 1st edn. (Cambridge University Press, 2003).
53. Guslienko, K., Sukhostavets, O. V. & Berkov, D. V. Nonlinear magnetic vortex dynamics in a circular nanodot excited by spin-polarized current. *Nanoscale Res. Lett.* **9**, 386 (2014).

Acknowledgements

Author contributions

T.T. designed the project. Y.I. performed the numerical simulations based on a program code developed by T.T. Y.I. also developed a model analysis with help of T.T. Y.I. and T.T. wrote the manuscript and prepared the figures.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-42951-z.

Correspondence and requests for materials should be addressed to T.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.