



OPEN

## An STDP-based encoding method for associative and composite data

Hong-Gyu Yoon & Pilwon Kim

Spike-timing-dependent plasticity (STDP) is a biological process of synaptic modification caused by the difference of firing order and timing between neurons. One of neurodynamical roles of STDP is to form a macroscopic geometrical structure in the neuronal state space in response to a periodic input by Susman et al. (*Nat. Commun.* 10(1), 1–9 2019), Yoon, & Kim. Stdp-based associative memory formation and retrieval. [arXiv:2107.02429v2](https://arxiv.org/abs/2107.02429v2) (2021). In this work, we propose a practical memory model based on STDP which can store and retrieve high dimensional associative data. The model combines STDP dynamics with an encoding scheme for distributed representations and is able to handle multiple composite data in a continuous manner. In the auto-associative memory task where a group of images are continuously streamed to the model, the images are successfully retrieved from an oscillating neural state whenever a proper cue is given. In the second task that deals with semantic memories embedded from sentences, the results show that words can recall multiple sentences simultaneously or one exclusively, depending on their grammatical relations.

Spike-timing-dependent plasticity (STDP) is a biological process of synaptic modification according to the order of pre- and post-synaptic spiking within a critical time window<sup>3–5</sup>, and considered to be critical for understanding the cognitive mechanisms such as temporal coding<sup>6–8</sup>, and the formation of associative memory<sup>9,10</sup>. In our separate work<sup>2</sup>, we analyzed an STDP-based neural model and showed that the model can associate multiple high-dimensional memories to a geometric structure in the neural state space which we call a *memory plane*. When exposed to repeatedly occurring spatio-temporal input patterns, the neural activity based on STDP transforms the patterns into the corresponding memory plane. Further, the stored memories can be dynamically revived with macroscopic neural oscillations around the memory plane if perturbed by a similar stimulus.

The presence and the function of the memory plane in the neural networks have caught attention in Ref.<sup>1</sup>, where it has been proposed that STDP can store transient inputs as imaginary-coded memories. In this work, we further emphasize a practical aspect of the memory plane, showing that it can play a central role in storing, retrieving, and manipulating structured information. Using the theoretical works in Ref.<sup>2</sup>, we intend to integrate an analytic and an implementation level description of the neural memory process based on the memory plane that is capable of handling high dimensional associative data.

In this work, we propose that a STDP-based memory model, combined with a proper encoding scheme, can store and retrieve a group structured information in the neuronal state space. Among a number of schemes for encoding compositional structure that have been proposed over the last few years, we adopt Tensor Product Representation (TPR)<sup>11</sup>. TPR is a general method for generating vector-space embeddings of internal representations and operations, which prove to contain a variety of structural information such as lists of paired items, sequences and networks.

We show that the STDP-based memory model with TPR can naturally provide a mechanism for segmenting continuous streams of sensory input into representations of associative bindings of items: first, we demonstrate an auto-associative memory task with a group of images. While the images are sequentially streamed into the system for storage, the corresponding information is internally stored in the connectivity matrix. Then the whole group of the images can be dynamically retrieved from the oscillating neural state, when the system is perturbed by a memory cue which is similar to any of the original images. In the second task for semantic manipulation, we use multiple semantic vectors to represent a sentence as a composite of words. Once several sentences are stored in the system via such semantic vectors, a single word can recall multiple sentences simultaneously or one exclusively, depending on their grammatical relations. This implies that the proposed method provides an alternative bio-inspiring approach to process multiple groups of associative data with composite structure.

Department of Mathematical Sciences, Ulsan National Institute of Science and Technology (UNIST), Ulsan Metropolitan City 44919, Republic of Korea. email: [pwkim@unist.ac.kr](mailto:pwkim@unist.ac.kr)

## Methods

**STDP-based memory model.** Our work follows the framework of standard firing-rate models<sup>1,12</sup>. We set the differential equation for the neural state as

$$\dot{\mathbf{x}} = -\mathbf{x} + \mathbf{W}\phi(\mathbf{x}) + \mathbf{b}(t), \quad (1)$$

where  $\mathbf{x} = [x_1 \cdots x_N]^\top \in \mathbb{R}^N$  is the state of  $N$  neuronal nodes and  $\mathbf{W} = (W_{ij}) \in \mathbb{R}^{N \times N}$  is a connectivity matrix with  $W_{ij}$  corresponding to the strength of synaptic connection from node  $j$  to  $i$ . Here  $\phi$  is a regularizing transfer function and  $\mathbf{b}(t)$  is a memory input.

The plasticity model that we propose is based on<sup>13</sup> as

$$\begin{aligned} \dot{W}_{ij}(t) = & -\gamma W_{ij}(t) + \rho \left( \underbrace{\int_0^\infty K(s)\phi(x_j(t-s))\phi(x_i(t)) ds}_{\text{pre- to post- firing}} \right. \\ & \left. + \underbrace{\int_0^\infty K(-s)\phi(x_j(t))\phi(x_i(t-s)) ds}_{\text{post- to pre- firing}} \right), \end{aligned} \quad (2)$$

where  $K$  is a temporal kernel and  $\rho$  is the learning rate. The parameter  $\gamma$  is the decaying rate of homeostatic plasticity. While the synapses are repeatedly excited by the external input, they also need to decay simultaneously to forget and dump the obsolete and irrelevant information gradually. For analytic simplicity, we use  $\phi(\mathbf{x}) = \mathbf{x}$  and a simplified Dirac-delta typed temporal kernel  $K(s)$  defined as

$$K(s) := \begin{cases} \delta(s - \tau) & s > 0 \\ -\delta(s + \tau) & s \leq 0, \end{cases} \quad (3)$$

with  $\tau > 0$ . Using this kernel instead of the conventional exponential one has two main consequences: first, it models a specific interspike timing exhibiting maximal concentration of synaptic changes near origin, which has been observed in related works. Second, it yields a simple delayed term through the convolution in Eq. (2) which is relatively feasible for analysis<sup>2</sup>.

Now after simplifications, the main model becomes

$$\begin{cases} \dot{\mathbf{x}} = -\mathbf{x} + \mathbf{W}\mathbf{x} + \mathbf{b}(t) \\ \dot{\mathbf{W}} = -\gamma\mathbf{W} + \rho(\mathbf{x}\mathbf{x}_\tau^\top - \mathbf{x}_\tau\mathbf{x}^\top), \end{cases} \quad (4)$$

where  $\mathbf{x}_\tau = \mathbf{x}(t - \tau)$  stands for the delayed synaptic response.

We use the memory input in the form of a sequential harmonic pulse as

$$\mathbf{b}(t) = \sum_{i=1}^n \sin(\omega t - \xi_i) \mathbf{m}_i, \quad (5)$$

where  $\mathbf{m}_1, \dots, \mathbf{m}_n \in \mathbb{R}^N$  are memory representations to be stored. Here  $\omega$  stands for the frequency of neural oscillations and  $\xi_i, i = 1, \dots, n$  stands for the sampling time for each component. The trajectory of the memory input  $\mathbf{b}(t)$  in (5) is periodic and embedded in a 2-dimensional plane  $S \subset \mathbb{R}^N$  which we call a memory plane with respect to the memory representations  $\mathbf{m}_1, \dots, \mathbf{m}_n$ . While the memory representations are distributed in the high dimensional neural state space  $\mathbb{R}^N$ , the memory plane  $S$  tends to be located in close proximity to the memory representations under a suitable condition<sup>2</sup>. Further, the system (4) has an asymptotically stable solution  $(\mathbf{x}^*, \mathbf{W}^*)$  that consists of a periodic solution  $\mathbf{x}^*(t)$  on the memory plane  $S$  and a constant connectivity matrix  $\mathbf{W}(t) = \mathbf{W}^* = \alpha(\mathbf{v}\mathbf{u}^\top - \mathbf{u}\mathbf{v}^\top)$  for some vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $S$ . We will see below that the matrix  $\mathbf{W}^*$  contains the essential information to retrieve the memory representations  $\mathbf{m}_1, \dots, \mathbf{m}_n$ . This implies, a convergence of  $\mathbf{x}(t)$  to a certain periodic oscillation,  $\mathbf{x}^*(t)$ , is a sign that the memories are stored in  $\mathbf{W}$  in a distributed way.

In the retrieval phase, we set  $\gamma = \rho = 0$  in Eq. (4) and use the connectivity matrix  $\mathbf{W} = \mathbf{W}^*$  as

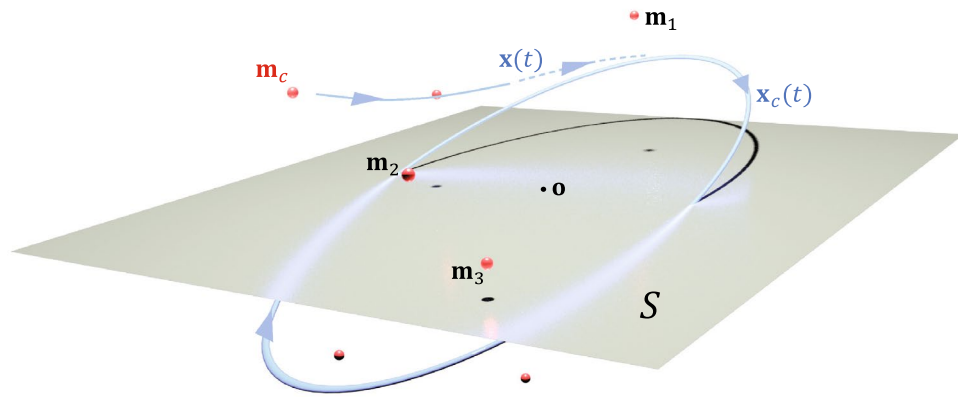
$$\dot{\mathbf{x}} = -\mathbf{x} + \mathbf{W}^*\mathbf{x} + \mathbf{b}(t). \quad (6)$$

where

$$\mathbf{b}(t) = \sin \omega t \mathbf{m}_c, \quad \mathbf{m}_c \in \mathbb{R}^N. \quad (7)$$

Here  $\mathbf{m}_c$  is the representation of memory cue. The retrieval system (6) has also an asymptotically stable periodic solution, say  $\mathbf{x}_c(t)$ . One can show that the trajectory of the periodic solution  $\mathbf{x}_c(t)$  is closely located to  $S$  if the memory cue  $\mathbf{m}_c$  is relevant to any of the memory representations  $\mathbf{m}_1, \dots, \mathbf{m}_n$ . Hence, as a neural state  $\mathbf{x}(t)$  is attracted to  $\mathbf{x}_c(t)$ , it is also bound to circle around the memory representations  $\mathbf{m}_1, \dots, \mathbf{m}_n$ . Figure 1 shows that the neural oscillation occurs near the memory plane and therefore in proximity to all the memory representations, when  $\mathbf{m}_c$  is given close to one of them.

**Encoding/decoding associative data with tag vectors.** Before storing and retrieving specific associative data in (4) and (6), respectively, we first need to properly encode them into memory representations  $\mathbf{m}_1, \dots, \mathbf{m}_n$ . It is reasonable to assume that the cognitive systems do not simply receive external inputs in a



**Figure 1.** Graphical illustration of the memory representations  $\mathbf{m}_1, \dots, \mathbf{m}_n$  and the corresponding memory plane  $S$ . The memory plane  $S$  is located in the subspace spanned by the memory representations and is shown to be close enough to them. A periodic orbit close to  $S$  can be used for efficient memory retrieval.

passive way, but rather actively pose them in the neural state space on acceptance. Suppose  $\mathbf{f}_1, \dots, \mathbf{f}_n \in \mathbb{R}^D$  are a series of the external inputs from the environment containing high dimensional associative information. We use a set of internal tags  $\mathbf{r}_1, \dots, \mathbf{r}_m \in \mathbb{R}^K$  to mark what classes the corresponding external inputs belong to. They may be used to indicate the order of sequence for the events (the first, the second, ..., the last) if the input is streamed through sequential observations, or types of sensors (visual, auditory, olfactory, tactile) if the input is a combination of senses, or the sentence elements (subject, predicate, object, modifier) if the input is a sentence composed of words. Such internal tags  $\mathbf{r}_1, \dots, \mathbf{r}_m$  can be formulated as low dimensional orthonormal vectors.

Following the vector embedding method for structured information<sup>11</sup>, we use the tensor product to encode a raw data  $\mathbf{f}_i$  into a memory representation  $\mathbf{m}_i$  as

$$\mathbf{m}_i = \mathbf{f}_i \otimes \mathbf{r}, \tag{8}$$

where  $\mathbf{r}$  is the tag vector corresponding to the raw data  $\mathbf{f}_i$ . If  $\mathbf{r}$  is of unit length, the original data can be exactly decoded from the memory representations by applying the right dot product of tagging vector as

$$\mathbf{m}_i \cdot \mathbf{r} = (\mathbf{f}_i \otimes \mathbf{r}) \cdot \mathbf{r} = \mathbf{f}_i(\mathbf{r} \cdot \mathbf{r}) = \mathbf{f}_i. \tag{9}$$

We say a neural state  $\mathbf{x} \in \mathbb{R}^N$  to be *retrievable* if  $\mathbf{x}$  is a linear combination of the memory representations  $\mathbf{m}_1, \dots, \mathbf{m}_n$ . For a retrievable state  $\mathbf{x}(t) = \sum_{j=1}^n c_j \mathbf{m}_j$ , a selective recovery of the original  $\mathbf{f}_i$  is possible by applying the right dot product with the corresponding tag vector as

$$\mathbf{x} \cdot \mathbf{r}_i = \left( \sum_{j=1}^n c_j \mathbf{m}_j \right) \cdot \mathbf{r}_i = c_i \mathbf{f}_i. \tag{10}$$

It can be shown that all the points on the memory plane  $S$  with respect to  $\mathbf{m}_1, \dots, \mathbf{m}_n$  are retrievable.

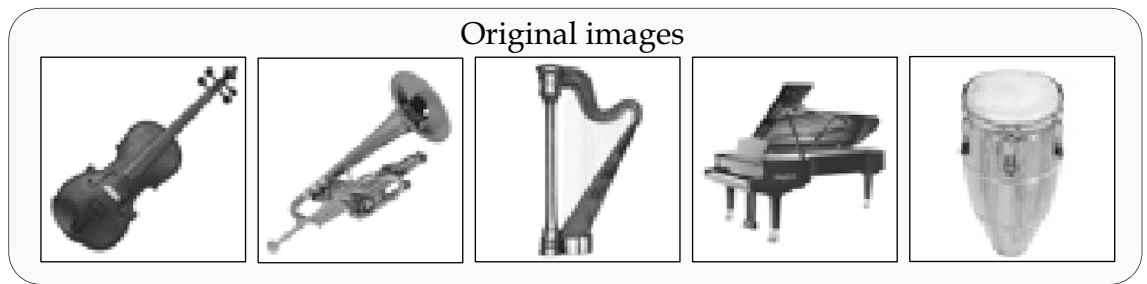
Let us describe how the encoding/decoding scheme is combined with the above memory models. In the storage phase, we encode the data  $\mathbf{f}_1, \dots, \mathbf{f}_n$  into the memory representations  $\mathbf{m}_1, \dots, \mathbf{m}_n$  using the tag vectors in Eq. (8), and run the system (4) with the memory input  $\mathbf{b}(t)$  in Eq. (5). To retrieve the original data, we run the retrieval system (6) with a certain cue  $\mathbf{m}_c$ . We wait until  $\mathbf{x}(t)$  converges to an oscillating trajectory,  $\mathbf{x}_c(t)$ , then try to evaluate  $\mathbf{x}(t) \cdot \mathbf{r}$  for retrieval. It has been shown<sup>2</sup> that  $\mathbf{x}_c(t)$  always has intersections with  $S$  which are therefore retrievable. However, since the neural state space  $\mathbb{R}^N$  is high dimensional, an irrelevant choice for  $\mathbf{m}_c$  leads to a trivial intersection near the origin which yields no meaningful retrieval. To the contrary, if  $\mathbf{m}_c$  is close to any of  $\mathbf{m}_1, \dots, \mathbf{m}_n$ , then  $\mathbf{x}(t)$  converges to a periodic solution  $\mathbf{x}_c(t)$  which is almost embedded  $S$  as illustrated in Fig. 1. Indeed, one can show that if  $\mathbf{m}_c$  is a scalar multiple of one of the memory representations, the corresponding  $\mathbf{x}_c(t)$  is completely embedded in  $S$  and is therefore retrievable for all  $t$ .

In the following section, we show through numerical tests for storage and retrieval that the STDP-based model (4) with the tensor product encoding can naturally provide a neural mechanism for segmenting continuous streams of sensory input into representations associative bindings of items.

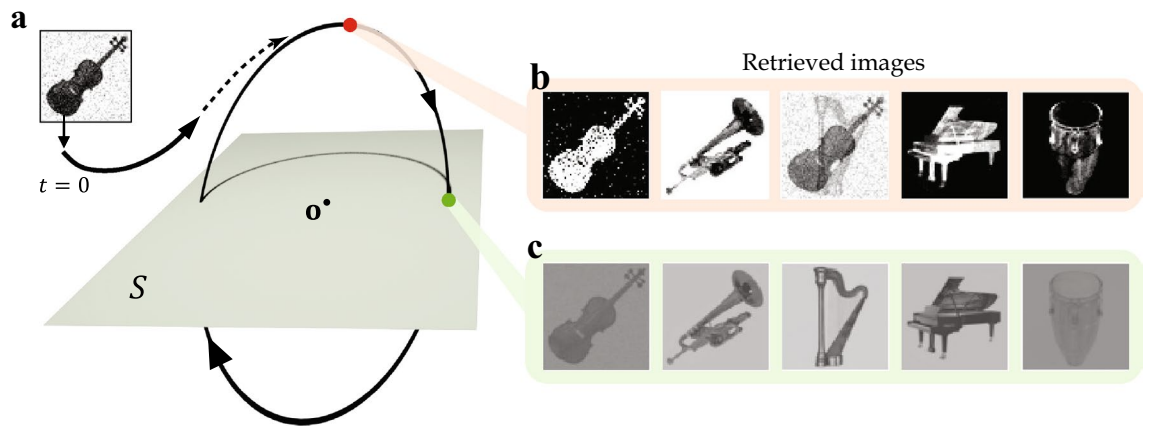
## Results

**Retrieval of grouped images.** We first demonstrate an auto-associative memory task that involves a group of images. This task uses five  $64 \times 64$  grayscale images of classical orchestral instruments in Fig. 2. The images are translated into external input vectors  $\mathbf{f}_i, i = 1, \dots, 5$  in  $\mathbb{R}^{64^2}$  and are combined into the memory representations as  $\mathbf{m}_i = \mathbf{f}_i \otimes \mathbf{r}_i, i = 1, \dots, 5$ . Here the tag vectors  $\mathbf{r}_i, i = 1, \dots, 5$  are orthonormal in  $\mathbb{R}^5$  and used as a placeholder for each image.

For numerical simulations in this article, the modified Euler’s method for delay equations has been universally used. In the first memory task, each  $64 \times 64$  image is translated to a vector  $\mathbf{f}_i$  as follows: every pixel is mapped to a value in  $[-\sigma, \sigma], \sigma > 0$ , depending on its brightness (pure black to  $-\sigma$  and pure white to  $\sigma$  linearly). Then



**Figure 2.** Grayscale images of  $64 \times 64$  pixels displaying classical orchestral instruments that are used for the memory input vectors  $\mathbf{f}_1, \dots, \mathbf{f}_5$ . All images were taken from Pixabay.

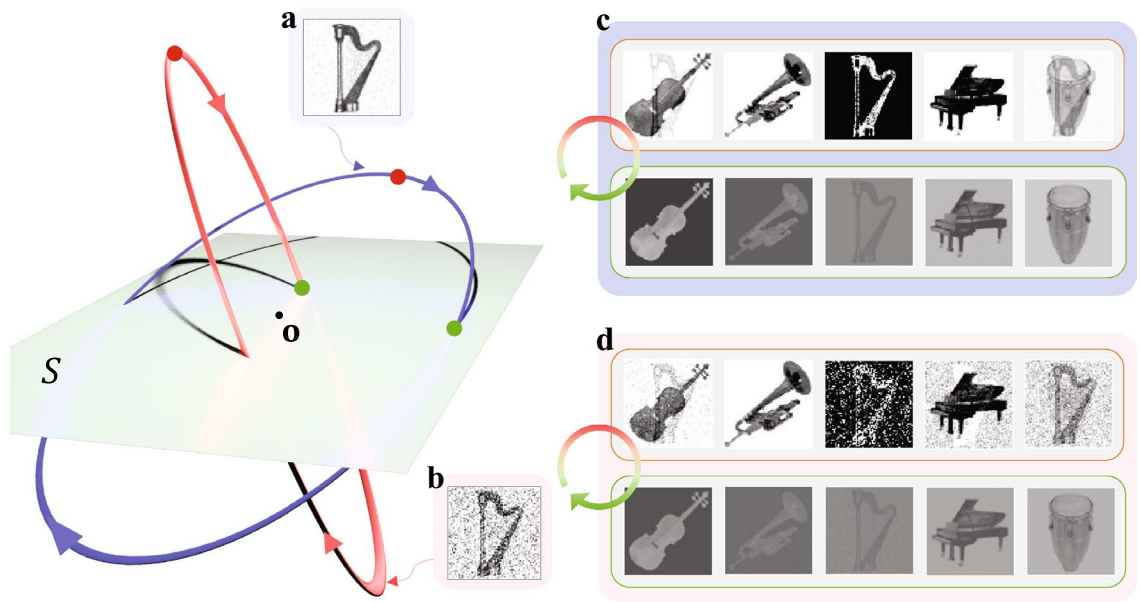


**Figure 3.** Auto-associative memory retrieval from a contaminated cue. (a) The noisy cue  $\mathbf{m}_c = \tilde{\mathbf{f}}_1 \otimes \tilde{\mathbf{r}}_1$  is generated from  $\tilde{\mathbf{f}}_1 = \sqrt{1 - \alpha^2} \mathbf{f}_1 + \alpha \boldsymbol{\zeta}$  and  $\tilde{\mathbf{r}}_1 = \sqrt{1 - \beta^2} \mathbf{r}_1 + \beta \boldsymbol{\eta}$ , where  $\boldsymbol{\zeta}, \boldsymbol{\eta}$  are Gaussian noise following  $\mathcal{N}_{\mathbb{R}^{64^2}}(0, \|\mathbf{f}_1\|)$  and  $\mathcal{N}_{\mathbb{R}^5}(0, 1)$ , respectively. The parameters are  $\alpha = 0.25$  and  $\beta = 0.2$ . (b) Snapshot of the retrieved images at the farthest point (red dot) from memory plane  $S$ . (c) Snapshot of the retrieved images at the intersection (green dot) of the orbit and the memory plane  $S$ . The timing of intersection  $t = t^\dagger > 0$  can be analytically determined as  $t^\dagger = (\tan^{-1} \omega + n\pi)/\omega, n \in \mathbb{Z}^2$ . Here, the contrast between the main object and the background are not well reproduced as the original images, since the brightness threshold  $\sigma = 0.005$  is used for drawing the retrieved images. One can obtain an original (either identical or exactly inverted) image by using a certain optimal value of  $\sigma$  for each retrieved image.

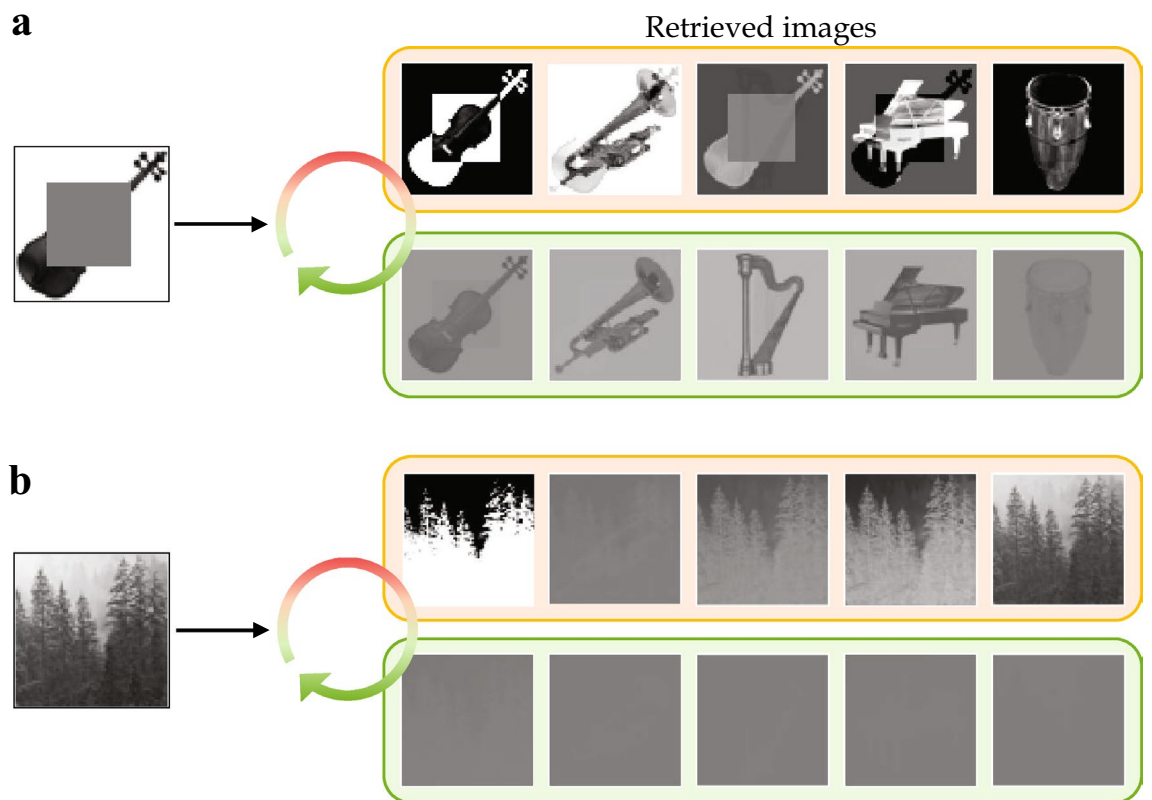
the resulted  $64 \times 64$  matrix is flattened to a vector  $\mathbf{f}_i$ . In the storage phase,  $\sigma = 0.02$  was used to maintain the magnitude for  $\mathbf{f}_i$  and  $\mathbf{m}_i$  at an appropriate level. Reconstructing the image from the vector can be done by performing the procedure in reverse order. The storage phase was proceeded for 40 seconds with the integration step size  $\Delta t = 0.1$  and  $\xi_i = \frac{\pi}{5}(i - 1)$  (5-evenly sequenced points on  $[0, \pi]$ ). The model parameters are chosen as  $\omega = 1.5, \gamma = \rho = 0.5$ , and  $\tau = \frac{\pi}{2\omega} = \frac{\pi}{3}$ . These specific choices were made based upon the analysis in [2], to achieve the convergence of  $\mathbf{W}(t) \rightarrow \mathbf{W}^*$  with simultaneously guaranteeing the sufficiently large magnitude of  $\mathbf{W}^*$ . As expected, stable convergence of connectivity to a certain  $\mathbf{W}^*$  is well achieved, when the arbitrary initial conditions  $\mathbf{x}(0)$  and  $\mathbf{W}(0)$  are appropriately small. The retrieval phase was proceeded for 15 seconds with  $\Delta t = 0.01$  for appropriately small arbitrary  $\mathbf{x}(0)$ . In Figs. 3, 4, and 5, the brightness threshold  $\sigma$  was adjusted to 0.002 for clear visibility, since the magnitude of the retrieved images are relatively small compared to original ones. Thus, any element of  $\mathbf{x} \cdot \mathbf{r}_i$  having value outside  $[-0.002 \ 0.002]$  is developed to a pixel of just pure black or white.

Figure 3 depicts the numerical simulation for the retrieval phase. For a better understanding of the process, a graphic illustration of the memory plane and the initial memory cue is given with the actual data. When the neural state  $\mathbf{x}(t)$  in Eq. (5) is continually perturbed by a noisy copy of one of the original images (violin), it approaches the memory plane  $S$ . Once the  $\mathbf{x}(t)$  converges to a limit cycle around  $S$  as stated in Theorem 2, the external input  $\mathbf{f}_1, \dots, \mathbf{f}_5$  can be reproduced by applying the tag vectors to  $\mathbf{x}(t)$ . In Fig. 3, we display two snapshots of the retrieved images obtained at two points on the orbit: Fig. 3b is taken at the farthest from  $S$  and Fig. 3c is at the intersection. It is notable that the retrieved images continuously oscillate, developing weak/strong and positive/negative images in turns. Such flashing patterns are generally different from image to image and are affected by the sequential order of the memory representations in Eq. (5) in the storage phase. Furthermore, due to the orthogonality of the tag vectors, the perfect images are acquired on the time instance when  $\mathbf{x}(t)$  penetrates  $S$ .

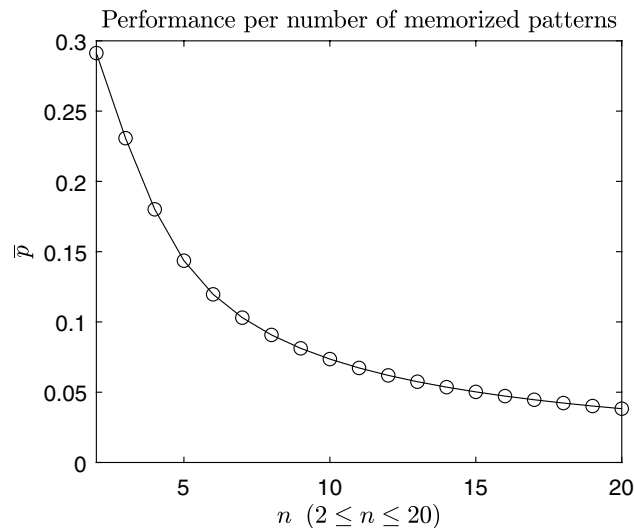
Figure 4 shows that the quality of the retrieved images depends on how close the memory cue is to the original image input. The cue with low-level noise in Fig. 4a leads to the orbit (blue) close to the memory plane  $S$ , producing the images of decent quality in Fig. 4c. However, if the cue is more contaminated with noise as in Fig. 4b,  $\mathbf{x}(t)$  approaches  $S$  at a relatively larger angle, making a stretched narrower elliptical orbit (red) that periodically



**Figure 4.** Comparison of retrieval quality according to the noise level in the cue. **(a)** The less noisy cue. The cue is generated in the same way as in Fig. 3 except using  $\mathbf{f}_3$  and  $\mathbf{r}_3$  instead of  $\mathbf{f}_1$  and  $\mathbf{r}_1$ . Gaussian noise are  $\mathcal{N}_{\mathbb{R}^{64^2}}(0, \|\mathbf{f}_3\|)$  and  $\mathcal{N}_{\mathbb{R}^5}(0, 1)$ , respectively. The parameters are  $\alpha = 0.1$  and  $\beta = 0.2$ . **(b)** The severely contaminated cue with  $\alpha = 0.7$ . **(c)** Snapshots of the retrieved images from the less noisy cue in (a), taken at the farthest point from  $S$  (top row) and at the intersection (bottom row). Here,  $p(t^\dagger) \approx 0.0271$ , and  $\bar{p} \approx 0.0899$ . **(d)** Snapshots of the retrieved images from the more noisy cue in (b), taken at the farthest point from  $S$  (top row) and at the intersection (bottom row). Here,  $p(t^\dagger) \approx 0.0194$ , and  $\bar{p} \approx 0.0688$ .



**Figure 5.** **(a)** Result of retrieval from a partially obstructed cue. Snapshots of the retrieved images are taken at the farthest point from  $S$  (top row) and at the intersection of the orbit and the memory plane (bottom row). Here,  $p(t^\dagger) \approx 0.0273$ , and  $\bar{p} \approx 0.1092$ . **(b)** Result of retrieval from an irrelevant cue. Snapshots of the retrieved images are taken at the farthest point from  $S$  (top row) and at the intersection of the orbit and the memory plane (bottom row). Here,  $p(t^\dagger) \approx 0.0000$ , and  $\bar{p} \approx 0.0015$ . In both (a) and (b), we used the noisy tag vector  $\mathbf{r}_1$  used in Fig. 3 for retrieval.



**Figure 6.** Memory capacity measured by the time-averaged performance: the memory representations were constructed with randomly generated 200 dimensional pattern vectors and 20 dimensional tag vectors.

gets far from  $S$ . Although the orbit from the severely contaminated cue still passes through the memory plane, it only does near the origin, providing relatively feeble images during a short time.

The retrieval can be performed with an incomplete cue. In Fig. 5a, the images are recalled from the partially obstructed cue. The original images can be recovered at a decent level, especially when  $\mathbf{x}(t)$  passes through the memory plane  $S$ . Figure 5b displays that an irrelevant cue (forest) fails to retrieve the original memory inputs. Indeed, it can be shown that a completely irrelevant cue results in a one-dimensional periodic orbit that keeps penetrating the memory plane back and forth just at the origin.

To quantitatively verify the performance, we introduce the scaled cosine similarity  $\mathbf{f}_i^\top \mathbf{g}_i(t) / \|\mathbf{f}_i\|^2$  between the original image  $\mathbf{f}_i$  and the retrieved image  $\mathbf{g}_i(t)$  at  $t$ . Note that, if  $\|\mathbf{g}_i(t)\| = O(\|\mathbf{f}_i\|)$ , this metric reflects the usual cosine similarity between two images and also ‘the memory intensity’  $\|\mathbf{g}_i(t)\| / \|\mathbf{f}_i\|$  together. Let  $p(t)$  denote the average of the scaled cosine similarity for the images  $\mathbf{f}_i, i = 1, \dots, n$  at  $t$ , and let  $\bar{p}$  denote the temporal average of  $p(t)$ . One can see that these two measures successfully capture the quality of retrieved images depending on given cues. See the caption of Figs. 4 and 5 for more detail.

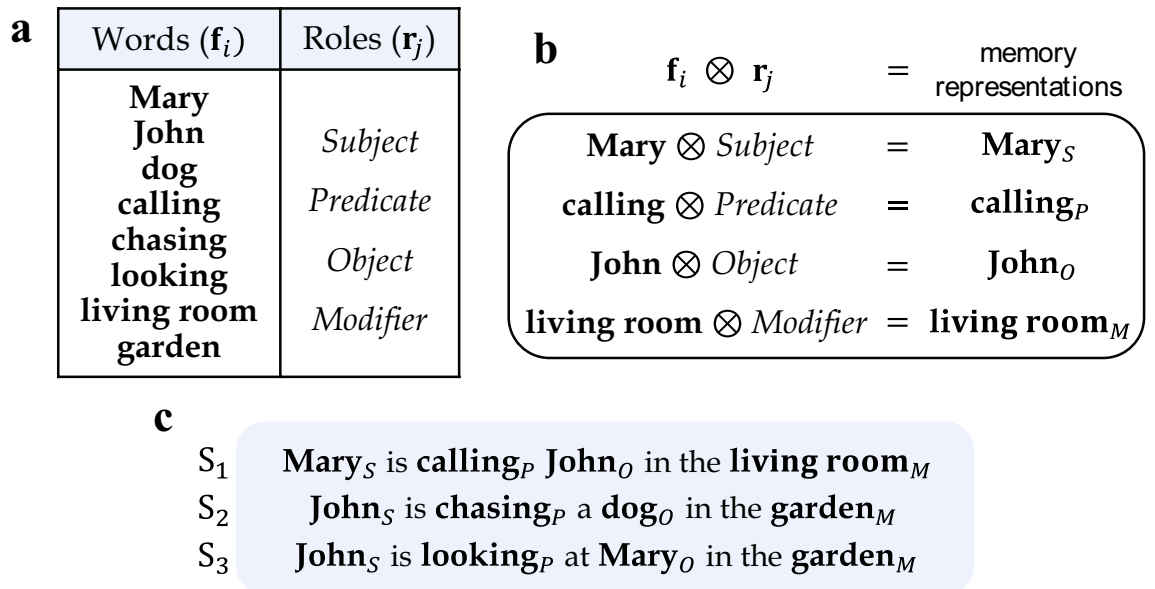
The memory capacity of the model measured by the time-averaged performance  $\bar{p}$  according to the number of memorized patterns is illustrated in Fig. 6. The performance  $\bar{p}$  is observed to decrease at least in a degree of  $n^{-\frac{1}{2}}$  as expected in Ref.<sup>2</sup>.

**Multiple groups of memory with composite structure.** This section deals with applications of the model to more complex associative memory. Suppose we have multiple groups of memory representations and have stored each group in the form of the memory plane using the system in Eq. (4). We are especially interested in the case where some memory representations belong to multiple groups. The following questions naturally arise: 1) Can the common memory component retrieve the corresponding multiple groups together? 2) Can a single memory group be selected by adding a further memory component in the cue? These questions are potentially related to high-level inference on memory.

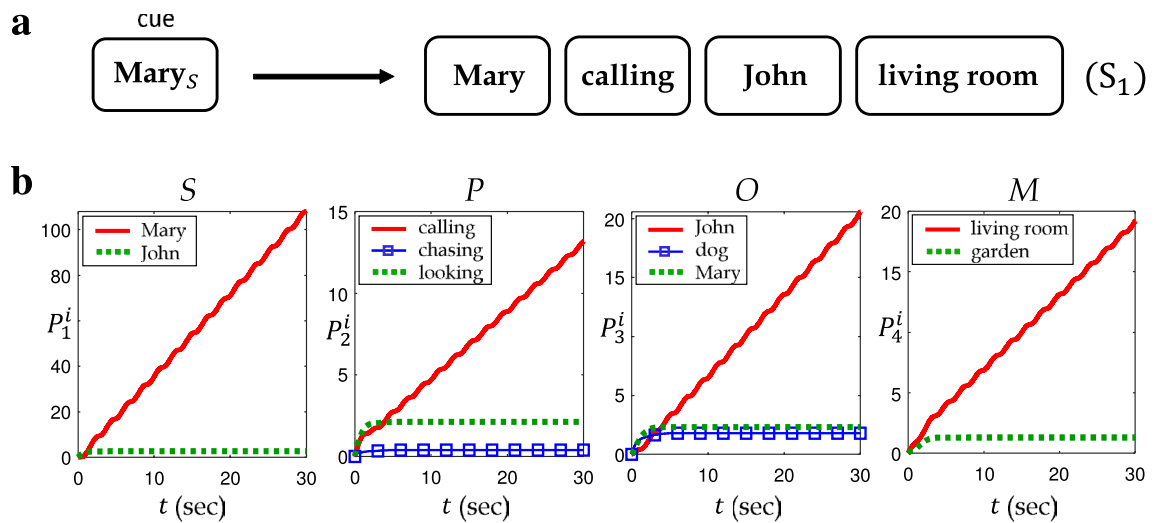
We also focus on compositional structure of memory representations created by the tag vectors. Memory inputs in this section are words and are collectively provided in the form of a sentence. We assume that each tag vector stands for the sentence element (subject, predicate, object, modifier) and is naturally bound to a word according to the role of the corresponding word in the sentence. Being activated by such a sequential stream of words, the system in Eq. (4) forms the memory plane which can be referred to as the encoding of the sentence.

For the simulation of semantic memory, we use three sentences composed of 8 words. Every word appearing in the sentence has one of the 4 roles (sentence elements). The vocabulary of the words and the roles are listed in Fig. 7a. We simply use arbitrarily chosen orthonormal sets  $\{\mathbf{f}_i\}_{i=1}^8$  for the words and  $\{\mathbf{r}_j\}_{j=1}^4$  for the roles, respectively. Figure 7b shows a couple of examples for memory representations each of which is a binding of a word and a role. Here the subindex on the right-hand side is used to express the corresponding role for the word. Our goal is to store the semantic information of sentences through Eq. (4) with the memory input  $\mathbf{b}(t)$  in Eq. (5). There are three sentences  $S_1, S_2$  and  $S_3$  listed in Fig. 7c that we use as the memory input in the simulation. Note that word **John** appears three times in the sentences, once in  $S_1$  as an object, and twice in  $S_2$  and  $S_3$  as a subject. Similarly, the words **Mary** and **garden** occur twice in a different context.

The memory connectivity  $\mathbf{W}_k^*, k = 1, 2, 3$  are obtained from separate single group learning on the sentences  $S_k, k = 1, 2, 3$ , respectively. We then set the combined memory connectivity for three sentences, i.e.,  $\mathbf{W}^* = \mathbf{W}_1^* + \mathbf{W}_2^* + \mathbf{W}_3^*$  for the collective retrieval phase. Here, we adopt the function



**Figure 7.** (a) List of words and roles used for the external input  $f_i$  and the tag  $r_j$ . (b) Descriptive explanation on constructing memory representations. (c) Three sentences  $S_1$ ,  $S_2$  and  $S_3$  generated by grouping memory representations. Note that the words **John**, **Mary** and **garden** appear several times in different contexts.



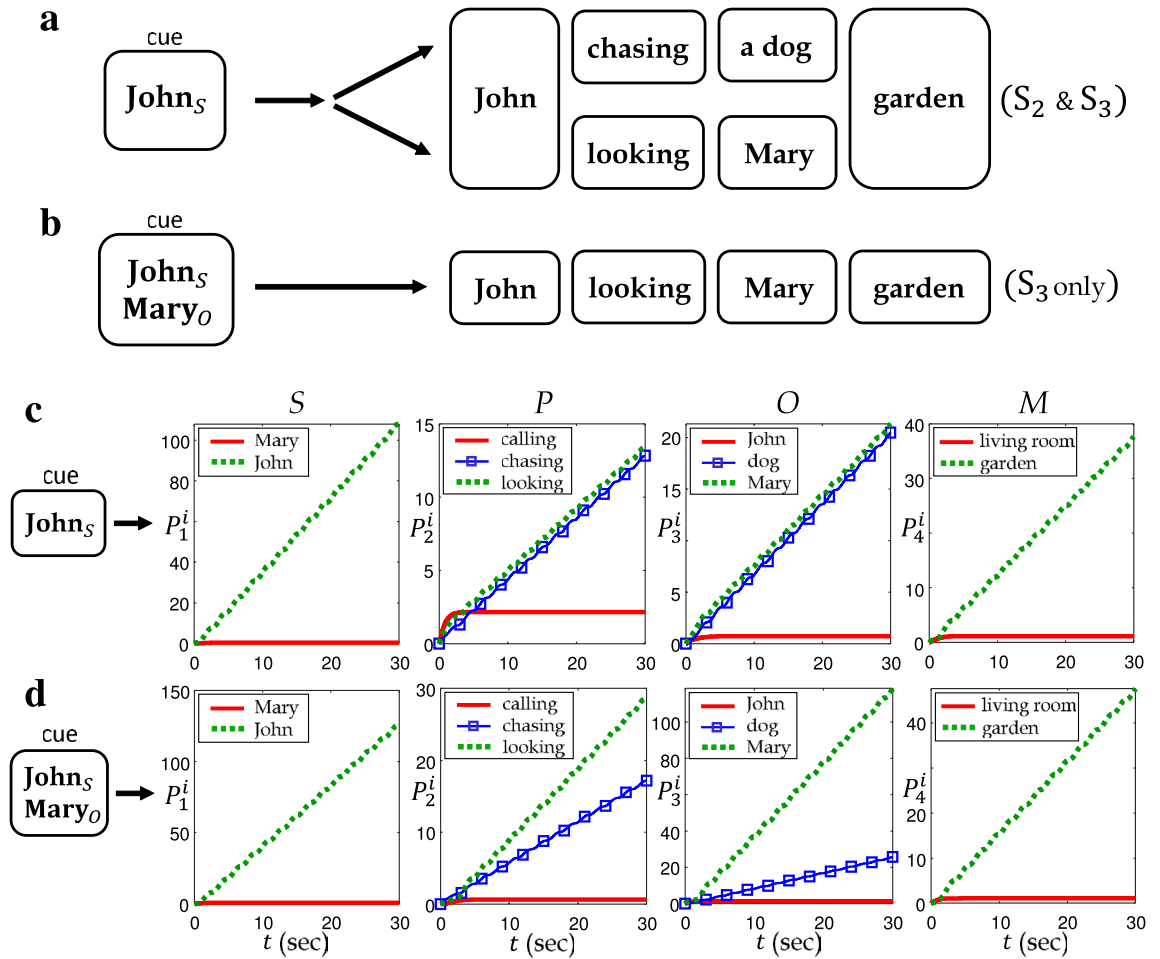
**Figure 8.** (a) Expected result of retrieval when the cue **Mary<sub>S</sub>** given. (b) The numerical result of retrieval by the cue **Mary<sub>S</sub>**. Dominant increasing values of  $P_j^i(t)$  are colored red, which turned out to correspond to  $S_1$ .

$$P_j^i(t) := \int_{t_0}^t \left| \mathbf{f}_i^\top(\mathbf{x}(s)) \cdot \mathbf{r}_j \right| ds, \quad i = 1, \dots, 8, j = 1, \dots, 4, \quad (11)$$

to measure how close the retrieved quantity is to the word  $f_i$  as the role  $r_j$ .

In the following semantics tasks, the retrieval was proceeded for 30 seconds with  $\Delta t = 0.01$  for appropriately small  $\mathbf{x}(0)$ . Multiple cues such as **John<sub>S</sub> + Mary<sub>O</sub>** in Fig. 9d are implemented by assigning each cue to its original sampling time through a harmonic pulse. In other words, the combined cue **John<sub>S</sub> + Mary<sub>O</sub>** is implemented as  $\mathbf{b}_c(t) = \sin(\omega t - \xi_1)(\mathbf{f}_2 \otimes \mathbf{r}_1) + \sin(\omega t - \xi_3)(\mathbf{f}_1 \otimes \mathbf{r}_3)$ .

In the first task of multiple composite memories, **Mary<sub>S</sub>** is given as the cue. Since **Mary** occurs as a subject only in  $S_1$ , one can expect the retrieved result to be  $S_1$  as in Fig. 8a. The numerical simulation of the retrieval process turned out to agree well with this expectation. Figure 8b compares the fitness of the words. The values of  $P_j^i(t)$  in Eq. (11) are evaluated while  $\mathbf{x}(t)$  is oscillating along a convergent orbit of Eq. (6). If  $P_j^i(t)$  keeps increasing with a large slope, the corresponding memory component  $f_i \otimes r_j$  can be identified as a dominantly retrieved one. The graphs in Fig. 8b show that such representations are **Mary<sub>S</sub>**, **calling<sub>P</sub>**, **John<sub>O</sub>** and **living room<sub>M</sub>**, which are well matched to  $S_1$ .



**Figure 9.** (a) Expected result of retrieval by a single component cue  $John_S$ . (b) Expected result of retrieval by multiple component cues  $John_S$  and  $Mary_O$ . (c) Numerical result of retrieval by a single component cue  $John_S$ . (d) Numerical result of retrieval by multiple component cues  $John_S$  and  $Mary_O$ .

The second task deals with the case with an ambiguous memory cue. Suppose that the memory component  $John_S$  is given as the cue. Since it occurs in the both sentence  $S_2$  and  $S_3$ , it is reasonable that the retrieval result should involve all the memory representations in both sentences as in Fig. 9a. This may be understood in that one of the fundamental capabilities of the brain is to examine all possible memories that contain the common cue, especially when the given cue is insufficient. However, this ambiguity can be eliminated by adding further cues. For example, if  $Mary_O$  is added as in Fig. 9b, the retrieval result should be narrowed down to  $S_3$  due to the extra constraint.

It turned out that the numerical simulations successfully capture the expected features of the memory retrieval process mentioned above. Figure 9c shows the results from the memory cue  $John_S$ . It is notable that  $chasing_p$  and  $looking_p$  in the second graph simultaneously increase with the almost same slope, indicating that they are equally dominant memory representations in retrieval. This is even clearer when compared to another memory component  $calling_p$  which is steady and negligible. The same pattern appears with  $John_O$  and  $dog_O$  in the third graph, both of which are dominant retrieval representations. The numerical results in Fig. 9d also reflect the retrieval tendency with additional memory cue. We provide the system in Eq. (6) with the extended memory input in Eq. (7) that consists of two memory representations  $John_S$  and  $Mary_O$ . Since the newly added cue  $Mary_O$  confines the retrieval result to the sentence  $S_3$  as in Fig. 9b, the memory representations in  $S_2$ ,  $chasing_p$  and  $dog_O$ , should be suppressed in retrieval. The second and third graphs in Fig. 9d show that, while  $chasing_p$  and  $dog_O$  increase (due to the common cue  $John_S$ ), the slope is smaller than that of  $looking_p$  and  $Mary_O$  in  $S_3$ , respectively. This implies that the dominantly retrieved representations are  $John_S$ ,  $looking_p$ ,  $Mary_O$  and  $garden_M$  which are matched to  $S_3$ .

### Discussion

There is now substantial evidence accumulated that neural oscillations are related to memory encoding, attention, and integration of visual patterns<sup>14–16</sup>. In Ref.<sup>1</sup>, the idea has been proposed that memories constitute stable dynamical trajectories on a two-dimensional plane in which an incoming stimulus is encoded as a pair of imaginary eigenvalues in the connectivity matrix. We extended such an idea further through a specific memory



system that can process a group of high dimensional associative data sets, by using the exact analytic relation between the inputs and the corresponding synaptic changes shown in Ref.<sup>2</sup>.

While the Hopfield network<sup>17</sup> retrieves static single data as a fixed point, the proposed model explores multiple data sets through neural oscillations. Compared to other neural-based models capable of conducting explicit and perfect retrieval of grouped data such as<sup>18–20</sup>, the model proposed in this work has the novelty of using a simple and continuous framework to reconstruct a group of multiple data in the general vector form. Moreover, the model produces the output through neural oscillations in reaction to an external cue, showing a potential link to a real memory process occurring in the brain.

We encode the input data with tag vectors based on the tensor representation, which has been proposed as a robust and flexible distributed memory representation<sup>11,21–24</sup>. This preprocess enables us to efficiently retrieve the stored data and, in addition, to deal with the composite structure in the data set. The ability to process associate multiple data sets with composite structures is essential in natural language understanding and reasoning. It has been shown that the proposed model can handle multiple sentences that describe distinct situations and can selectively allow the recall cue to arouse a group of associative memories according to its semantic relevance.

From a practical perspective, our results suggest an alternative approach for a memory device. The conventional von Neumann architecture is non-scalable and its performance is limited by the so-called von Neumann bottleneck between nonvolatile memories and microprocessors. On the other hand, operating data with artificial synapses is benefiting from a parallel information process consuming a small amount of energy per synapse. Moreover, conventional digital memory systems convert the inputs to a binary code and save it in a separate storage device, likely destroying the correlation information by such physical isolation. The proposed model is based on continuous dynamical systems and provides a simple and robust approach to deal with a sequence of associative high-dimensional data. Processing data in the continuous and distributed system results in the plastic storage of the correlated information in the synaptic connections.

In this work, we used a continuous model based on the average of local spiking activities of neurons. The model can be seen as a continuous approximation of Spiking Neural Network(SNN) which has been recognized as a promising architecture for bio-inspired neuromorphic hardware. There have been several studies showing the dynamical correspondence between SNNs and their firing-rate approximations including Wilson–Cowan model<sup>25–27</sup>. However, reproducing the memory process in this paper through SNN is substantial yet interesting challenge in practice, considering it requires establishing a proper conversion between a series of high dimensional data and spiking patterns.

Received: 6 August 2021; Accepted: 4 March 2022

Published online: 18 March 2022

## References

1. Susman, L., Brenner, N. & Barak, O. Stable memory with unstable synapses. *Nat. Commun.* **10**(1), 1–9 (2019).
2. H.-G. Yoon, & P. Kim. Sdp-based associative memory formation and retrieval. [arXiv:2107.02429v2](https://arxiv.org/abs/2107.02429v2) (2021).
3. Bliss, T. V. & Collingridge, G. L. A synaptic model of memory: Long-term potentiation in the hippocampus. *Nature* **361**(6407), 31–39 (1993).
4. Bi, G.-Q. & Poo, M.-M. Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **18**(24), 10464–10472 (1998).
5. Caporale, N. & Dan, Y. Spike timing-dependent plasticity: A hebbian learning rule. *Annu. Rev. Neurosci.* **31**, 25–46 (2008).
6. Blum, K. I. & Abbott, L. F. A model of spatial map formation in the hippocampus of the rat. *Neural Comput.* **8**(1), 85–93 (1996).
7. Rao, R. P. & Sejnowski, T. J. Spike-timing-dependent hebbian plasticity as temporal difference learning. *Neural Comput.* **13**(10), 2221–2237 (2001).
8. Gerstner, W., Kempter, R., Van Hemmen, J. L. & Wagner, H. A neuronal learning rule for sub-millisecond temporal coding. *Nature* **383**(6595), 76–78 (1996).
9. Tsodyks, M. Spike-timing-dependent synaptic plasticity—the long road towards understanding neuronal mechanisms of learning and memory. *Trends Neurosci.* **25**(12), 599–600 (2002).
10. Szatmáry, B. & Izhikevich, E. M. Spike-timing theory of working memory. *PLoS Comput. Biol.* **6**(8), e1000879 (2010).
11. Smolensky, P. Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artif. Intell.* **46**(1–2), 159–216 (1990).
12. Dayan, P. *et al.* Theoretical neuroscience: Computational and mathematical modeling of neural systems. *J. Cogn. Neurosci.* **15**(1), 154–155 (2003).
13. Kempter, R., Gerstner, W. & Van Hemmen, J. L. Hebbian learning and spiking neurons. *Phys. Rev. E* **59**(4), 4498 (1999).
14. Singer, W. & Gray, C. M. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* **18**(1), 555–586 (1995).
15. Gupta, N., Singh, S. S. & Stopfer, M. Oscillatory integration windows in neurons. *Nat. Commun.* **7**(1), 1–10 (2016).
16. Rutishauser, U., Ross, I. B., Mamelak, A. N. & Schuman, E. M. Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature* **464**(7290), 903–907 (2010).
17. Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.* **79**(8), 2554–2558 (1982).
18. C. Yáñez-Márquez, L. P. Sánchez-Fernández, & I. López-Yáñez. Alpha-beta associative memories for gray level patterns. in *International Symposium on Neural Networks*, pp. 818–823, Springer (2006).
19. C. Yáñez-Márquez, M. E. Cruz-Meza, F. A. Sánchez-Garfias, & I. López-Yáñez. Using alpha-beta associative memories to learn and recall rgb images. in *International Symposium on Neural Networks*, pp. 828–833, Springer (2007).
20. Yáñez-Márquez, C. *et al.* Theoretical foundations for the alpha-beta associative memories: 10 years of derived extensions, models, and applications. *Neural Process. Lett.* **48**(2), 811–847 (2018).
21. Hintzman, D. L. Minerva 2: A simulation model of human memory. *Behav. Res. Methods Instrum. Comput.* **16**(2), 96–101 (1984).
22. P. Kanerva, *Sparse distributed memory*. MIT press (1988).
23. Humphreys, M. S., Bain, J. D. & Pike, R. Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychol. Rev.* **96**(2), 208 (1989).
24. Pollack, J. B. Recursive distributed representations. *Artif. Intell.* **46**(1–2), 77–105 (1990).
25. Brette, R. Philosophy of the spike: Rate-based vs. spike-based theories of the brain. *Front. Syst. Neurosci.* **9**, 151 (2015).

26. Chow, C. C. & Karimipناه, Y. Before and beyond the wilson-cowan equations. *J. Neurophysiol.* **123**(5), 1645–1656 (2020).
27. Wallace, E., Benayoun, M., Van Drongelen, W. & Cowan, J. D. Emergent oscillations in networks of stochastic spiking neurons. *PLoS ONE* **6**(5), e14804 (2011).

### Acknowledgements

P. Kim was supported by National Research Foundation of Korea (2017R1D1A1B04032921). P. Kim was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2021R1A4A1032924).

### Author contributions

H.Y. developed the theoretical formalism, performed the analytic calculations and performed the numerical simulations. P.K. devised the main conceptual ideas and wrote the main manuscript text.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to P.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022