

## Regular Paper

# Properties of Associative Memory Model with the $\beta$ -th-order Synaptic Decay

RYOTA MIYATA<sup>1,2</sup> TORU AONISHI<sup>1</sup> JUN TSUZURUGI<sup>3</sup> KOJI KURATA<sup>4,a)</sup>

Received: January 30, 2013, Revised: March 20, 2013/June 18, 2013,  
Accepted: July 25, 2013

**Abstract:** Many associative memory models with synaptic decay such as the forgetting model and the zero-order decay model have been proposed and studied so far. The previous studies showed the relation between the storage capacity  $C$  and the synaptic decay coefficient  $\alpha$  in each synaptic decay model. However, with the exceptions of a few studies, they did not compare the network retrieval performance between different synaptic decay models. We formulate the associative memory model with the  $\beta$ -th-order synaptic decay as an extension of the zero-order decay model. The parameter  $\beta$  denotes the synaptic decay order or the degree of the synaptic decay term, which enables us to compare the retrieval performance between different synaptic decay models. Using numerical simulations, we investigate the relation between the synaptic decay coefficient  $\alpha$  and the storage capacity  $C$  of the network by varying the synaptic decay order  $\beta$ . The results show that the properties of the synaptic decay model are constant for a large decay order  $\beta$ . Moreover, we search the minimum  $\beta$  to avoid overloading and the optimal  $\beta$  to maximize the network retrieval performance. The minimum integer value of  $\beta$  to avoid overloading is  $-1$ . The optimal integer value of  $\beta$  to maximize the network retrieval performance is 1, i.e., the degree of the forgetting model, and the suboptimal integer  $\beta$  is 0, i.e., that of the zero-order synaptic decay model.

**Keywords:** auto-associative memory,  $\beta$ -th order synaptic decay, zero-order synaptic decay, forgetting process, synaptogenesis, storage capacity

## 1. Introduction

The task of storing information as a memory and recalling it from the partial or noisy cue is fundamental for the brain. The Hopfield model [4], the best-known associative memory model, could reproduce the form of memory in the brain. By changing the synaptic weights, this model stores memory patterns in the network. Thereafter, this network retrieves the stored pattern even if given its noisy one. The properties of associative memory model have been investigated over several decades (for reviews see Ref. [14]).

The ordinary Hopfield model, however, has a critical number of memory patterns that can be stably stored:  $0.138N$ , where  $N$  is the system size [1], [2], [13], [16]. If the number of embedded memory patterns is larger than  $0.138N$ , the network cannot retrieve any memory pattern due to overloading.

It has been reported that elimination and generation of synapses (synaptogenesis [17], [18]) continues to take place in certain regions of the postnatal brain including the hippocampal regions, which are widely considered to function as associative

memory. Synaptogenesis, i.e., replacement of old synapses by new ones, is crucial for the formation of neural networks, and thus it affects neural network functions especially memory formation. Furthermore, according to the previous neurophysiological experiment [18], synapses with smaller weights tend to be eliminated with higher probability.

In our previous work, we investigated the effect of synaptogenesis on memories, modeling it mathematically: the associative memory model with the zero-order synaptic decay [9]. We demonstrated the possibility that synaptogenesis plays a role in maintaining recent memories while avoiding overloading. In the zero-order synaptic decay model, all of the synaptic weights approach zero by increments at each learning step for embedding a new pattern. This synaptic decay process is consistent with the characteristics of the synaptogenesis: the smaller weight the synapse has, more easily elimination of the synapse occurs, and it is equivalent to erasing the old memory traces gradually. This enables the network to retrieve recently embedded memory patterns stably while avoiding overloading.

If the dynamics of synaptic weights in the network have decay [3], [6], [7], [9], [11] or saturation [15], overloading does not occur. The associative memory model with the forgetting process, an alternative network model to avoid overloading, was proposed by Mézard et al. [7]. In the forgetting model, every time the network learns a new pattern, the synaptic weight decreases at a rate proportional to its value. In this scheme, the old memory traces are erased exponentially with time. Many synaptic decay models such as the forgetting model have been proposed and studied

<sup>1</sup> Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Yokohama, Kanagawa 226–8502, Japan

<sup>2</sup> Research Fellow of the Japan Society for the Promotion of Science, Chiyoda, Tokyo 102–0083, Japan

<sup>3</sup> Faculty of Engineering, Okayama University of Science, Okayama 700–0005, Japan

<sup>4</sup> Faculty of Engineering, University of the Ryukyus, Nakagami, Okinawa 903–0213, Japan

<sup>a)</sup> kurata@mibai.tec.u-ryukyu.ac.jp

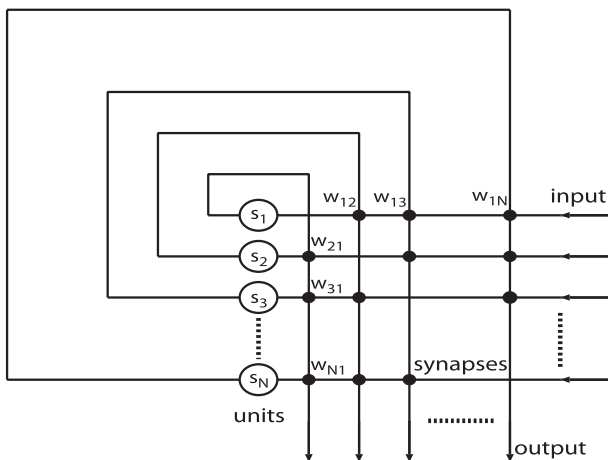
so far [3], [6], [7], [9], [11]. Some researchers gave a statistical mechanical analysis of each synaptic decay model using the spin glass analogy [5], [7], [8], [10], [12]. They showed the relationship between the storage capacity  $C$  and the synaptic decay coefficient  $\alpha$  in each synaptic decay model.

From an engineering viewpoint, these models raise a key question what is the optimal synaptic decay model to maximize the network retrieval performance. With the exceptions of a few studies [10], hitherto the previous studies did not compare the network retrieval performance between different synaptic decay models. We explore the optimal synaptic decay model, formulating the associative memory model with the  $\beta$ -th-order synaptic decay as an extension of the zero-order decay model. The parameter  $\beta$  denotes the synaptic decay order, or the degree of the synaptic decay term, which enables us to compare the retrieval performance between different synaptic decay models. When the decay order  $\beta$  is 1, this synaptic decay model is equivalent to the forgetting model. Using numerical simulations, we investigate the relation between the synaptic decay coefficient  $\alpha$  and the storage capacity  $C$  of the network by varying the synaptic decay order  $\beta$ . The results show that the properties of the model are constant for a large decay order  $\beta$ .

## 2. Model

We begin by formulating the  $\beta$ -th-order synaptic decay process, which includes synaptic reset (i.e., elimination and generation of synapses) in the associative memory model. As shown in **Fig. 1**, we consider a recurrent network model with  $N$  formal neurons and  $N(N-1)$  synaptic connections.

The associative memory model consists of the memory storage process and memory retrieval process. The synaptic decay and synaptic reset are included in the storage process. For the purpose of mathematical tractability and simplicity, we assume that the time-scale of synapse dynamics in the storage process is far different from that of network dynamics in the retrieval process. Under such an assumption of time-scale separation, the storage process and retrieval process can be separated from one another.



**Fig. 1** Schematic diagram of a recurrent network functioning as an auto-associative memory. Neurons are numbered  $i = 1, 2, \dots, N$  and are characterized their respective activities,  $s_1, s_2, \dots, s_N$ , which take  $\pm 1$ , respectively. Presynaptic neuron  $j$  is connected to postsynaptic neuron  $i$  through a synapse with efficacy (weight)  $w_{ij}$ .

Set out below is the outline of auto-associative memory functions that are implemented in our model. In the memory storage process,  $M$  patterns  $\xi^1, \xi^2, \dots, \xi^M$  are embedded in the network one by one, modifying synaptic weights in accordance with the Hebbian learning rule. Every time the network learns a new pattern, for example, all the synaptic weights decay at a constant speed (zero-order synaptic decay [9]) or decay exponentially (forgetting process [7]). In the zero-order synaptic decay model, the synapse whose weight is sufficiently small disappears, and then the new synaptic connection is created. Note that the synaptic weights are fixed during the memory retrieval process. The network with the fixed synaptic weights retrieves the recent memory patterns even if given the noisy ones.

### 2.1 Synapse Dynamics in the Storage Process

The index  $\mu (= 1, 2, \dots, M)$  of memory pattern  $\xi^\mu$  is the learning step, which represents the time when the pattern was stored. Note that  $\mu = M$  represents the most recently stored pattern, and that the smaller  $\mu$  is, the older  $\xi^\mu$  is. Each element  $\xi_i^\mu$  ( $i = 1, 2, \dots, N$ ) of the  $\mu$ -th memory pattern  $\xi^\mu$  takes  $\pm 1$ , and is generated independently with the probability,

$$\text{Prob}[\xi_i^\mu = \pm 1] = \frac{1}{2}. \quad (1)$$

We now consider the following synaptic dynamics with the synaptic decay:

$$w_{ij}(\mu) = w_{ij}(\mu - 1) + \Delta w_{ij}, \quad (2)$$

$$\Delta w_{ij} = -\alpha \text{sgn}(w_{ij}(\mu - 1)) |w_{ij}(\mu - 1)|^\beta + \xi_i^\mu \xi_j^\mu, \quad (3)$$

where  $\text{sgn}(x)$  is the sign function

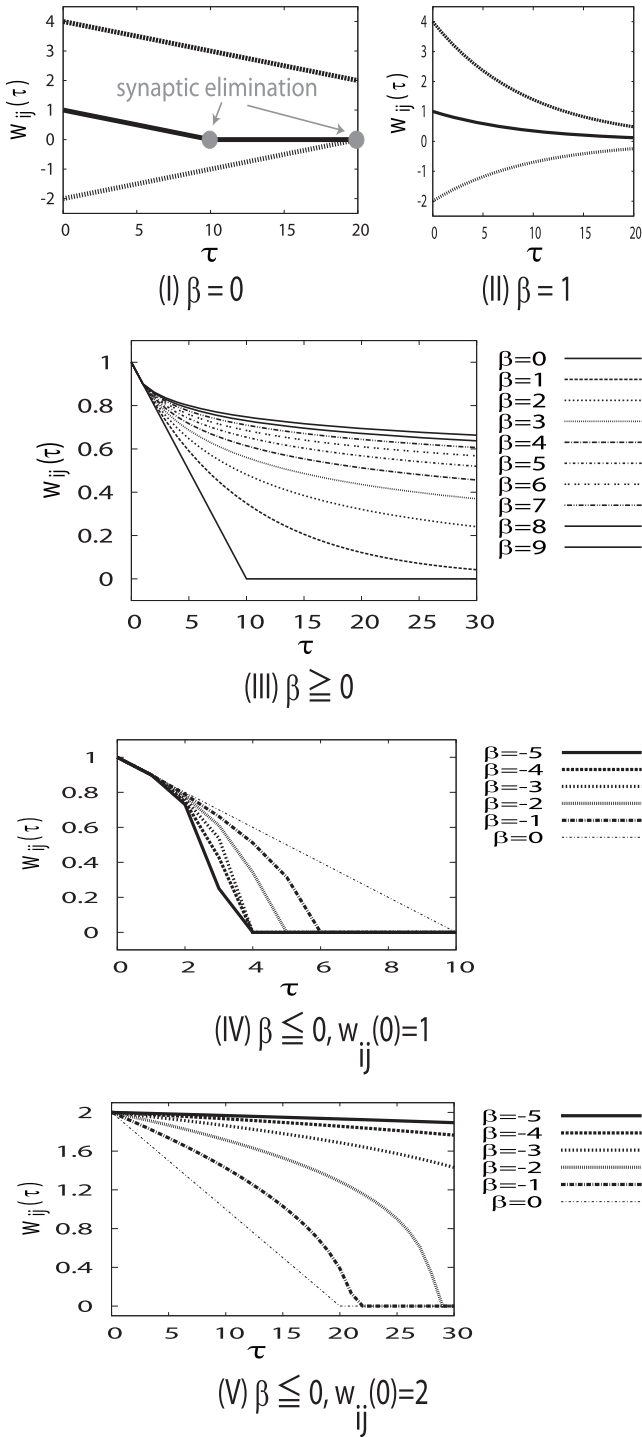
$$\text{sgn}(x) = \begin{cases} 1, & x \geq 0, \\ -1, & \text{otherwise.} \end{cases} \quad (4)$$

$w_{ij}$  denotes the weight of synaptic connection from unit  $j$  to  $i$ , and it is symmetrical, i.e.,  $w_{ij} = w_{ji}$ . The network is assumed to have no self-interaction,  $w_{ii} = 0$ .  $\Delta w_{ij}$  denotes the amount of synaptic modification at each learning step.

The second term of the right-side Eq. (3) represents the correlation (generalized Hebbian) learning rule. This rule is local in that the change to synaptic weight between units  $i$  and  $j$  for embedding a new pattern depends only on the activities of these two units and not on those of other units.

On the other hand, the first term of the right-side Eq. (3) represents the synaptic decay process. The parameter  $\beta$  denotes the synaptic decay order, or the degree of the synaptic decay term. When  $\beta = 0$ , the parameter  $\alpha$  denotes the synaptic decay speed. Every time the network learns a new pattern according to the Hebbian rule,  $|w_{ij}|$  decreases by a constant value  $\alpha$ . This is called the zero-order synaptic decay process [9]. **Figure 2** (I) illustrates the zero-order synaptic decay dynamics:  $\Delta w_{ij} = -\alpha \text{sgn}(w_{ij}(\mu - 1))$ . In the zero-order decay model, all the synapses decay at a constant speed  $\alpha$ . The smaller  $|w_{ij}|$  is, the earlier it reaches zero. If the sign of the synaptic weight inverts as a result of the synaptic decay process, the synaptic weight is reset ( $w_{ij} = 0$ ) and then a new pattern is stored:

$$w_{ij}(\mu) = \xi_i^\mu \xi_j^\mu, \quad \text{if } |w_{ij}(\mu - 1)| < \alpha |w_{ij}(\mu - 1)|^\beta. \quad (5)$$



**Fig. 2** Synaptic decay dynamics (i.e., the first term of the right-side Eq. (3)). (I) The zero-order decay model [9],  $\Delta w_{ij} = -\alpha \text{sgn}(w_{ij}(\mu - 1))$ ; (II) the forgetting model [7],  $\Delta w_{ij} = -\alpha w_{ij}(\mu - 1)$ .  $\alpha = 0.1$ . Abscissa represents the time step  $\tau$  and ordinate does the synaptic weight  $w_{ij}(\tau)$ . The two panels plot examples of synaptic decay dynamics with three different initial synaptic weights:  $w_{ij}(0) = 4, 1, -2$ . The synaptic weight of the zero-order synaptic decay model reaches zero in finite time whereas that of the forgetting model does not. When  $w_{ij}(\tau)$  becomes zero, elimination of the synapse occurs. (III) The  $\beta$ -th-order synaptic decay dynamics when  $\beta \geq 0$  and  $w_{ij}(0) = 1$ :  $\Delta w_{ij} = -\alpha \text{sgn}(w_{ij}(\mu - 1)) |w_{ij}(\mu - 1)|^\beta$ . The larger  $\beta$  is, the more slowly  $|w_{ij}(\tau)|$  decays. The synaptic decay dynamics when  $w_{ij}(0) > 1$  and  $\beta \geq 0$  have the same characteristics as the case in which  $w_{ij}(0) = 1$  and  $\beta \geq 0$ . (IV)  $w_{ij}(0) = 1, \beta \leq 0$ . The larger  $|\beta|$  is, the faster  $|w_{ij}(\tau)|$  converges to zero. (V)  $w_{ij}(0) = 2, \beta \leq 0$ . As  $|\beta|$  increases,  $|w_{ij}(\tau)|$  hardly decays at all.

This procedure simply represents elimination of the sufficiently small synapse and birth of a new synaptic connection. We here assume that if the elimination of synaptic connection between units  $j$  and  $i$  occurs, the new synapse rebuilds the connection  $w_{ij}$ .

When  $\beta = 1$  in Eq. (3), the parameter  $\alpha$  denotes the synaptic decay rate. Every time the network learns a new pattern,  $|w_{ij}|$  decays exponentially. This is called the forgetting process [7] or the first-order synaptic decay process. Figure 2 (II) illustrates the dynamics of forgetting process:  $\Delta w_{ij} = -\alpha w_{ij}(\mu - 1)$ . In the forgetting model, the synaptic weight decreases at a rate proportional to itself. Thus, the synaptic weight does not converge to zero in finite time. Unlike the zero-order synaptic decay model, the elimination and generation of synapses do not occur in the forgetting model.

Next, we consider the higher order of synaptic decay ( $\beta > 1$ ). Figure 2 (III) illustrates the  $\beta$ -th-order synaptic decay dynamics  $\Delta w_{ij} = -\alpha \text{sgn}(w_{ij}(\mu - 1)) |w_{ij}(\mu - 1)|^\beta$  when  $\beta$  is more than or equal to 0. When  $\beta \geq 1$ , the synaptic weight does not converge to zero in finite time as shown in this figure. Thus, the elimination and generation of synapses do not occur in the model with the higher synaptic decay order than 0. Furthermore, the larger  $\beta$  is, the more slowly the synaptic weight decays. Thus, the higher-order decay model ( $\beta > 1$ ) tends to store the older memory traces than the zero-order decay model and the forgetting model.

We also consider the case when the decay order is negative ( $\beta < 0$ ). Figures 2(IV)-(V) illustrate the  $\beta$ -th-order synaptic decay dynamics  $\Delta w_{ij} = -\alpha \text{sgn}(w_{ij}(\mu - 1)) |w_{ij}(\mu - 1)|^\beta$  when  $\beta$  is less than or equal to 0. When  $\beta \leq 0$ , the synaptic weight converge to zero in finite time. Thus, the elimination and generation of synapses also occurs in the model with the lower synaptic decay order than 0. The synaptic reset procedure described by Eq. (5) is applied to the network model with the negative decay order. Moreover, as shown in Fig. 2 (IV), the larger  $|\beta|$  is, the faster the synaptic weights converge to zero when  $w_{ij} \leq 1$ . However, the synaptic decay dynamics with the negative decay order ( $\beta < 0$ ) is more complicated than those with the positive one ( $\beta > 0$ ). As shown in Fig. 2 (V), as the absolute value  $|\beta|$  increases, the synaptic weights hardly decay at all.

In the next section, we investigate the effect of synaptic decay order on the properties of the associative memory model using numerical simulations. In the storage process, all of the initial synaptic weights are set to zero  $w_{ij}(0) = 0, (i, j = 1, 2, \dots, N)$ , and Eqs. (2), (3) are repeated  $M$  times in order to determine the synaptic weights.

### 2.2 Network Dynamics in the Retrieval Process

In the retrieval process, the network restore pattern vectors to the original stored state using the synaptic weights after learning,  $w_{ij}(M)$ . Note that the synaptic weights are fixed during the retrieval process ( $w_{ij} = w_{ij}(M)$ ).

We assume that all the units work synchronously at discrete time  $t (= 1, 2, \dots)$ . The network dynamics can be determined by

$$s_i(t + 1) = \text{sgn}(h_i(t)), \tag{6}$$

$$h_i(t) = \sum_{j=1(\neq i)}^N w_{ij} s_j(t), \tag{7}$$

where  $s_i(t)$  is the state of unit  $i$  at time  $t$ .  $h_i(t)$  is the local field of unit  $i$  at time  $t$ , whose sign determines whether the next state of unit  $i$  is firing ( $s_i(t + 1) = 1$ ) or not ( $s_i(t + 1) = -1$ ).

Equation (8) defines the time evolution of the system state. For any symmetric connection matrix  $w_{ij}$ :  $w_{ij} = w_{ji}$ , the network system has finite possible states. Starting from any arbitrary initial state, the system state of the ordinary associative memory model always reaches an equilibrium or a periodic solution, and the period is known to be no more than 2. According to Eqs. (8), (9), the system state is updated until it converges.

### 3. Simulation Results

For simplicity, we assumed that neither the total number of units  $N$  nor that of synapses  $N(N - 1)$  does change over time in the following numerical simulations. The network consisting of 1,000 units learned  $M$  patterns,  $\xi^1, \xi^2, \dots, \xi^M$ , one by one. The number of patterns to be stored was set to 1,000 (i.e.,  $M = N$ ), which was much larger than the critical value of the ordinary Hopfield model:  $0.138N$ . In the storage process, all of the initial synaptic weights were set to zero,  $\{w_{ij}(0)\} = 0$ . Every time the network learned a memory pattern, the absolute value of synaptic weight  $|w_{ij}|$  decays in proportional to the exponentiation  $|w_{ij}|^\beta$ . Note that the synaptic weights are then fixed during the retrieval process. In the retrieval process, each stored pattern  $\xi^\mu$  ( $\mu = 1, 2, \dots, M$ ) was given as an initial state  $s(0)$ . This is because it has been reported that the storage capacity of the associative memory network when a stored pattern corrupted with a small noise was given as  $s(0)$  is almost the same as that when the originally stored pattern was given [1], [12]. Considering the periodical solution of the network dynamics, the system is forced to stop when the present state  $s(t)$  equals the second to last one  $s(t - 2)$ . Simulations were carried out on a computer by varying the decay coefficient  $\alpha$ .

First, we investigate how many and how old patterns are retrievable in the network with the  $\beta$ -th-order synaptic decay. Because the synaptic decay and synaptic elimination processes implies erasing old memory gradually, a limit exists on the number of retrievable patterns. As the criterion of successful recall, we used the overlap between the  $\mu$ -th memory pattern  $\xi^\mu$  and the stationary system state  $s^\mu$  when the initial system state is set to  $s(0) = \xi^\mu$ ,

$$m^\mu = \frac{1}{N} \sum_{i=1}^N \xi_i^\mu s_i^\mu. \tag{8}$$

In calculation of the overlap, we used  $s(t)$  which first fulfilled the stop condition of retrieval dynamics ( $s(t) = s(t - 2)$ ) as a stationary state  $s^\mu$ . If an initial state  $s(0)$  is equal to or close to an originally stored pattern  $\xi^\mu$  and the stored pattern has not been erased in the network with synaptic decay, the memory retrieval state is stable:  $m^\mu \approx 1$ . Otherwise  $m^\mu \approx 0$ . If  $m^\mu \geq 0.8$ , we regarded  $\xi^\mu$  as the retrievable pattern, and we counted it. The storage capacity  $C$  is defined as the total number of the retrievable patterns.

Figure 3 plots the overlap  $m^\mu$  as a function of the learning step  $\mu$  for several values of the decay coefficients  $\alpha$  in the four  $\beta$ -th-order synaptic decay models other than the zero-order decay model and the forgetting model ( $\beta \neq 0, 1$ ). As shown in the fig-

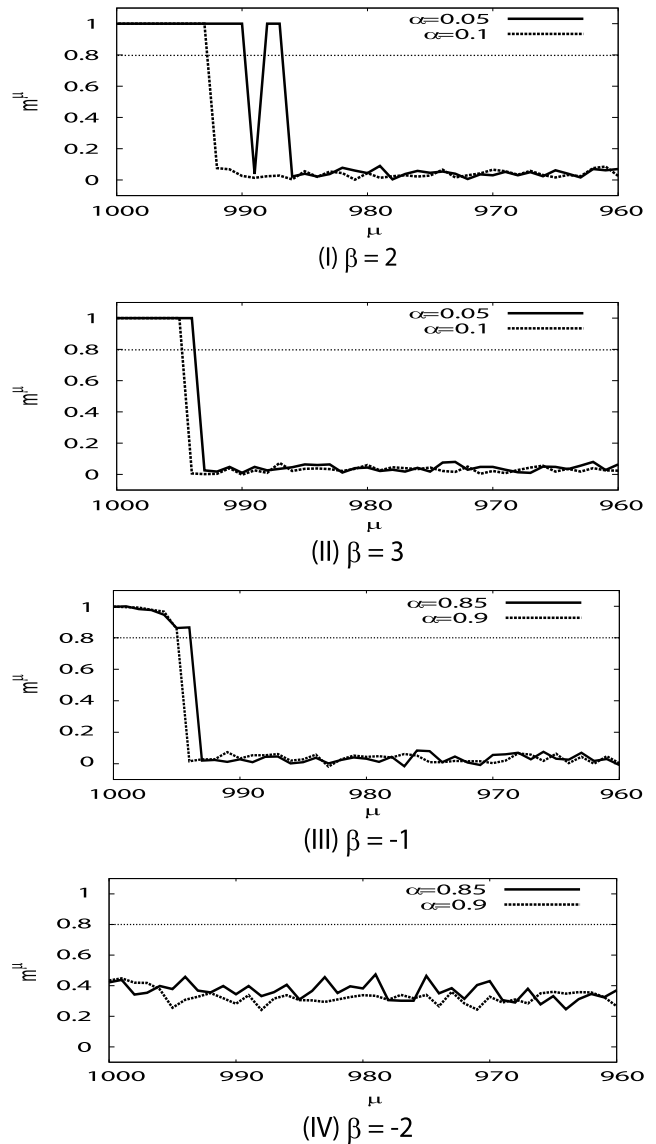
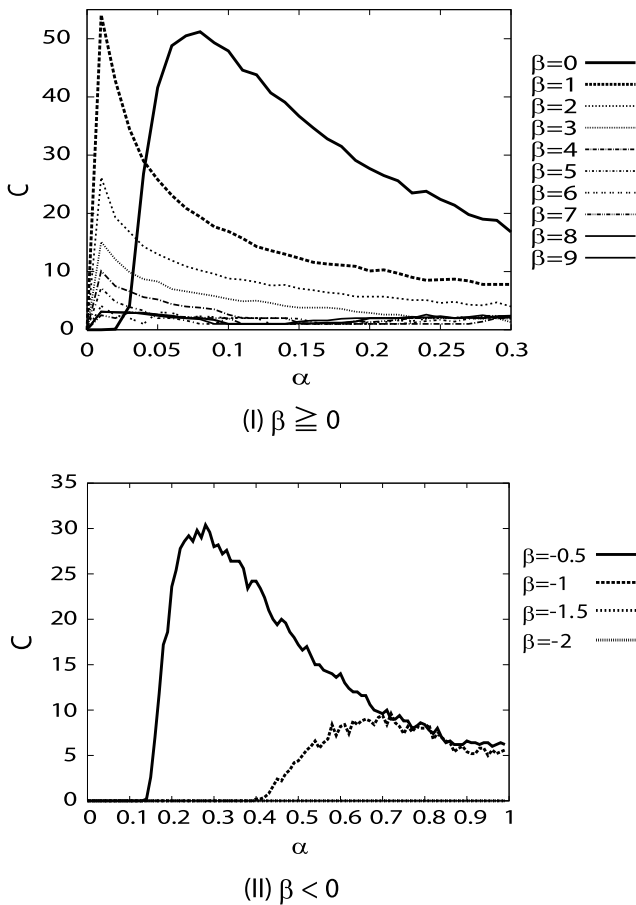


Fig. 3 Overlap  $m^\mu$  as a function of the learning step  $\mu$  in the  $\beta$ -th-order synaptic decay model other than the zero-order synaptic decay model and the forgetting model (i.e.,  $\beta \neq 0, 1$ ). Note that  $\mu = M$  represents the most recently stored pattern, and that the smaller  $\mu$  is, the older  $\xi^\mu$  is.  $N = M = 1,000$ . 1 sample. Horizontal dashed curve is the threshold for successful recalling  $m^\mu = 0.8$ . (I)  $\beta = 2$ , (II)  $\beta = 3$ , (III)  $\beta = -1$ , (IV)  $\beta = -2$ . The synaptic decay models stably retrieve the recent memories except for the case (IV). When  $\beta = -2$ , the network cannot retrieve any stored patterns.

ures (I)-(IV), only the recent memories ( $\mu > 980$ ) can be stably retrieved ( $m^\mu \approx 1$ ) in the synaptic decay model except for  $\beta = -2$ . Figure 3 predicts a phase transition phenomenon depending on the learning step  $\mu$ . As  $\mu$  decreases, the memory retrieval state ( $m^\mu \approx 1$ ) becomes unstable, and the so-called spin-glass state ( $m^\mu \approx 0$ ) appears. On the other hand, when  $\beta = -2$ , the network cannot recall any stored patterns due to overloading.

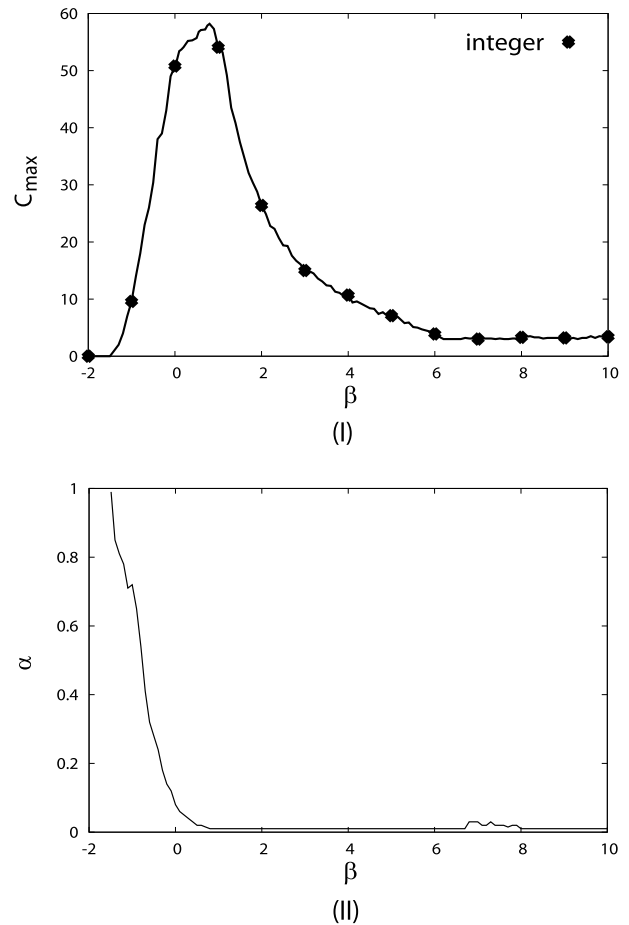
Figure 3 also shows the number of retrievable patterns depends on the decay coefficient  $\alpha$  in every synaptic decay model. Figure 4 illustrates the storage capacity  $C$  of each synaptic decay model as a function of  $\alpha$  for each degree of the synaptic decay term  $\beta$ . As shown in Fig. 4 (I), the zero-order decay model ( $\beta = 0$ ) and the forgetting model ( $\beta = 1$ ) have a remarkable memory retrieval performance compared with the other  $\beta$ -th-order decay



**Fig. 4** Storage capacity  $C$  as a function of the decay coefficient  $\alpha$  at each  $\beta$ -th-order synaptic decay model. (I)  $\beta = 0, 1, \dots, 9$ . (II)  $\beta = -2, -1.5, -1, -0.5$ . Each curve shows the average of 10 samples.  $N = M = 1,000$ . A point at the intersection of the abscissa with each curve corresponds to the minimum  $\alpha$  to avoid overloading ( $C = 0$ ). A peak of each curve corresponds to the maximal storage capacity  $C_{\max}$  of the synaptic decay model at each decay order  $\beta$ .

models. An increase in the decay order  $\beta$  degrades the network retrieval performance. When  $\beta > 6$ , there is little difference in the memory properties between the synaptic decay models. Moreover, when the decay order is negative ( $\beta < 0$ ), the increase in the absolute value  $|\beta|$  causes a dramatic drop in the network retrieval performance as shown in Fig. 4 (II). When  $\beta \leq -1.5$ , the storage capacity  $C$  becomes zero for any decay coefficient  $\alpha$ . In other words, the network which has the synaptic decay order less than or equal to  $-1.5$  cannot avoid overloading.

Finally, we show the optimal synaptic decay order to maximize the network retrieval performance. We here define the maximal storage capacity of the  $\beta$ -th-order synaptic decay model,  $C_{\max}$ , as the maximal value of  $C$  at each  $\beta$ . **Figure 5** (I) illustrates the maximal storage capacity  $C_{\max}$  as a function of the decay order  $\beta$ . The optimal decay order to maximize  $C_{\max}$  is  $\beta = 0.8$ . When  $\beta$  is an integer, the forgetting model, whose decay order is  $\beta = 1$ , has the largest  $C_{\max}$ . The zero-order decay model ( $\beta = 0$ ) has the second largest  $C_{\max}$ . As the absolute value  $|\beta|$  increases,  $C_{\max}$  decreases. When  $\beta \leq -1.5$ , the maximal storage capacity becomes zero, i.e., the network cannot avoid overloading ( $C_{\max} = 0$ ). When  $\beta \geq 6$ , on the other hand, the maximal storage capacity is maintained approximately constant:  $C_{\max} \approx 3$ . **Figure 5** (II) shows the relation between  $\alpha$  at  $C_{\max}$  and  $\beta$ . When  $\beta > 0.8$ , the optimal synaptic de-



**Fig. 5** (I) Maximal storage capacity  $C_{\max}$  as a function of the synaptic decay order  $\beta$ . The curve plots the average of 50 samples.  $N = M = 1,000$ . Dots represent  $C_{\max}$  when  $\beta$  is an integer value. (II) The relation between the synaptic decay coefficient  $\alpha$  at  $C_{\max}$  and the decay order  $\beta$ .

coefficient  $\alpha$  to maximize the storage capacity is almost 0.01. The  $\alpha$  at  $C_{\max}$  approaches 1 as  $\beta$  decreases. Note that  $\alpha = 1$  represents erasing a whole memory pattern at each learning step for embedding the new pattern.

#### 4. Conclusion

We investigated the properties of the associative memory model with the  $\beta$ -th-order synaptic decay. We demonstrated the relationship between the storage capacity  $C$  and the synaptic decay coefficient  $\alpha$  at each decay order  $\beta$ . In the cases of  $\beta = 0, 1$ , the obtained results are completely equal to those of the zero-order decay model [9] and the forgetting model [7], respectively. Numerical simulations revealed that the synaptic decay order model can stably retrieve the recently stored patterns while avoiding overloading when  $\beta > -1.5$ . When  $\beta \geq 6$ , the properties of the associative memory models become almost the same.

Furthermore, we searched the optimal decay model that has the best performance of memory retrieval by varying the decay order  $\beta$ . We evaluate the network retrieval performance using the maximal storage capacity  $C_{\max}$ . The optimal decay order to maximize  $C_{\max}$  is  $\beta = 0.8$ . When  $\beta$  is an integer, the forgetting model ( $\beta = 1$ ) has the largest  $C_{\max}$ , and the zero-order synaptic decay model ( $\beta = 0$ ) has the second largest  $C_{\max}$ . We also carried out computer simulations for  $N = 2,000$ , and got the same results as

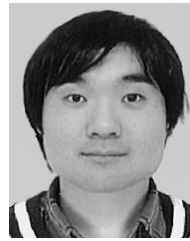
those for  $N = 1,000$ .

We here investigate only the maximal retrieval performance of the  $\beta$ -th-order synaptic decay model when an originally stored pattern was given as the initial retrieval state. In other words, we just study how many and how old patterns are unforgotten in the network with synaptic decay. In the associative memory model, starting with an initial state close to a stored pattern, the retrieval state monotonically approaches the stored one. However, starting with an initial state which is not so close to a memorized one, the state once approaches it but then goes away from it. This suggests that there is the basin of the attractor of a stored pattern [1], [12]. We will derive this critical initial distance which guarantees the success of recalling at each  $\beta$ -th-order decay model.

**Acknowledgments** This work was supported by KAKENHI 24-9230.

### References

- [1] Amari, S. and Maginu, K.: Statistical neurodynamics of associative memory, *Neural Networks*, Vol.1, pp.63–73 (1988).
- [2] Amit, D. J., Gutfreund, H. and Sompolinsky, H.: Storing infinite numbers of patterns in a spin-glass model of neural networks, *Physical Review Letters*, Vol.55, pp.1530–1533 (1985).
- [3] Date, A. and Kurata, K.: A property of neural networks of associative memory models with replacing units, *Artificial Life and Robotics*, Vol.12, pp.291–294 (2008).
- [4] Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities, *Proc. National Academy of Sciences*, Vol.79, pp.2554–2558 (1982).
- [5] Kimoto, T. and Okada, M.: Sparsely encoded associative memory with forgetting process, *IEICE Trans. Inf. Syst.*, Vol.E85-D, pp.1938–1945 (2002).
- [6] Ishikawa, M.: Structural learning with forgetting, *Neural Networks*, Vol.9, pp.509–521 (1996).
- [7] Mézard, M., Nadal, J.P. and Toulouse, G.: Solvable models of working memories, *Journal of Physique*, Vol.47, pp.1457–1462 (1986).
- [8] Mimura, K., Okada, M. and Kurata, K.: Associative memory model with forgetting process using nonmonotonic neurons, *IEICE Trans. Inf. Syst.*, Vol.E81-D, pp.1298–1304 (1998).
- [9] Miyata, R., Aonishi, T., Tsuzurugi, J. and Kurata, K.: Properties of Hopfield model with the zero-order synaptic decay, *Artificial Life and Robotics*, Vol.17, pp.163–167 (2012).
- [10] Miyata, R., Kurata, K. and Aonishi, T.: Sparsely encoded Hopfield model with unit replacement, *IEICE Trans. Inf. Syst.*, Vol.E95-D, pp.2124–2132 (2012).
- [11] Nadal, J.P., Toulouse, G., Changeux, J.P. and Dehaene: Network of formal neurons and memory palimpsests, *Europhysics Letters*, Vol.1, pp.535–542 (1986).
- [12] Okada, M., Mimura, K. and Kurata, K.: Associative memory with forgetting process -analysis by statistical neurodynamics, *IEICE Trans. Inf. Syst.*, Vol.J77-D-II, pp.1178–1180 (1994).
- [13] Okada, M.: A hierarchy of macrodynamical equations for associative memory, *Neural Networks*, Vol.8, pp.833–838 (1995).
- [14] Okada, M.: Notions of associative memory and sparse coding, *Neural Network*, Vol.9, pp.1429–1458 (1996).
- [15] Parisi, G.: Asymmetric neural networks and the process of learning, *Journal of Physics A*, Vol.19, pp.675–680 (1986).
- [16] Shiino, M. and Fukai, T.: Self-consistent signal-to-noise analysis and its application to analogue neural networks with asymmetric connections, *Journal of Physics A*, Vol.25, pp.375–381 (1992).
- [17] Tominaga-Yoshino, K., Kondo, S., Tamotsu, S. and Ogura, A.: Reti-tive activation of protein kinase A induces slow and persistent synap-togenesis in cultured hippocampus, *Neuroscience Research*, Vol.44, pp.357–367 (2002).
- [18] Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J. and Kasai, H.: Principles of Long-Term Dynamics of Dendritic Spines, *The Journal of Neuroscience*, Vol.28, pp.13592–13608 (2008).



**Ryota Miyata** received his B.E. and M.E. degrees from University of the Ryukyus in 2009 and 2011, respectively. He is currently a doctoral student at Inter-disciplinary Graduate School of Science and Engineering, Tokyo Institute of Tech-nology, and a JSPS research fellow DC. His research interests include theoretical neuroscience and machine learning.



**Toru Aonishi** received his B.E. degree from Kyusyu Institute of Technology in 1993, and his M.E. and D.E. degrees from Osaka University in 1995 and 1998, re-spectively. He was a researcher at RIKEN Brain Science Institute from 1998 to 2004. He is currently an associate professor at Interdisciplinary Graduate School of Sci-ence and Engineering, Tokyo Institute of Technology. His re-search focuses on non-linear dynamics, statistical mechanics and theoretical neuroscience.



**Jun Tsuzurugi** received his B.S. degree from Tokyo Metropolitan University in 1998, and his M.E. degree from Kyushu University in 2002, and his Ph.D. degrees from Nara Institute of Science and Tech-nology and Kyoto University in 2002 and 2013, respectively. From 2002 to 2003, he was an assistant at Fukui National In-stitute of Technology. He is currently a lecturer at Okayama Uni-versity of Science.



**Koji Kurata** received his B.S., M.S., and Ph.D. degrees from The University of Tokyo in 1981, 1983 and 1990, respec-tively. From 1984 to 1990, he was a research associate at The University of Tokyo. From 1990 to 2000, he was an as-sociate professor at Osaka University. He is currently a professor at the Faculty of Engineering, University of the Ryukyus. His research interests in-clude computational aspects of neural networks, especially those of topological mapping organization. He is also interested in pat-tern formations in homogeneous fields.